

Assessing the impact of academic literacy interventions in higher education: an evaluation design

I Fouché
24910503

Thesis submitted for the degree *Doctor Philosophiae* in *Linguistics and Literary Theory* at the Potchefstroom Campus of the North-West University

Supervisor: Prof. T van Dyk
Co-supervisor: Dr G Butler

May 2016

ACKNOWLEDGEMENTS

I would like to give thanks to the following people, without whom this thesis would not have been possible.

To my supervisors, Prof. Tobie van Dyk and Dr Gustav Butler, whose guidance and mentorship have led me to constantly re-evaluate my assumptions, gain new insights, and ultimately become a better academic. Thank you.

To my family, whose love and support means everything to me. You have shaped me into the person I am today, and without you, none of this would have been possible.

To my friends – without your encouragement, sympathy, and support, this journey would have been infinitely more difficult. In particular, I would like to thank Gerda Gilson and Terrence Carney, both of whom have walked this journey before me, for supporting me throughout the past years by being sounding boards, sharing copious bottles of wine, and offering encouragement on a weekly basis – your friendship is invaluable.

To colleagues at both the North-West University and the University of Pretoria who have helped me in countless ways over the past three years: Dr Henk Louw, Dr Elsa Meihuizen and Ms Leischa Lazenby, for your assistance in the data-gathering process; Dr Suria Ellis and Ms Erika Fourie, for assisting me with all the statistical analyses for this study; and to the countless other colleagues who have offered support, be it administrative, academic or otherwise.

To all the students and academic literacy specialists who participated in this study.

Soli Deo gloria.

Pretoria

May, 2016

ABSTRACT

Responsibly designing academic literacy interventions is becoming increasingly important in a higher education environment where resources for student development are scarce. Providing proof of the effectiveness of such interventions is equally important – academic literacy specialists must show that what they are doing has a significant and meaningful impact on student success. If certain aspects of an intervention were shown not to work optimally, these should be addressed. This cycle of providing evidence of an intervention’s successes and shortcomings, and addressing any such shortcomings, is the goal of impact measurement. However, very few studies have attempted to comprehensively measure the impact of academic literacy interventions, probably because measuring impact in the social sciences is a challenging undertaking.

The goal of this study was to develop an evaluation design that could be used to effectively and responsibly measure the impact of a wide range of academic literacy interventions. The first step in developing this evaluation design was to survey the available literature on impact measurement, specifically in the field of academic literacy, so as to propose a theoretically sound design and accompanying research instruments. Through a process of inquiry, piloting and critical reflection, several of these research instruments were adapted to make them applicable to the wide variety of academic literacy interventions that are presented in the South African context.

After having proposed an initial evaluation design, this design was verified and validated by i) implementing it in measuring the impact of an academic literacy course in the South African context and ii) obtaining feedback from academic literacy specialists from across the country on how the design could be further improved to suit their respective contexts. After having critically reflected on the implementation process, and after having analysed responses from academic literacy specialists, a revised evaluation design and accompanying research instruments were proposed. These should assist researchers in comprehensively and responsibly measuring the impact of a wide range of academic literacy interventions, and consequently benefit the field of academic literacy as a whole.

KEY TERMS

Academic literacy; impact; effect; programme evaluation; evaluation design; student success; language testing

TABLE OF CONTENTS

Acknowledgements	i
Abstract	ii
List of figures	ix
List of tables	x
Appendices	xi
CHAPTER 1 INTRODUCTION.....	1
1.1 Contextualisation.....	1
1.2 Problem statement	2
1.3 Research design.....	2
1.4 Research questions	6
1.5 Objectives.....	6
1.6 Instruments	6
1.7 Parameters of the study	7
1.8 Overview and structure of the study	8
1.9 Ethical considerations	10
1.10 Contribution of the study	10
CHAPTER 2 LITERATURE REVIEW	12
2.1 Introduction	12
2.2 Key concepts	13
2.2.1 Impact	13
2.2.2 Success.....	14
2.2.3 Stakeholders.....	15
2.2.4 Language assessment.....	16
2.2.5 Programme evaluation	16
2.2.6 Evaluation design.....	17
2.3 Underpreparedness for higher education studies	17

2.4	Academic literacy.....	19
2.4.1	Defining ‘academic literacy’	19
2.4.2	A conceptual framework for academic literacy.....	22
2.4.3	Academic literacy levels as an indicator of academic success	27
2.4.4	Identifying at-risk students	29
2.4.5	Academic literacy interventions	31
2.5	Programme evaluation	33
2.5.1	Paradigms in programme evaluation	36
2.5.2	Evaluating impact	37
2.5.2.1	Considerations when evaluating impact	38
2.5.2.2	Difficulties when measuring the impact of an intervention.....	40
2.5.3	Methodology for evaluating impact.....	42
2.5.3.1	Types of assessment in academic literacy courses	45
2.5.3.1.1	Valid and reliable academic literacy tests	45
2.5.3.1.2	Assessing writing assignments	49
2.5.3.1.3	Perceptions of impact	53
2.6	Conclusion.....	58
CHAPTER 3	TOWARDS IMPACT MEASUREMENT: AN OVERVIEW OF APPROACHES FOR ASSESSING THE IMPACT OF ACADEMIC LITERACY ABILITIES (ARTICLE 1)	60
3.1	Prelude to Chapter 3.....	60
3.2	Introduction	60
3.3	Previous studies on the impact of academic literacy interventions.....	63
3.4	Discussion and conclusion	76
CHAPTER 4	IMPACT MEASUREMENT: TOWARDS CREATING A FLEXIBLE EVALUATION DESIGN FOR ACADEMIC LITERACY INTERVENTIONS (ARTICLE 2).....	80
4.1	Prelude to Chapter 4.....	80
4.2	Introduction	80
4.3	Aim of the chapter.....	82

4.4	Instruments for assessing the impact of academic literacy interventions	82
4.4.1	Assessing improvement in students' academic literacy levels	84
4.4.1.1	Instrument 1: Reliable, validated generic academic literacy tests	85
4.4.1.2	Instrument 2: Language for specific academic purposes (LSAP) tests	89
4.4.1.3	Instrument 3: Writing assignments	92
4.4.2	Assessing the extent to which academic literacy abilities are necessary and applied in students' content subjects.....	96
4.4.2.1	Instrument 4: Determining academic literacy ability acquisition and academic literacy needs in content subjects – student questionnaire.....	97
4.4.2.2	Instrument 5: Determining academic literacy needs in content subjects – lecturer questionnaire.....	100
4.4.2.3	Instrument 6: Content analysis of study material	101
4.4.2.4	Instrument 7: Correlating academic literacy results with other variables	102
4.4.2.5	Instrument 8: Qualitative feedback from primary stakeholders	103
4.5	An evaluation design for academic literacy interventions	105
4.6	Conclusion.....	108

CHAPTER 5 IMPACT MEASUREMENT: QUANTITATIVELY DETERMINING THE IMPROVEMENT IN STUDENTS' ACADEMIC LITERACY LEVELS AT A SOUTH AFRICAN UNIVERSITY (ARTICLE 3)110

5.1	Prelude to Chapter 5.....	110
5.2	Introduction	111
5.3	Background	112
5.4	Design and method.....	114
5.5	Results and discussion.....	115
5.5.1	Instrument 1: Using a generic academic literacy test	115
5.5.1.1	Discussion	118
5.5.2	Instrument 2: Using a generic extended writing assignment (assessed by means of a rubric)	120
5.5.2.1	Discussion	122

5.5.3 Instrument 3: Quantitatively assessing an extended writing assignment	124
5.5.3.1 Discussion	128
5.5.4 Instrument 4: Correlating academic literacy achievements with other variables	130
5.5.4.1 Discussion	130
5.6 Summary of findings	132
5.7 Conclusion.....	134

CHAPTER 6 AN “ENLIGHTENING COURSE THAT EMPOWERS FIRST YEARS”?: A HOLISTIC ASSESSMENT OF THE IMPACT OF A FIRST-YEAR ACADEMIC LITERACY COURSE (ARTICLE 4).....136

6.1 Prelude to Chapter 6.....	136
6.2 Introduction	137
6.3 Background	138
6.4 Design and method.....	139
6.5 Results and discussion.....	139
6.5.1 Instrument 1: Using a generic academic literacy test	139
6.5.1.1 Discussion	140
6.5.2 Instrument 2: Using a generic extended writing assignment (assessed by means of a rubric)	141
6.5.2.1 Discussion	143
6.5.3 Instrument 3: Quantitatively assessing an extended writing assignment .	145
6.5.3.1 Discussion	146
6.5.4 Instrument 4: Student questionnaire	147
6.5.4.1 Discussion	149
6.5.5 Instrument 4: Correlating academic literacy achievements with other variables	155
6.5.5.1 Discussion	155
6.6 Conclusion.....	156

CHAPTER 7	IMPACT MEASUREMENT: REFINING AN EVALUATION DESIGN FOR ACADEMIC LITERACY INTERVENTIONS THROUGH SELF-REFLECTION AND FEEDBACK (ARTICLE 5)	159
7.1	Prelude to Chapter 7	159
7.2	Introduction	159
7.3	Background to the development of the evaluation design	160
7.4	A summary of the originally proposed evaluation design	162
7.5	Reflecting on the challenges experienced in implementing the design	165
7.5.1	The insider / outsider dilemma	166
7.5.2	The use of an academic literacy test as pre- and post-test	167
7.5.3	The use of an extended writing assignment	167
7.5.4	The use of questionnaires	169
7.6	Feedback from academic literacy specialists	170
7.6.1	Quantitative feedback from academic literacy specialists	171
7.6.2	Qualitative feedback from academic literacy specialists	174
7.7	Revised evaluation design	182
7.8	Conclusion	186
CHAPTER 8	CONCLUSION AND RECOMMENDATIONS	188
8.1	Introduction	188
8.2	Summary of the study	188
8.3	Conclusions and contribution of the study	191
8.4	Limitations	192
8.5	Recommendations for further research	194
8.6	Conclusion	196
	BIBLIOGRAPHY	197

List of figures

Figure 2.1: Linguistic, cognitive and sociocultural dimensions of literacy	23
Figure 2.2: An interactional framework for language use	25
Figure 2.3: The Bachman and Palmer construct of language ability	26
Figure 4.1: Characteristics of an effective writing assessment	94
Figure 4.2: Proposed evaluation design for academic literacy interventions.....	107
Figure 5.1: Questions categorised based on abilities assessed.....	117
Figure 5.2: Questions categorised based on abilities assessed – re-categorised	117
Figure 5.3: Categories used to calculate accuracy in written assignments	126
Figure 7.1: Evaluation design for academic literacy interventions	184

List of tables

Table 2.1:	Academic literacy abilities assessed in the PTEEP	48
Table 5.1:	Improvement between the pre- and post-intervention academic literacy tests (1 st semester)	119
Table 5.2:	Improvement between pre- and post-intervention writing assignments (assessed by means of rubric) (1 st semester)	123
Table 5.3:	Reduction of language mistakes between pre- and post-assignments (1 st semester)	128
Table 5.4:	Increase in academic vocabulary usage pre- and post-assignments (1 st semester)	129
Table 5.5:	Non-parametric correlations using Spearman's rho (1 st semester)	131
Table 6.1:	Improvement between the pre- and post-intervention academic literacy tests (2 nd semester)	141
Table 6.2:	Improvement between the pre- and post-intervention writing assignments (assessed by means of rubric) (2 nd semester)	144
Table 6.3:	Reduction in language mistakes between pre- and post-assignments (2 nd semester)	146
Table 6.4:	Increase in academic vocabulary usage between pre- and post-assignments (2 nd semester)	147
Table 6.5:	Importance of abilities in content subjects, according to students (figures expressed as percentages)	150
Table 6.6:	Extent to which abilities are addressed in academic literacy course (figures expressed as percentages)	152
Table 6.7:	Non-parametric correlations using Spearman's rho (2 nd semester)	156
Table 7.1:	Academic literacy specialists' perceptions of usefulness of instruments	172
Table 7.2:	Academic literacy specialists' perceptions of usefulness of instruments, categorised into types of interventions	173

Appendices

Appendix A: Rubric for assessing writing assignments	210
Appendix B: The C3-model for evaluation of text quality	211
Appendix C: Storch and Tapper quantitative analysis	212
Appendix D: Need-press questionnaire on academic literacy abilities.....	214
Appendix E: Lecturer questionnaire on academic literacy abilities.....	217
Appendix F: Revised need-press questionnaire on academic literacy abilities	219
Appendix G: Coordinator questionnaire	221
Appendix H: Revised lecturer questionnaire on academic literacy abilities	222
Appendix I: Revised writing rubric	223
Appendix J: Rubric for oral presentations	225

CHAPTER 1

INTRODUCTION

1.1 Contextualisation

It is generally acknowledged that the South African secondary education system no longer sufficiently prepares students for higher education studies (Shay, 2015:431; Bharuthram, 2012:206; Smit, 2012:369; Van Dyk, Zybrands, Cillié & Coetzee, 2009:333; Higher Education South Africa, 2008:3). University throughput rates are directly affected by this underpreparedness: 73% of students do not complete their degrees in regulation time, while only 45% of students graduate at all (Scott, Ndebele, Badsha, Figaji, Gevers & Pityana, 2013:43). Low academic literacy levels (see Section 2.4.1) play a major role in the fact that such a low percentage of students are able to complete their degrees – either in the prescribed time, or at all (see Terraschke & Wahid, 2011:173; Davies, 2010:xi; Defazio, Jones, Tennant & Hooke, 2010:34; Van Dyk *et al.*, 2009:333; Archer, 2008:248; Holtzman, Elliot, Biber & Sanders, 2005:285; Stephen, Welman & Jordaan, 2004:42-51; Leibowitz, 2001:44). The relationship between students' academic literacy levels and the likelihood of them succeeding in higher education is further explored in Section 2.4.3.

From the above, it can be deduced that many students are in need of academic literacy support. However, effective academic literacy support often depends on the availability of significant resources – something that is becoming increasingly scarce at universities (see, for example, Hornsby & Osman, 2014:712-713; Kwiek, Lebeau & Brown, 2014:6; Feast, 2002:70; Leibowitz, 2001:4-5). Reduced financial resources translate into fewer lecturers, markers, venues, etc. Having to justify the necessity of drawing on these limited resources might be the reason why universities have been increasingly focusing on the efficiency and quality of programmes, whereas previously the focus has been almost exclusively on helping students access universities (Yeld, 2010:26). In so far as academic literacy interventions are concerned, these interventions need to prove that they meaningfully contribute to student success. This study accepts Jansen's (2004:3) view of student success which focuses on the cognitive and associated abilities that enable each person to master academic information so as to proceed to the next year of studies.

Academic literacy programmes therefore need to show how they equip students with the abilities that students will utilise to ultimately succeed in their studies.

1.2 Problem statement

Universities have an ethical obligation to put measures in place that aim to help students overcome barriers to success. One type of intervention that universities have the power to put in place is responsibly designed academic literacy interventions, i.e. interventions that have impact. Higher education institutions, and in particular the various departments, units and centres that offer academic literacy interventions, are responsible for ensuring that these interventions have the highest impact that they can possibly have (cf. Song, 2006:422). This means that the impact of academic literacy interventions on students' academic success must be comprehensively and effectively measured.

As the review in Chapter 3 shows, very few studies attempt to assess the impact of academic literacy interventions. Those that do follow varied approaches and usually investigate only one or two aspects of an intervention's impact. Several questions often remain unanswered: What contributes to the impact of an intervention? How can this be measured? What is the long-term effect of the intervention? (Van Dyk, Cillié, Coetzee, Ross & Zybrands, 2011a:493). Some studies, like that done by Archer (2008:248), attempt to answer as many of these questions as possible by using a “multi-faceted approach [that enables] a holistic and contextualized picture (...) to emerge”.

Thus far, however, no study has proposed an evaluation design that would systematically assess the impact of an academic literacy intervention on student success. The current study intends to address this gap by developing a multi-faceted evaluation design for measuring the impact of academic literacy interventions on student success.

1.3 Research design

Lynch (1996:156) defines a design as “a methodological strategy for program evaluation”, including “methods and sources of data collection” (Bamberger, Rugh & Mabry, 2012:613). The proposed evaluation design fits this description, and allows

researchers to follow Bachman and Palmer's (2010:20) advice that assessment¹ (for example that of an academic literacy intervention) must be systematic, "designed and carried out according to clearly defined procedures that are methodical and open to scrutiny", and substantively grounded (i.e. the assessment must be based on verifiable areas of content, for example, practice that is currently accepted in the relevant field of study).

A strength of the evaluation design proposed in this study is that it was constructed through a participatory process. Several guidelines mentioned by Alderson and Scott (1992:39-41) were followed in the process of creating this evaluation design. Firstly, it was designed by drawing on the experience of other researchers through a perusal of the literature (see Chapters 2, 3 and 4). Secondly, the input of other academic literacy specialists was incorporated in the construction and testing of the instruments (see Chapters 4 and 7). In the validation process (see Chapters 5 and 6), academic literacy lecturers at the North-West University assisted in collecting the data, and the expertise of a statistician was obtained in the analysis of the data. I further relied on my two promoters as well as peer reviewers in the drafting and finalisation of this final document.

One of the cornerstones of the proposed evaluation design is that data must be triangulated. Cheetham, Fuller, Mcivor and Petch (1992:10) argue that "effectiveness derives from a variety of perspectives and assumptions". A researcher must consider a variety of sources and perspectives on the impact of an intervention to truly obtain a holistic, comprehensive view thereof, contributing to the validity of the study. Moreover, there is a possibility that one technique or source might be inherently biased; using a range of sources, however, could eliminate such potential bias (Lynch, 1996:60). Patton (2002:93) states that the triangulation of data sources as well as analytical perspectives are necessary to increase the credibility and the accuracy of findings in a postpositivist framework. It is specifically useful to include both quantitative and qualitative components in any evaluation study (Lynch, 1992:94). Ross (2009:776) agrees: "A mixed mode approach is more likely to provide evidence that there has been a program effect".

¹ Though their advice refers to testing, it is equally applicable to a broader view of assessment, such as determining the impact of academic literacy programmes.

The evaluation design proposed in this study (see Chapters 4 and 7) draws on qualitative and quantitative feedback from a variety of primary and secondary stakeholders – specifically students, academic literacy lecturers and content-subject lecturers. Using mixed methods also enables the evaluator to answer both the “what” questions (for example, what abilities improved the most) that are generally answered by quantitative data, as well as the “how” and “why” questions (for example, how the course addresses students’ needs) that are generally answered by means of qualitative data (Dudley-Evans, 1995:128). Further, the evaluation design’s recommendation that multiple measures be used to assess impact also addresses the problem of many impact studies being limited to selective viewpoints, for example “determining the satisfaction of participants, rather than anything beyond that” (Cilliers & Herman, 2010:254) (also see Section 2.5.2). Using a variety of research instruments strengthens the research design and makes qualified conclusions possible (cf. Lynch, 1996:74).

The instruments recommended in the evaluation design proposed in the current study address the three dimensions of literacy proposed by Kern (2000:16-17, 38), namely sociocultural, linguistic and cognitive/metacognitive dimensions (see Section 2.4.2 and Chapter 7). The sociocultural dimension is addressed by, for example, considering the general academic and subject-specific conventions as well as vocabulary that are used in academic discourse. The linguistic dimension is addressed by considering students’ ability to use lexical, morphological, syntactic and semantic knowledge, to structure their writing appropriately and to link ideas effectively. The cognitive/metacognitive dimension is addressed by, for example, assessing students’ ability to construct an argument or synthesise information.

The proposed evaluation design should assist the researcher in adhering to four general standards of evaluation that are generally agreed upon, namely utility, feasibility, propriety and accuracy (see Section 2.6.3) (Joint Committee on Standards for Educational Evaluation, 2014; Newcomer, Hatry & Wholey, 2010:7). The selection of research instruments can, for example, assist evaluators when they need to consider the variety of primary and secondary stakeholders that should be consulted (i.e. utility) and guide them in using research instruments that are likely to produce valid and reliable results (i.e. accuracy). Using valid and reliable instruments that are widely used, and following the guidelines of studies which have used these instruments, could contribute

to ethical evaluation (i.e. propriety). Keeping in mind the guidelines suggested for the revised evaluation design (see Chapter 7) would also assist researchers in ensuring that their evaluation plans are workable in real world settings (i.e. feasibility).

Despite having followed the above guidelines in the creation of the evaluation design, it must be remembered that doing research on impact, especially in the field of education, is generally a challenging undertaking. One reason for this is that the many factors involved in student success are difficult to control, and isolating any of these factors so as to establish the impact of a specific element will be difficult, if not impossible (Howes, 2003:148). Research into academic literacy interventions is no exception. In this field, factors that must be kept in mind, but which cannot necessarily be controlled, include the exposure to academic writing that students receive in their mainstream subjects as well as academic literacy-related feedback from sources outside the academic literacy classroom. A further challenge is that it is usually not feasible to form control groups at universities where academic literacy interventions are part of almost all students' curricula, as is the case at most South African universities. In such cases, creating artificial control groups might bring along serious ethical considerations. To meaningfully determine the impact of an academic literacy intervention, therefore, alternative methods need to be used.

Quasi-experimental designs must be considered in cases where no control groups are available. When using these designs, it might not be possible to make the same, unambiguous causal claims regarding the effect of interventions than might be possible with traditional experimental designs (Lynch, 1996:73-74). However, stronger causal claims can be made by, for example, using a pre-test/post-test design that uses multiple measures, in combination with instruments such as questionnaires concerning perceptions of the programme by teachers and students (Lynch, 2003:25; Jick, 1979:602). As argued in Section 2.5.3, triangulating different types of evidence in a research design can be a useful validity check where traditional experimental designs are not feasible. Furthermore, supplementing the pre-test/post-test design with multiple perspectives allows the researcher to depart from a purely positivist research design and to move towards a postpositivist, mixed-methods research design (Lynch, 2003:27), as discussed in Section 2.5.1. To strengthen the validity of findings in an evaluation process, Lynch (1996:60) states that the obvious strategy to use would be triangulation.

1.4 Research questions

This study intends to answer the following questions:

- Question 1: Which methods, and which combinations of methods, have been used to assess the impact of academic literacy interventions in the past?
- Question 2: What would an appropriate evaluation design to measure the impact of an academic literacy intervention, based on trends and suggestions in the literature, consist of?
- Question 3: Can this evaluation design be used effectively to assess the impact of a specific academic literacy intervention?
- Question 4: To which extent can the evaluation design be generalised to other settings, so as to assess the impact of a variety of academic literacy courses in a variety of contexts?

1.5 Objectives

In this study, I aim to:

- Objective 1: determine which methods have been used in existing studies (as reported on in the literature) to assess the impact of academic literacy interventions;
- Objective 2: develop an evaluation design, based on trends and suggestions in the literature, to assess the impact of an academic literacy intervention on student success;
- Objective 3: use this evaluation design to assess the impact of a specific academic literacy intervention; and
- Objective 4: determine to which degree this evaluation design can be generalised to other settings, so as to assess the impact of a variety of academic literacy interventions in a variety of contexts.

1.6 Instruments

The aim of this study was to create an evaluation design consisting of various instruments which could be used to evaluate the impact of a variety of academic literacy

interventions. During the verification and validation process of this study, several existing instruments were adapted to suit the purposes of the study. A detailed discussion of these instruments can be found in Sections 4.4 and 7.7 of this study.

Instruments that were administered or applied to staff members, students, or student assignments in the course of the current study were as follows:

- a rubric for assessing writing assignments (Appendix A);
- the Storch and Tapper quantitative analysis (Appendix C);
- the need-press questionnaire on academic literacy abilities (Appendix D);
- the Test of Academic Literacy Levels (See Section 2.5.3.1.1 for a more detailed discussion of this test); and
- a coordinator questionnaire (Appendix G).

1.7 Parameters of the study

The parameters of the study are as follows:

- although the evaluation design might be applicable to other contexts, it was verified and validated in the South African context – future research may determine whether it could be applicable to other contexts;
- during the implementation phase of this research design (see Chapters 5 and 6), the aim was to determine the short-term impact of one academic literacy course – no research was conducted beyond students' first year of study;
- only the English versions of the academic literacy courses (AGLE111 and AGLE121) at the North-West University's Potchefstroom Campus were used when the evaluation design was practically implemented (Chapters 5 and 6); and
- although the study falls within the field of programme evaluation, the focus was exclusively on measuring the impact of academic literacy interventions.

Before the structure of this study is discussed in more detail, one important aspect should be kept in mind. It is important to note that since academic literacy interventions vary vastly (for example in terms of content and purpose), flexibility would have to be a key characteristic of any evaluation design which aims to assess the impact of such

interventions. Any researcher who wishes to use the evaluation design that is proposed at the end of this study would have to be able to adapt the research instruments in the evaluation design in a manner that would most effectively assess the impact of the academic literacy intervention in question, based on its context and specific needs. This has been taken into consideration as far as possible in the proposed evaluation design.

1.8 Overview and structure of the study

This study is an article-based PhD. Thus, five articles form the core of the current study. To avoid confusion, the articles are referred to by their corresponding chapter numbers. Thus, throughout this study, Article 1 is referred to as Chapter 3, Article 2 as Chapter 4, Article 3 as Chapter 5, Article 4 as Chapter 6, and Article 5 as Chapter 7. The five articles (Chapters 3 to 7) are preceded by this introduction (Chapter 1) and a literature review (Chapter 2). The articles are followed by a conclusion (Chapter 8). The current section outlines Chapters 2 to 8 and, where applicable, briefly describes the methodology employed. This is elaborated upon in the chapters themselves.

Chapter 2 provides an in-depth review of the literature. This literature review aims to elaborate on concepts which, due to the nature of academic articles, could not be expanded on in the articles (Chapters 3 to 7) themselves.

The first of the five articles is contained in Chapter 3. The chapter provides a thorough overview of studies which have attempted to assess the impact of a variety of academic literacy interventions in the past. The strengths and weaknesses of approaches followed are discussed, and the main methods of assessing impact up to this point in the literature are summarised. This chapter is a vital foundation upon which the evaluation design proposed in Chapter 4 is based.

The second article is contained in Chapter 4. Chapter 4 builds on Chapter 3 in that a theoretical evaluation design is proposed based on the instruments which have been identified in the literature thus far. Weaknesses which were identified in Chapter 3 are addressed in Chapter 4 by providing guidelines from the literature for the use of each proposed instrument. Where possible, specific instruments (for example questionnaires and rubrics) are proposed and included as appendices.

Chapter 5 contains the third article around which this study is built. This chapter reports on the first part of a two-part process in which the proposed evaluation design was tested and verified by applying it to an academic literacy course at the North-West University's Potchefstroom Campus². In this chapter, the impact of the first module of the academic literacy course, Basic Skills in Academic Literacy (AGLE111), which is presented during the first semester of the year, is evaluated. Only empirical data is reported on in this chapter. This entails an investigation of whether students' academic literacy abilities had improved by considering a pre- and post- academic literacy test as well as examining whether, and to which degree, students' writing abilities improved by examining pre- and post- writing assignments. Students' writing assignments are assessed by means of two measures, namely a rubric as well as quantitative measures. Furthermore, students' results in the AGLE111 module are correlated to the results of their other subjects.

Chapter 6 (Article 4) reports on the second part of the two-part process of testing and verifying the proposed evaluation design. This chapter reports on the second module, Advanced Skills in Academic Literacy (AGLE121), which is presented during the second semester of the year, as well as the impact of both modules when seen as a unit. Students' results on an academic literacy pre- and post-test as well as the results with regard to their writing abilities (by considering pre- and post- writing assignments) are again reported on. In addition, students' perceptions of the year-long intervention in its totality are considered by analysing student questionnaires. Finally, students' results in the academic literacy module are correlated to the results of their mainstream modules.

The last of the five articles is contained in Chapter 7. This chapter discusses the validation of the evaluation design. This is done by consulting, by means of a questionnaire containing both closed and open-ended questions, a variety of academic literacy course-coordinators at various universities, with the purpose of determining whether the evaluation design could be applied to a variety of academic literacy courses in various settings. In addition to that, the researcher's own experience with applying the evaluation design to an academic literacy course is discussed, and strengths, weaknesses and limitations that emerged from this implementation are considered. Any further

² The Centre for Academic and Professional Language Practice at the North-West University's Potchefstroom campus is currently in the process of re-evaluating its academic literacy courses. An important aspect of this re-evaluation is to assess the impact of existing courses.

recommendations regarding the possible refinement of the evaluation design are made at this point. This chapter is thus both conceptual and empirical in nature.

Chapter 8 concludes this study by summarising its main findings, making conclusions based on the research conducted, indicating the contribution of the study, discussing its limitations, and making recommendations for future research.

1.9 Ethical considerations

Ethical clearance was obtained from the North-West University for all research instruments that were used in this study. Participants were given the opportunity to provide consent (in writing) to their data (test results, essays and survey responses) being used for research purposes, and only data of participants who provided such consent were used. All research instruments contained a disclaimer informing participants about the rationale for the study. It was made clear that participation was voluntary and that participants could opt out of the study at any time. Disclaimer forms were attached to all research instruments, making it clear that no students would be disadvantaged by participating in the study. At no point during this study are students or lecturers identified by name, student number or personnel number.

1.10 Contribution of the study

The chief contribution of the study is the development of a comprehensive, verified and validated (see Section 1.8) evaluation design for assessing the impact of academic literacy interventions. The outcomes of implementing such an evaluation design will be manifold. Firstly, it could be used to enable course developers to determine whether their academic literacy interventions have the desired impact, and to adjust their course content and approach accordingly; as De Klerk, Van Deventer and Van Schalkwyk (2006:149) state, “a study should not only be reflective, but should also provide guidelines and directives for restructuring the programmes to meet current challenges”. Where interventions are shown to have a meaningful impact, it will become easier to justify the costs and other resources necessary to run such interventions meaningfully. The course developers of various academic literacy interventions would be able to learn from each other’s failures and successes by considering the impact of their respective

interventions. In fact, successfully measuring the impact of academic literacy interventions would eventually strengthen academic literacy as a discipline.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

The gap between the secondary and higher education systems has steadily widened over the past years, to the extent that many students who enter higher education institutions are not sufficiently prepared for the demands of tertiary studies (Cliff, 2014:322; Department of Higher Education and Training, 2014:6; Butler, 2013:71; Van Dyk *et al.*, 2009:333; Higher Education South Africa, 2008:3). In spite of this, higher education institutions are under great pressure to increase throughput rates. This is because South Africa firstly has a very low higher education participation rate of approximately 17% (Department of Higher Education and Training, 2014:30) as opposed to a 70% participation rate in North America and Western Europe (Council on Higher Education, 2009:4), and secondly has very low throughput³ and pass rates (Yeld, 2010:26). At the same time, resources are becoming increasingly scarce in higher education institutions (see Section 1.1), resulting in institutions being increasingly hesitant to spend money on student support interventions without proof that such programmes have a significant impact on student success.

A low level of academic literacy is considered to be one of the major contributing factors to poor student throughput, not only in South Africa, but also internationally (Terraschke & Wahid, 2011:173; Davies, 2010:xi; Defazio *et al.*, 2010:34; Archer, 2008:248). Many studies (e.g. Petersen-Waughtal & Van Dyk, 2011; Van Dyk, 2010; Cliff & Hanslo, 2009; Weideman, 2009; Van Dyk & Weideman, 2004) have examined possible placement and diagnostic tests to identify students who are in need of additional academic literacy support during their first year of university studies. Furthermore, various studies (e.g. Wingate & Tribble, 2012; Parkinson, Jackson, Kirkwood & Padayachee, 2007; Lea, 2004) have described a range of academic literacy support programmes designed to address students' poor academic literacy levels. However, very few studies have attempted to measure the impact of such interventions on students' academic literacy levels.

³ Cliff and Hanslo (2009:265) define throughput as “the rate at which selected students graduate with a meaningful and portable qualification”.

The increased focus on throughput and pass rates, combined with the reality of a resource-scarce higher education environment, is possibly why higher education institutions are increasingly focusing on graduation rates, general efficiency and the quality of programmes (Yeld, 2010:26). Song (2006:422), for example, states that “[t]o justify the extra funding support, administrators of [academic literacy] programs are often under pressure to provide evidence of students’ superior academic performance over time”. In order for academic literacy programmes to justify their own existence, they need to prove that they have a real and significant impact on student success.

2.2 Key concepts

Before this problem is discussed in more detail, it might be worthwhile to first set the parameters for how certain key concepts are defined for the purpose of the current study. These key concepts are *impact*, *success*, *stakeholders*, *language assessment*, *programme evaluation* and *evaluation design*.

2.2.1 Impact

The current study treats the terms “impact”⁴ and “effect” synonymously. The Concise Oxford English Dictionary (2002:710) defines “impact” as “a marked effect or influence”. De Vos, Fouché, Strydom and Delpont (2011:459-460) define impact as follows: “A programme effect, or impact, refers to a change in the target population or social conditions that have been brought about by the programme – that is, a change that would not have occurred had the programme been absent”. De Graaff and Housen (2009:727) state that the term “effect” refers to an observable change in learners’ knowledge, disposition or behaviour, which can be ascribed to a specific instructional intervention (the authors do qualify that this might be in interaction with contextual variables). It follows that an intervention’s effectiveness depends on the extent to which

⁴ I take note of Bak’s (2004:22) opinion that the term “impact studies” should be avoided unless the researcher has many resources and “a clear, limited and easily controlled situation”, and that studies that do not fit these criteria should rather refer to the “consequences” of an intervention. However, since a large body of research already exists that does refer to the term “impact” rather than the arguably more ambiguous term “consequences”, this study will refer to “impact” and “effect” throughout, whilst acknowledging that assuming a “direct, linear causal connection” is likely problematic for the majority of academic literacy interventions, as is the case with most social sciences studies.

desired and actual outcomes intersect (De Graaff & Housen, 2009:727-728). It is, however, worthwhile to keep in mind this warning by Cheetham *et al.* (1992:9-10):

Despite much apparently straightforward use of the word, ‘effectiveness’ is not something which has an object-like reality ‘out there’ waiting to be observed and measured. Like any other data, empirical evidence about the effectiveness of (...) programmes is a product of data collection procedures and the assumptions on which they are based. The concept of effectiveness derives from particular ways of thinking and makes sense only in relation to its context. (...) The challenge is to arrive at working definitions of effectiveness in specific situations, and hence of methods of studying it, which do not permanently lose sight of its conceptual context.

Keeping the above argument in mind, impact (or effect) will, for the purposes of the current study, be seen as i) the observable improvement in academic literacy abilities between the onset and the completion of an academic literacy intervention, and ii) the extent to which these abilities are necessary and applied in students’ content subjects.

2.2.2 Success

“Success” can be interpreted in many ways. This study accepts Jansen’s (2004:3) view which focuses on the cognitive and associated abilities that enable each person to master academic information so as to proceed to the next year of studies. Cliff and Yeld (2006:20) focus on these cognitive and associated abilities, and view success as “the student having perceived the nature of the language and thinking demands placed upon him or her and having made appropriate responses to those demands”. Tinto and Pusser’s (2006:1) definition, in contrast, is similar to the second part of Jansen’s definition in that they view student success as the completion of a higher education qualification. However, their definition of “persistence” is perhaps closer to Jansen’s definition, in that they define persistence as “the enrolment of individuals over time”. It is interesting to note Engstrom and Tinto’s (2008:47) argument that academic preparation is the most important indicator for student success (though his focus is on low-income students, this argument can surely be made for all students). Van Dyk and Van de Poel (2013:45) add that other factors that influence success are “[s]tudents’ motivation and ability to adapt to new ways of pursuing, interpreting, organising, producing and communicating knowledge, and to get accustomed to the norms, standards, procedures, values and linguistic forms that constitute academic life”. The goal of this study is to create an evaluation design that could be used to determine the extent to which an academic

literacy intervention could assist students in becoming acculturated into the higher education environment through some of the factors mentioned by Van Dyk and Van de Poel.

2.2.3 Stakeholders

It is important to keep in mind the various stakeholders that might be affected by the evaluation of a programme. “The term stakeholder (...) reveals the notion of people who hold some sort of stake in the assessment or evaluation judgement or decision” (Lynch, 2003:15). As Brindley (1998:67) aptly states, just as companies will not launch a new product without thorough market analysis, it would also not be wise for educational institutions to put in place assessment processes without consulting with relevant stakeholders. In a study on the impact of an academic literacy intervention on student success, stakeholders could be grouped into two categories: those who request the evaluation, and those who are affected by the evaluation (Lynch, 2003:10). Examples of stakeholders are academic literacy lecturers, lecturers not teaching the academic literacy course but who expect their students to acquire certain abilities from the subject, students (and even their guardians, as these would want to know that the standard of education they are paying for is of a high quality), sponsors, funders, administrators, universities, policy makers, assessment specialists, academic literacy coordinators at other universities, society at large, etc. (Van Dyk & Van de Poel, 2013:44; Ross, 2009:758; Lynch, 2003:16; Brindley, 1998:67). These stakeholders can further be categorised into primary stakeholders (those who will be directly affected by the evaluation or study, like students, whoever commissioned the evaluation or study, and lecturers), secondary stakeholders (those who are not directly involved in the assessment or evaluation, but who still have occasional contact with the intervention, like content-subject lecturers) and tertiary stakeholders (people who have little or no contact with the specific intervention, but who might still be interested in the outcomes of a study assessing its impact) (Lynch, 2003:16). Primary level stakeholders should have a voice in determining the goals of the evaluation. Secondary level stakeholders may also have a voice in setting goals for the study, but might be relied on less than the primary level stakeholders. Tertiary level stakeholders will likely only learn about the evaluation outcomes through published literature.

2.2.4 Language assessment

Lynch (2003:1) defines language assessment as “the range of procedures used to investigate aspects of individual language learning and ability, including the measurement of proficiency, (...) determination of achievement in relation to syllabus objectives and analysis of ability to perform specific tasks”. According to Bachman and Palmer (2010:20), what differentiates assessment from other information collecting techniques is that assessment must be systematic (“designed and carried out according to clearly defined procedures that are methodical and open to scrutiny”) as well as substantively grounded (based on verifiable areas of content, for example course outcomes, currently accepted practice in the field, etc.). “Language testing assumes that the aspects of language ability that we are interested in are things that can be measured, however imperfectly” (Lynch, 2003:4). Although academic literacy testing covers a broader field than merely language proficiency, the principles discussed above still apply. Academic literacy testing is discussed in more detail in Section 2.5.3.1.1.

2.2.5 Programme evaluation

According to Babbie (2005:360), evaluation research can be defined as the “process of determining whether a social intervention has produced the intended result”. Weinbach (2005:2) similarly defines it as “the systematic use of research methods to make judgments about the effectiveness and the overall merit, worth, or value of some form of (...) practice”. According to Brown (2001:15), programme evaluation in an educational context concerns the continuous gathering, analysis and synthesis of data so as to improve not only the various facets of a curriculum, but also the curriculum as a whole. Because accountability has become increasingly important in recent years, De Vos *et al.* (2011:449) argue that stakeholders want to see evidence about which aspects of a programme or intervention work, how these work, and how they can be improved.

With regards to the evaluation of language programmes, there is often an overlap between programme evaluation and language assessment. This is because programme evaluation frequently uses language assessments, based on which judgements and decisions are made, and actions are taken (Bachman & Palmer, 2010:21; Lynch, 2003:1).

The concept of “design research” is also relevant to programme evaluation. Design research centres around artefacts which were created either in the form of a construct, method or model (Cole, Purao, Rossi & Sein, 2005:3-4; Hevner, March, Park & Ram, 2004:77). The evaluation design proposed in this study could be considered such an artefact (see Section 7.2), which is then used as a tool in programme evaluation. Programme evaluation as a whole is discussed more comprehensively in Section 2.5 of this literature review.

2.2.6 Evaluation design

Lynch (1996:156) states that “a methodological strategy for program evaluation can be conceptualized as a design”. Bamberger *et al.* (2012:613) expand on this by stating that an evaluation design can be seen as “the plan for conducting an evaluation” – this includes “methods and sources of data collection, deliverables, methods of analysis, and a timeline”. The purpose of the current research is to propose such an evaluation design for effectively and comprehensively assessing the impact of an academic literacy intervention on student success.

2.3 Underpreparedness for higher education studies

South Africa has a serious problem with regards to its university attrition rates. A study by Scott, Yeld and Hendry (2007) shows that approximately 30% of all undergraduate students will have dropped out before having completed their first year of study. This trend continues in students’ subsequent years of study; after five years, a mere 30% of undergraduate students graduated from university, with a total of 56% having left without graduating, and 14% still being in the system. This picture has not improved over the past six years. A recent study by Scott *et al.* (2013:43) indicates that as many as 55% of all enrolled students currently leave university without graduating, and only 27% of all students graduate their 3- and 4-year qualifications in regulation time. At best, therefore, a total of 44% of students who register for undergraduate studies are likely to obtain their degrees. Similar trends can be identified internationally (see, for example, Holder, Jones, Robinson & Krass, 1999; Johnson & Buck, 1995; Johnes & Taylor, 1989).

Universities acknowledge that academic underpreparedness puts students at risk of not achieving success at higher education institutions, both in South Africa and the rest of the world (Prinsloo & Subotzky, 2011:17). Students' underpreparedness, claim De Klerk *et al.* (2006:3), means that their "innate ability may be masked by deficiencies in knowledge, skills and academic proficiencies, that they are likely to perform below their potential and, in a significant number of cases, will fail when they may have the ability to pass". Cliff and Hanslo (2009:274) state that "[u]nderprepared students, almost by definition, are unlikely to respond to typical higher education academic tasks in theoretically or practically anticipated ways". The number of these underprepared students is staggering: the National Benchmark Tests Project's (NBTP) first full-scale pilot in 2009 showed that half of all new students would have to receive some type of academic support to enable them to obtain a degree in a reasonable time (Yeld, 2010:29). Its 2010 results suggested that 71% of all candidates would need academic literacy support (Yeld, 2010:30). This trend is echoed in national and international literature on the topic (e.g. Van der Merwe & Van der Merwe, 2009; Baer, Cook & Baldi, 2006).

The South African government, and in turn universities, lose millions of Rands due to academically underprepared students who fail to obtain their degrees (Nair, 2004:94-95). Van Dyk (2005:40) explains that "[a]cademically underprepared students (...) contribute to universities losing vast amounts of money in subsidy, as a proportion of the subsidy is only obtained when a student graduates". At the same time, universities are under a lot of pressure to accept an increasing number of diverse and underprepared students (Jama, Mapesela & Beylefeld, 2008:992-993; Basson, 2006:1, 11; Botha & Cilliers, 1999:144; Holder *et al.*, 1999:19-20) due to the national and international trend towards the massification of education (see Teichler [1998] and Calderon [2012] for a comprehensive discussion of this trend). The Department of Higher Education and Training (2014:30), for example, aims to increase participation in universities from the current rate of 17.3% to a rate of 25% by 2030. Already, the current participation rate is twice as high as it was in 1994. The challenge that the higher education sector faces, therefore, is to assist the masses of students who enter universities and other higher education facilities each year to overcome the barriers that prevent them from reaching their full potential. As Van Dyk and Van de Poel (2013:46) argue, "when students enter [higher education] they are allowed (epistemological) access and consequently these students have to be supported to

develop their awareness and abilities through assimilating, understanding, embracing, questioning, interacting and engaging with the codes and conventions of academia”.

Researchers agree that insufficient academic literacy proficiency is one of the major facets of underpreparedness for higher education (Butler, 2013:71; Terraschke & Wahid, 2011:173; Davies, 2010:xi; Defazio *et al.*, 2010:34; Van Dyk *et al.*, 2009:333; Archer, 2008:248; Holtzman *et al.*, 2005:285; Stephen *et al.*, 2004:42-51; Feast, 2002:83; Holder *et al.*, 1999:27-28). The concept of “academic literacy” is the focus of the following section.

2.4 Academic literacy

2.4.1 Defining “academic literacy”⁵

Before academic literacy can be defined, it might be useful to first situate it in a theoretical discipline. Butler (2013:74) argues that the field of applied linguistics is probably the best theoretical “home” for academic literacy. If Weideman’s (2003b:28) definition of applied linguistics as “a solution to a language teaching problem” is accepted (cf. Lynch, 1996:1), then academic literacy interventions clearly fall within the parameters of this discipline (Butler, 2013:74).

Defining academic literacy is not an easy or straightforward task, as there is no single definition of academic literacy that is accepted universally (Butler, 2013:75; Van Dyk & Van de Poel, 2013:47). In fact, Parkinson, Jackson, Kirkwood and Padayachee (2008:12) note that “academic literacy, which is a more restricted notion than literacy as a whole, and might thus be expected to be clearly one thing, is (...) interpreted differently by different groups”.

Academic literacy has been viewed through various paradigms. Possibly the most traditional of these is the skills-based paradigm (cf. Bachman & Palmer, 1996:75). Lynch

⁵ I agree with Van Dyk and Van de Poel (2013:56) that academic literacy is more than, and sometimes even different from, language proficiency. However, much of the research in this field does not refer to academic literacy *per se*, and therefore certain studies referred to in this review might use terms such as “language proficiency”, “academic language abilities”, “proficiency in the language of teaching and learning”, “English for Academic Purposes” etc. For the purposes of the current study, these terms should for the most part be seen in the context of academic literacy research.

(2003:46) points out that a long tradition exists “of discussing and categorising language tests [and by extension, all academic literacy activities] based on the four skill areas of reading, listening, writing and speaking”. Boughey (2013:33) argues that, in many instances, this is still the case. Yet, as Lynch (2003:46) points out, most language activities involve a combination of several language skills. Thus, viewing language as a set of distinct and neutral “skills” becomes almost nonsensical. Furthermore, following a skills-based approach could lead to interventions overemphasising certain skills and neglecting others, thus ignoring and even losing the interrelatedness of typical university-level tasks (Butler, 2013:75). As a result, the skills-based view of academic literacy has become increasingly contested in academic literacy research (Butler, 2013:75; Jacobs, 2013:128; Archer, 2006:450; Weideman, 2003b:28-29).

A more recent and popular view of academic literacy is that of the New Literacies Studies. According to Butler (2013:76), this view “supports a ‘social practices’ account of academic literacy” and stresses the fact that academic literacy practices cannot be separated from the “norms, values and ways of thinking and behaving in distinct discourse communities”. These discourse communities often refer to specific academic disciplines, such as law, biology or history. Taylor, Ballard, Beasley, Bock, Clanchy and Nightingale (1988:5) argue that the term “academic literacy” has no meaning outside the specific culture (or cultures) of universities. Leibowitz (2001:2) also focuses on the cultural conventions of academic literacy. These conventions, once known and understood by students, enable them to appropriately participate in specific academic ‘Discourses’ (Boughey, 2000). The word ‘Discourses’ here is capitalised intentionally, and refers to Gee’s notion of Discourse. According to Gee (2008:155), a Discourse is...

...composed of distinctive ways of speaking/listening and often, too, writing/reading *coupled* with distinctive ways of acting, interacting, valuing, feeling, dressing, thinking, believing (...) so as to enact specific socially recognizable identities engaged in specific socially recognized activities.

Lea and Street (1998:158), in discussing educational research into student writing, similarly identify various models in a New Literacies Studies paradigm, namely “academic literacy”, “academic socialisation” (both implicit in Gee’s definition of Discourse) as well as “study skills”. The authors argue that the academic literacy model encapsulates the constructs of academic socialisation and study skills. Academic literacy is, on the one hand, a new way of communicating that students need to appropriate to be

accepted into this unfamiliar culture in which they wish to succeed, but it is also a new way of thinking about knowledge and the world that students should acquire.

Though the New Literacies Studies have undoubtedly enriched our conception of academic literacy, this theory has also been criticised in recent years. Even proponents of this theory have admitted that the theory has not yet been able to contribute to practice. Lillis (2003:192), for example, states that “[w]hilst powerful as an oppositional frame, that is as a critique of current conceptualisations and practices surrounding student writing, academic literacies has yet to be developed as a design frame (...) which can actively contribute to student writing pedagogy as both theory and practice”. Lea (2004:741) agrees with Lillis, and adds that “the focus of this body of research, both on critique and primarily on student writing, might also indicate why the work has not yet been taken up by educational developers concerned with pedagogy and practice more broadly, rather than specifically with supporting assignment writing”. Other researchers, like Van Dyk and Van de Poel (2013:51), criticise the theory for focusing too much on specific groups, specifically traditionally marginalised students. In brief, it could be argued that “[n]owadays, the Academic Literacies movement tends to be nothing more than criticism guided by postmodern reasoning” (Van Dyk & Van de Poel, 2013:50).

Some definitions do not strictly fall in either the domain of the skills-based approach or that of the New Literacies Studies. Kaburise (2012:78), for example, differentiates between general English courses and academic literacy courses by saying that support in the latter is not focused on improving students’ English *per se*, but rather at improving their ability to “manipulate and apply this knowledge in performing varying academic tasks”. Van Dyk and Van de Poel (2013:46) argue that literacy has three dimensions: a social dimension where information is exchanged; a cognitive dimension where students organise, understand and reason about information; and a linguistic dimension that relates to language itself (cf. Kern, 2000:38). Cliff and Yeld (2006:19) define academic literacy as “students’ capacities to engage successfully with the demands of academic study in the medium of instruction of the particular study environment”.

Jacobs (2013:136-137) calls for a “shared ontology regarding academic literacy work” in South Africa, since, she argues, the field of academic literacy is currently “chaotic” due to widely different conceptions of academic literacy work and research. Van Dyk and

Van de Poel (2013:50) agree that “there is a lack of interdisciplinary collaboration with regard to developing theory and responsibly designing practices to enhance academic literacies that will truly benefit students”. They recommend placing how knowledge is constructed in various academic fields at the centre of how we see academic literacy. Van Dyk and Van de Poel (2013:58-59) aptly summarise what many of the researchers mentioned above seem to imply:

In our opinion the answer does not lie in debates on autonomous or ideological approaches to academic literacy development, often found in the literature. The former refers to the perspective that students have problems and that their problems can be fixed; literacy then resides in the individual’s ability to decode text. The latter refers to the perspective that students should be provided with the opportunity to develop agency and gain a voice; literacy then resides in shared values belonging to shared discourse communities. The answer rather lies in the academic acculturation of students, which entails much more than just socialising students into academic discourses, but rather socialising them into the norms, values and practices of the academic community at large, as well as within disciplines.

“Academic acculturation”, these authors argue, is hard work for students; students must become members of academic communities, acquire the practices of these communities, become academically literate and ultimately master the academic discourse (in all its facets) (Van Dyk & Van de Poel, 2013:59). Van Dyk and Van de Poel (2013:56) accordingly argue for an “open, non-restricted view of language” and define academic literacy as “being able to use, manipulate, and control language and cognitive abilities for specific purposes and in specific contexts”. The current study accepts this comprehensive definition of academic literacy as a working definition.

2.4.2 A conceptual framework for academic literacy

Several conceptual frameworks exist that attempt to explain the use and acquisition of academic English, for example the New Literacies Studies Paradigm (Lea & Street, 1998), Systemic Functional Linguistics (SFL, or the ‘Sydney School’) (Martin, 2012; Halliday, 1993), the BICS/CALP framework (Cummins & Swain, 1986), the ESP Genre Analysis framework (Swales, 1990), and the Four Resources Model (Luke & Freebody, 1999). However, none of these seem to fully encapsulate all the dimensions of academic literacy. The New Literacies Studies Paradigm, for example, has been criticised for remaining too theoretical, not actively contributing to student writing pedagogy, and

being difficult to implement in large-scale projects due to its labour intensiveness (Van Dyk & Van de Poel, 2013:48-50). The use of the BICS/CALP framework in South African higher education has also been criticised, since this framework was originally developed for assisting children speaking minority languages in bilingual education, and has not been proven to be applicable to the South African higher education context (Coetzee-Van Rooy, 2010:25-26). Other frameworks, like SFL, the ESP Genre Analysis Framework and the Four Resources Model, focus only on specific aspects of academic literacy, such as language structures, the genres that are used when communicating in certain contexts, or reading in academic contexts.

Kern (2000:16-38) situates academic literacy in a sociocognitive framework. He argues that literacy (including academic literacy) consists of three dimensions: linguistic (how language is used), cognitive (active thinking and problem solving), and sociocultural/psychological (acculturation into the specific conventions of particular discourse communities) (cf. Van Dyk & Van de Poel, 2013:46). These three dimensions, which are summarised in Figure 2.1, overlap and are interdependent. According to Kern, literacy cannot be understood without considering all three of these perspectives.

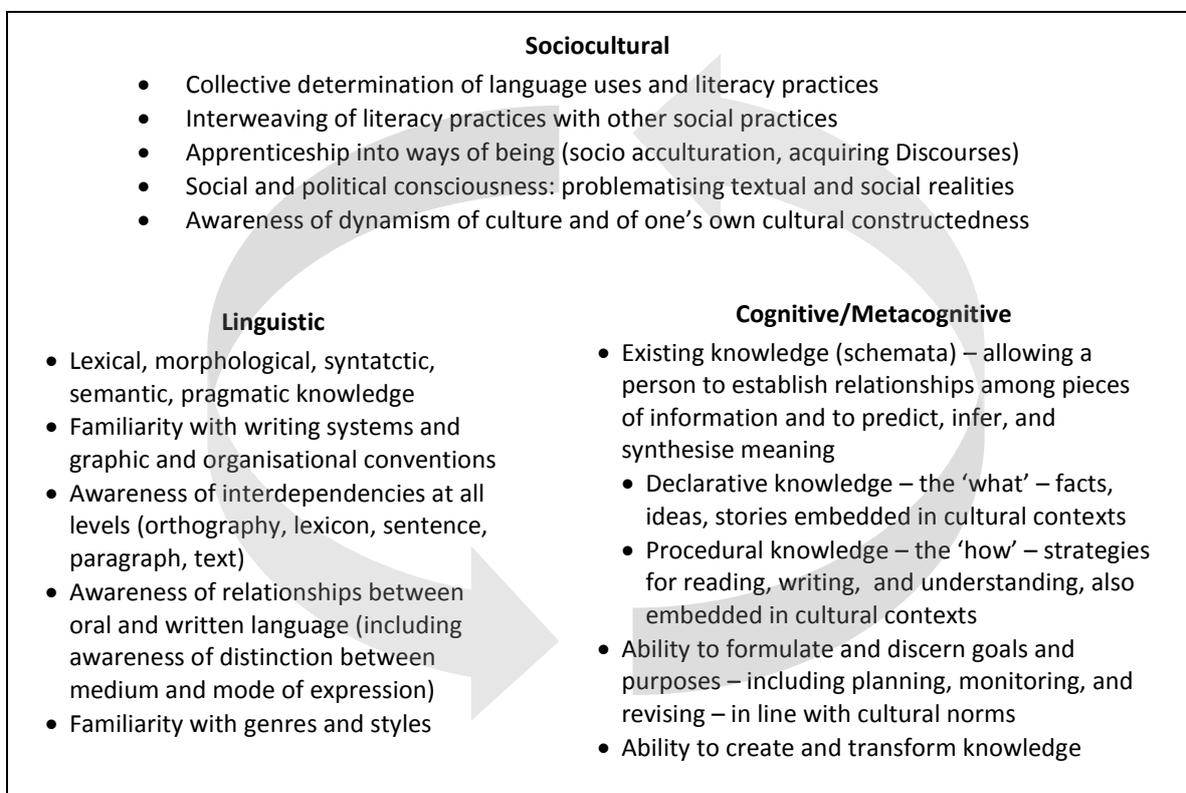


Figure 2.1: Linguistic, cognitive, and sociocultural dimensions of literacy (Kern, 2000:38)

Kern (2000:16-17) expands on these dimensions by discussing several principles of literacy. He argues that literacy involves interpretation, collaboration, conventions, cultural knowledge, problem solving, reflection and self-reflection, as well as language use. The sociocultural, linguistic and cognitive dimensions, he argues, are “infused in each of the seven principles of literacy” (Kern, 2000:38). He elaborates on these principles as follows:

1. Literacy involves *interpretation*. Writers and readers participate in double acts of interpretation – the writer interprets the world (events, experiences, ideas, and so on), and the reader then interprets the writer’s interpretation in terms of his or her own conception of the world.
2. Literacy involves *collaboration*. Writers write for an audience, even if they write for themselves. Their decisions about what must be said, and what can go without saying, are based on their understanding of their audience. Readers in turn must contribute their motivation, knowledge, and experience in order to make the writer’s text meaningful.
3. Literacy involves *conventions*. How people read and write texts is not universal, but governed by cultural conventions that evolve through use and are modified for individual purposes.
4. Literacy involves *cultural knowledge*. Reading and writing function within particular systems of attitudes, beliefs, customs, ideals, and values. Readers and writers operating from outside a given cultural system risk misunderstanding or being misunderstood by those operating on the inside of the cultural system.
5. Literacy involves *problem solving*. Because words are always embedded in linguistic and situational contexts, reading and writing involve figuring out relationships between words, between larger units of meaning, and between texts and real or imagined worlds.
6. Literacy involves *reflection and self-reflection*. Readers and writers think about language and its relations to the world and themselves.
7. Literacy involves *language use*. Literacy is not just about writing systems, nor just about lexical and grammatical knowledge; it requires knowledge of how language is used in spoken and written contexts to create discourse.

(Kern, 2000:16-17)

Although Kern’s (2000) framework is aimed at second and foreign language learners, it can be applied effectively to academic literacy. In fact, it could be argued that academic literacy is a second (or foreign) language to all first-year students. For the purposes of the current study, the conception of academic literacy will be situated in this framework.

As language use forms a significant part of academic literacy though, it is also useful to consider Bachman and Palmer’s framework of language competence. Bachman and Palmer (1996:61-62) begin by defining language use as “the creation or interpretation of intended meanings in discourse by an individual, or as the dynamic and interactive negotiation of intended meanings between two or more individuals in a particular situation”. Therefore, they believe that “language ability must be considered within an interactional framework of language use” (Bachman & Palmer, 1996:62). The authors state that multiple complex interactions among the characteristics of the language users and the characteristics of the language use situation are involved in language use (see Figure 2.2). The characteristics of the language user include topical knowledge, language knowledge and personal characteristics (Bachman & Palmer, 1996:62-64). These authors argue that the nature of language use (for example, how language is used naturally in specific contexts) must be kept in mind if we want to generalise interpretations from language tests and other activities to domains outside the language intervention or test (Bachman & Palmer, 2010:33). Thus, in an academic literacy context, the language tasks in classrooms and tests need to closely approximate how students would use language in their other subjects.

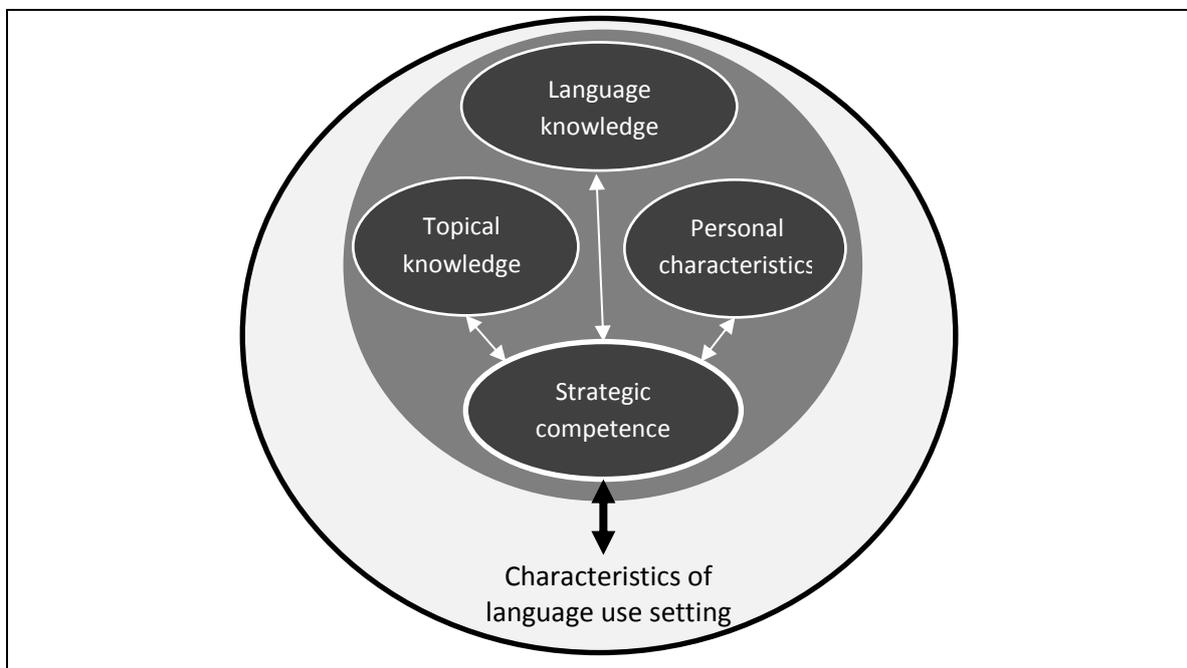


Figure 2.2: An interactional framework for language use

(Bachman & Palmer, 1996:63)

Language ability is seen as an interplay between language knowledge and strategic competence (Bachman & Palmer, 2010:33). Van Dyk and Weideman (2004:8) create a useful visual representation of Bachman and Palmer’s construct of language ability in Figure 2.3.

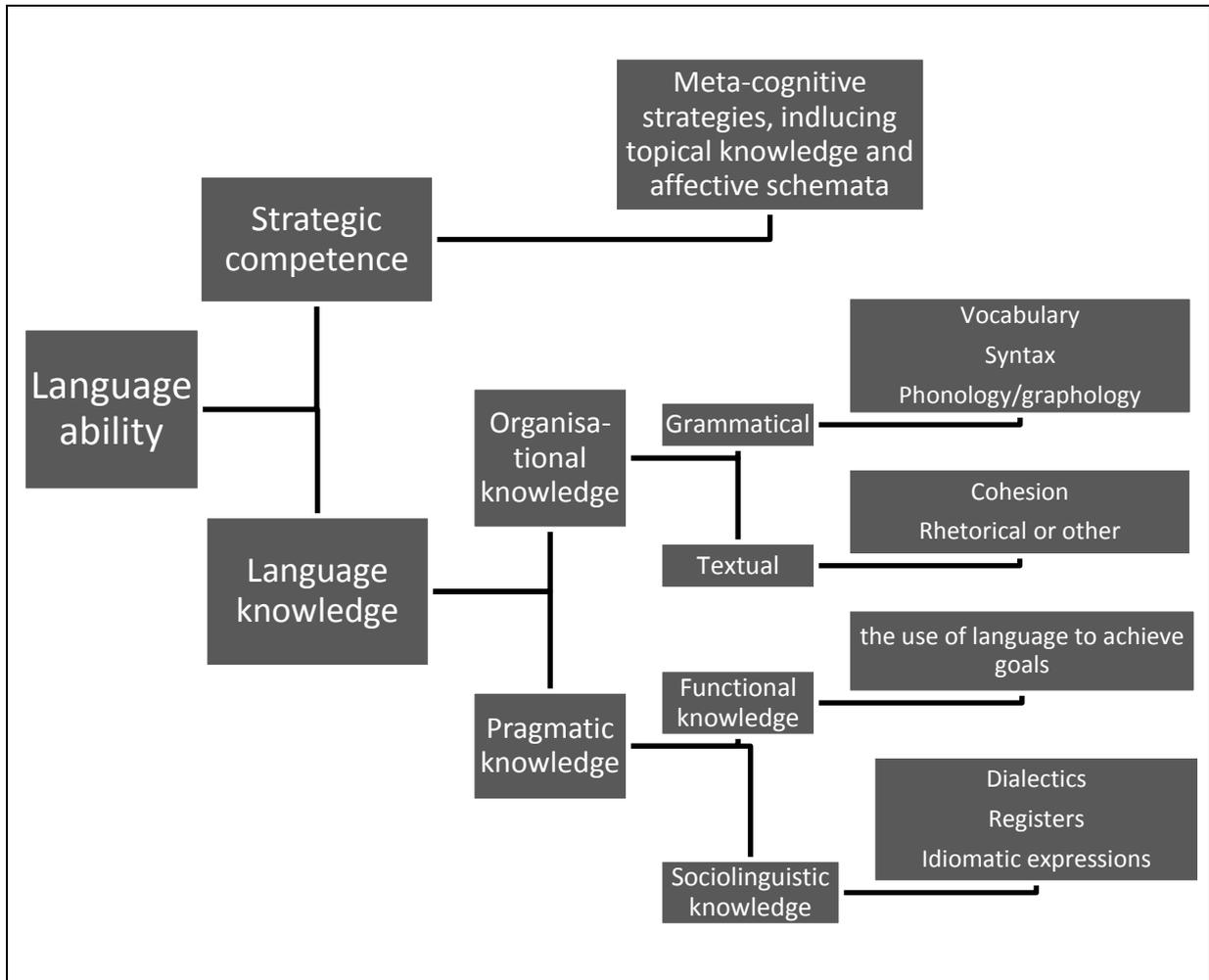


Figure 2.3: The Bachman and Palmer construct of language ability

Adapted from Bachman and Palmer (1996:63) by Van Dyk and Weideman (2004:8)

There has, however, been criticism against this construct of language ability (cf. Van Dyk & Weideman, 2004:8-9), specifically because certain areas are not always as distinct as they would appear in Figure 2.3. Bachman and Palmer (1996:66) themselves warn that this framework must be contextualised for each particular language situation and use. Bachman and Palmer (1996:76) further argue that instead of speaking of specific “skills”, it is much more useful to consider any language task in terms of the language ability in

combination with the task characteristics (the authors call this an “ability-task”) (cf. Bachman & Palmer, 2010:55-56). According to Bachman and Palmer (1996:75-76):

We would thus not consider language skills [i.e. reading, writing, speaking and listening] to be part of language ability at all, but to be the contextualized realization of the ability to use language in the performance of specific language use tasks. We would (...) argue that it is not useful to think in terms of skills, but to think in terms of specific activities or tasks in which language is used purposefully.

Accordingly, throughout the current study the term “abilities” is used when referring to outcomes related to academic literacy.

Douglas (2000:38) agrees and argues that instead of thinking of the four skills as distinct constructs to be tested, it might be more useful to think about them as “the means by which that ability is realized in the performance of tasks in actual language use situations”. For example, though one might refer to a specific test as a writing test, in reality it might be a test that requires reading abilities, critical thinking abilities, and possibly even speaking and listening abilities, with only the ultimate outcome being presented in the form of writing (also see Weideman’s [2013:13-14] example of calling a test a ‘listening’ test for convenience sake, even though it is understood that it is not just the isolated ability of listening that is really being tested). When the terms “writing”, “reading”, “listening” and “speaking” are discussed in the current study for the purposes of identifying an assessment or task, the supposition will be that, in fact, the language assessment or task referred to consists of a complex myriad of several language abilities.

For the purposes of the current study, these two frameworks (i.e. that of Kern [2000] and Bachman and Palmer [1996]) are seen as complementary, and underlie this study’s conception of academic literacy.

2.4.3 Academic literacy levels as an indicator of academic success

Low academic literacy levels in students’ language of teaching and learning are generally seen as a major reason why many undergraduate students with high academic potential are unable to complete their studies in the prescribed time (see, amongst others, Keeve, Naudé & Esterhuysen, 2012:147; Petersen-Waughtal & Van Dyk, 2011:99; Yeld, 2010:29; Baik & Greig, 2009:402; Bojuwoye, 2002:287; Hyland & Hamp-Lyons, 2002:1; Pretorius, 1995:33). Farnill and Hayes (1996), for instance, statistically show a

correlation between students' English language proficiency and their cumulative first-year marks. Van Dyk and Van de Poel (2013:52) explain that language "could be considered the cornerstone of literacy and literacy, in turn, is crucial for academic success". Cliff and Yeld (2006:19-22) also point to the strong relationship between language and performance at university. What is more, communication ability in general is seen as one of the most important predictors of career success (Sinclair, 1995:41).

There are various reasons why students with low levels of academic literacy are likely to have difficulties with the demands of academic work. Firstly, such students will find it difficult to meaningfully interact with prescribed material (Petersen-Waughtal & Van Dyk, 2011:101; Van Dyk, 2005:38; Stephen *et al.*, 2004:45-46; Amos & Quinn, 1997:187) and thus to learn new content (Yeld, 2010:26). Secondly, students will struggle to produce acceptable writing in an academic context; in other words, to produce and to convey knowledge (Petersen-Waughtal & Van Dyk, 2011:101; Stephen *et al.*, 2004:45-46; Yeld, 2001:130; Amos & Quinn, 1997:226). Academic literacy becomes an even greater obstacle to student success if the language of teaching and learning is a student's second or third language (Leibowitz, 2001:22-23, 28-29). Moreover, few lecturers are willing to assist in students' acquisition of academic literacy, as they see poor academic literacy levels as "a failure of the schooling system, and [thus] not a legitimate concern of the academe" (Moore, 1998:84). In addition, lecturers are rarely equipped to assist students who display poor academic literacy levels (Van Dyk & Van de Poel, 2013:52; Misselhorn, 1997).

Rosenthal (1996:24) points out that language proficiency is not always a product of intelligence, motivation or academic ability. Instead, she argues, it is an ability that is practised and improved over time. Eskey (1983:322) agrees that many students need guidance and explicit instruction in the process of becoming proficient in the use of academic language. Several studies indicate that academic literacy abilities do not develop without explicit instruction.

De Graaff and Housen (2009:729) state that a variety of studies have confirmed that, in general, second language learners who have received language instruction have higher levels of language proficiency than those learners who have not been instructed. Holder *et al.* (1999:27) indicate that students' academic literacy levels at the beginning of their

studies are statistically linked to them completing their degrees in the minimum required time. They state that if students' low academic literacy levels are not addressed, "the predictive validity of the literacy ratings will be maintained" (cf. Farnill & Hayes, 1996:264). Thompson (1990:102) adds that the writing proficiency of students with low language abilities is unlikely to significantly improve without explicit writing instruction. If we accept these arguments, it follows that academic literacy courses must have the potential of impacting on students' academic language usage (in that they guide students towards appropriate academic language usage, and also explicitly instruct students in such language usage), which in turn will impact on their academic success. The first step, then, would be to identify students who are at risk of failing their higher education studies due to low academic literacy levels.

2.4.4 Identifying at-risk students

Academic literacy is of course not the only possible barrier to student success. Several barriers, for example socio-economic circumstances, secondary education, being a first-generation student, a heritage of unequal distribution of resources (particularly in South Africa), personal problems, inherent factors (for example a lack of intrinsic motivation and locus of control), poor time-management strategies and many more can prevent students from reaching their full potential (Keeve *et al.*, 2012:146-169; Mcghie, Van Der Walt, Van Schalkwyk, Walt & Schalkwyk, 2012:31; Van Dyk *et al.*, 2009:333; Stephen *et al.*, 2004:43-45; Penrose, 2002:455). There is very little that universities can do about many of these factors. Where universities are however able to impact on students' chances for success, they have to do all in their power to do so. It would seem that the potential for improving students' academic literacy levels is one area in which universities have the potential to effect a change. Universities therefore become ethically obliged to firstly identify the students who are most at risk of dropping out of university because of insufficient academic literacy levels, and to secondly put in place effective interventions that could assist students in overcoming this obstacle.

It is important that “at risk” students be identified as soon as possible (Stephen *et al.*, 2004)⁶. According to Potter and Van der Merwe (1993), first-year assessment tests are excellent indicators of at-risk students. There are various tests available for identifying students who are at risk of failing their higher education studies. Examples of these tests are the ELSA-Plus (Van Dyk & Weideman, 2004:3), the Placement Test in English for Educational Purposes (PTEEP) (Cliff & Hanslo, 2009:268-274), the academic literacy section of the National Benchmark Test (NBT) (Yeld, 2010:29), and the Test of Academic Literacy Levels (TALL) (Van Dyk, 2005). Poor academic literacy levels are not restricted to undergraduate studies; as a result, the Test of Academic Literacy for Postgraduate Students (TALPS) has been developed to assess the ability of postgraduate students to deal with the language demands of postgraduate degrees (Butler, 2009). These tests do not all have the same purpose in mind. The NBT, for example, is mainly used for selection purposes, while the TALL is used for placement purposes. The proposed study will use the TALL as part of a design for evaluating students’ academic literacy levels. Between 2000 and 2005, the TALL consistently identified approximately a third of the student population at three major South African universities as being “at risk of not completing their studies in the prescribed time” due to poor academic literacy levels (Van Dyk, 2005:45). These findings closely correspond to those of, for example, Scott *et al.* (2007), as discussed in Section 2.3.

As stated in Section 2.4.2, Bachman and Palmer (1996:75-76) argue that “it is not useful to think in terms of [language] ‘skills’, but [that one should rather] think in terms of specific activities or tasks in which language is used purposefully”. Ideally, test developers should follow this advice, and tests’ constructs should reflect this by being based on language tasks rather than language skills, so as to more effectively identify at-risk students. The term “test construct” can be defined as “an ability or set of abilities that will be reflected in test performance, and about which inferences can be made on the basis of test scores” (Davies, Brown, Elder, Hill, Lumley & Mcnamara, 1999:31).

⁶ One aspect that researchers should take into account when considering the issue of at-risk students with regards to academic literacy is the possibility that all students, and not just certain so-called “weaker” students, are in need of academic literacy support programmes (see for example Petersen-Waughtal and Van Dyk [2011:101]). I agree with this sentiment, but it is not within the parameters set by this study to prove or disprove the necessity of academic literacy courses for all students. However, should the proposed evaluation design for assessing the impact of academic literacy courses prove effective, further research could be conducted on the impact of these courses on so-called “stronger students” who were not initially identified as being at-risk.

Douglas (2000:111) adds that a test construct makes “the nature of the ability one wants to measure” explicit. The test construct of the TALL is discussed in more detail in Section 2.5.3.1.1.

2.4.5 Academic literacy interventions

Because of the extent to which academic literacy levels seem to influence student success, universities have no choice but to support under-prepared students by means of mechanisms that attend to factors like academic literacy which affect students’ academic success (Van der Merwe & Van der Merwe, 2009:299; Botha & Cilliers, 1999:144). There is a wide range of approaches to academic literacy interventions aimed at assisting to increase students’ academic literacy levels. This section briefly describes the most prominent among them.

Most academic language proficiency courses at universities can be classified as English for Academic Purposes (EAP) courses. For the purposes of this study, EAP courses will be seen as synonymous with academic literacy courses. “EAP is normally considered to be one of two branches of English for Specific Purposes [ESP], the other being EOP [English for Occupational Purposes]” (Flowerdew & Peacock, 2001:11). Winberg, Wright, Birch and Jacobs (2013:91) define ESP as a framework where language and content are integrated, with systematic connections that link the development of thinking and relevant learning. Robinson (1980:13) describes an ESP course as a course that “is based on a rigorous analysis of students’ needs and should be ‘tailor-made’”, while Hyland (2009:202) states that ESP courses identify “the demands placed by academic or workplace contexts on communicative behaviors and developing pedagogic practices by which these behaviors can be taught”. In fact, argues Robinson (1980:13), “[a]ny ESP course may differ from another in its selection of skills, topics, situations and functions and also language”. EAP courses are English courses that focus on the acquisition of the literacy abilities needed in academic contexts. They are usually taught to students who need English for study purposes (Kennedy & Bolitho, 1984:4-5). EAP courses generally address two issues, “namely to improve students’ overall language proficiency and to introduce them to the linguistic conventions and academic skills that they need for their university study” (Terraschke & Wahid, 2011:174). EAP courses can be interdisciplinary (Clark, 1998:124); that means that the same generic course might well be presented to,

amongst others, commerce, natural science and humanities students (e.g. Van Wyk, 2002), as is the case with the course focused on in Chapters 5 and 6. However, the recent trend has been towards discipline specific EAP courses (cf. Carstens, 2013a). Examples of such courses are courses in law (e.g. Ngwenya, 2010), health sciences (e.g. Defazio *et al.*, 2010; Van Dyk *et al.*, 2009), natural sciences (e.g. Fouché, 2009; Goodier & Parkinson, 2005; Jacoby, Leech & Holten, 1995), human sciences (e.g. Pienaar, 2005; Kasper, 1994), engineering (e.g. Butler & Van Dyk, 2004) and economic and management sciences (e.g. Goodier & Parkinson, 2005; Connor & Kramer, 1995).

Proponents of discipline-specific EAP courses argue that students should become part of the discourse communities of their various academic disciplines (Johns, 1995:280; Berkenkotter, Huckin & Ackerman, 1991:191). Berkenkotter *et al.* (1991:191) argue that “students entering academic disciplines need a specialised literacy that consists of the ability to use discipline-specific theoretical and linguistic conventions to serve their purpose as writers”. In addition to general academic discourse rules that are true for all academic discourse, students also need to acquire a more specific set of rules that apply to particular disciplines, and in some cases even individual courses (Johns, 1995:278-280). Although possible differences in the writing conventions of various faculties play an important role in justifying discipline-specific EAP courses, it could be argued that an even more important factor is that students appear to be more motivated by this type of course (see, for example, Ngwenya, 2010:78; Zambo & Cleland, 2005:157; Rosenthal, 1996:141; Hutchinson & Waters, 1987:48).

Approaches which are often used in EAP courses can be broadly categorised under generic, subject-specific, collaborative (generally a subsection of subject-specific interventions), and limited purpose academic literacy interventions (Van de Poel & Van Dyk, 2015:169-173). Examples of approaches that might fall under subject-specific academic literacy interventions include the following: Content-Based Instruction (CBI) (Stoller, 2004; Grabe & Stoller, 1997), which integrates content- and language-learning objectives; the broader discipline-specific approach (Braine, 1995; Kennedy & Bolitho, 1984), where the EAP course focuses on and uses material from a broad discipline, for example the natural sciences or law; and genre-based approaches, such as SFL and ESP approaches (Martin, 2012; Hyon, 2001; Paltridge, 2001; Jacoby *et al.*, 1995; Flowerdew, 1993; Halliday, 1993; Swales, 1990), which generally focus on teaching specific genres

of text that are relevant to various disciplines. Approaches that might be categorised as collaborative academic literacy interventions include: the adjunct model (Rosenthal, 1996; Johns, 1995; Snow & Brinton, 1988), which pairs and coordinates EAP courses and content-area courses; team taught ESP teaching (Dudley-Evans, 1995:304-308), in which a content specialist and a language specialist teach the same class, with the content specialist focusing on content-related matter, and the language specialist focusing on language-related matter; and the New Literacies approach (Street, 2008, 2004; Lea & Street, 1998; Street, 1993; Gee, 1991; Street, 1984), which considers language and social literacy not as technical skills, but rather as social practices that are practised within institutions which in turn are sites of discourses and power.

Clearly, EAP (or academic literacy) courses in many different forms are used extensively nationally and internationally. Furthermore, as can be seen from the literature above, the structure, advantages and disadvantages of these courses have been reported on extensively. However, very few studies have attempted to measure the impact of these courses on student success (Mhlongo, 2014:47; Butler, 2013:79-80; Terraschke & Wahid, 2011:174; Carstens & Fletcher, 2009b:319; Storch & Tapper, 2009:208; Holder *et al.*, 1999:20). Studies which have made this attempt are discussed in Chapter 3 of this study. This gap in the discourse regarding academic literacy courses necessitates a comprehensive consideration of how the impact of these courses could be effectively assessed. That, then, is the major goal of the current study.

2.5 Programme evaluation

There is an increasing demand across the world for systematic data on programme performance; funders want to know that the programme adds value to its target audience, staff members want to understand the strengths and weaknesses of their programme so that they can improve it, etc. (Newcomer *et al.*, 2010:7). According to Lynch (2003:2), programme evaluation serves two key purposes in educational contexts. The first purpose is to ascertain to which extent the programme's objectives have been met, while the second purpose is to determine what the links are between students' achievement and the specific programme or intervention. Once the above has been determined, inferences about the effectiveness of specific components of the intervention can be made (Lynch, 2003:7). Ideally, this should translate into finding ways in which the intervention can

subsequently be made even more effective (Newcomer *et al.*, 2010:6). In this way, argues Brown (2001:15), the evaluation process becomes an ongoing needs assessment which leads to continuous programme development (cf. Bachman & Palmer, 2010:25).

Although formative evaluation can be used in programme evaluation, it is more likely that summative evaluation might be used (Lynch, 2003:7, 10; Henning, 1987:3). The purpose of summative evaluation is to “assess impact and to provide information that can be fed into repeat versions or related activities”; therefore, it is useful for “durable courses” (Dudley-Evans & St John, 1998:128). Summative evaluation is often linked to a positivist research paradigm (Lynch, 2003:21), which is discussed in more detail in Section 2.5.1.

There exist varied approaches to programme evaluation. Brousselle and Champagne (2011), for example, discuss programme theory evaluation, and how logic analysis can be used in this type of evaluation. In logic analysis, the theory of a programme is evaluated based on scientific literature reviews and other expert judgements (Brousselle & Champagne, 2011:69-70) (cf. Ross, 2009:759-761; Renger & Titcomb, 2002). Another approach is that of implementation monitoring, “to serve as a guarantee that the program intervention was in fact delivered in ways consonant with the program theory” (Ross, 2009:766). Neither of these, however, include an analysis of the effects of a programme. It is precisely these effects, however, that form the focus of the current study.

Lynch’s (1990) context-adaptive model (CAM) for language programme evaluation is considered an appropriate organising principle for the current study. This flexible and iterative model consists of the following steps: 1) establish the evaluation’s audience and goals, 2) develop a context inventory, 3) develop a preliminary thematic framework, 4) develop a data collection design, 5) collect the data, 6) analyse the data; and 7) report on the evaluation (Lynch, 1996:3-9).

The first two steps of this model provide a useful contextualisation of the research that is to be conducted. Lynch (1996:3) argues that the first step in evaluating a programme would be to identify and consult relevant stakeholders. The goals of the evaluation will depend on having identified, and possibly consulted with, these stakeholders. The next step, argues Lynch (1996:5), is to identify the characteristics of the programme. He

recommends an inventory, in the form of a checklist, of possible aspects to keep in mind. These include the following:

- 1) Availability of a comparison group
- 2) Availability of valid and reliable measures of language ability
- 3) Identification of different types of evaluation expertise (e.g. statistical analysis)
- 4) Timing of the evaluation
- 5) Selecting students for the programme
- 6) Characteristics of the students
- 7) Characteristics of the staff
- 8) Size and intensity of the programme
- 9) Availability of resources for the programme
- 10) Perspective and purpose of the programme
- 11) Social and political climate surrounding the programme

Note that all of the above points might not be equally relevant to all programmes. Lynch (1996:6) recommends that the evaluator use the above inventory by considering which aspects are relevant to the intervention being evaluated, and focusing on those aspects. Once the first two steps have been considered, the evaluator can proceed to construct a preliminary thematic framework based on the outcomes of these steps. “A preliminary thematic framework provides a conceptualization of the program in terms of the salient issues and themes that have emerged from the determination of audience and goals and the elaboration of the context inventory” (Lynch, 1996:6). Lynch (1996:6) further argues that “[a]rticulating this framework provides the evaluator with a focus that will guide the collection and analysis of evaluation data”. The data collection design forms the fourth step of the CAM, and is informed by the preliminary thematic framework. Likewise, the collection and analysis of data is informed by the data collection design. These three steps (i.e. preliminary thematic framework, data collection design, and data collection and analysis) are integral to the implementation phase of the current study (see Chapters 5 and 6), and should be kept in mind before selecting instruments from the proposed evaluation design (see Section 7.7).

2.5.1 Paradigms in programme evaluation

Lynch (2003:4) further states that programme evaluation will generally be based on either a positivist or an interpretivist paradigm. The positivist view is based on the notion of objectivity, assumes that the researcher will be neutral, and holds the belief that causal relationships can be discovered, and that results can be generalised to larger populations. Due to its focus on objectivity, this research paradigm relies on quantitative research. Conversely, the interpretivist view⁷ is based on the notion that all research must be subjective and that knowledge is socially constructed (Kreuger & Neuman, 2006:72-82; Henning, Van Rensburg & Smit, 2004:17-20; Lynch, 2003:4); accordingly, it relies on qualitative research. Lynch (2003:20-21) argues that assessment and evaluation are inherently political. Very often there is an explicit request for statistical evidence from primary-level stakeholders, like the funders of the intervention, or even from tertiary level stakeholders. In such cases, a positivist paradigm might be called for.

Though many research designs can be identified as either being purely positivist or interpretivist, some researchers have argued that there might be a middle ground between these two approaches (see for example Brindley [1998] for a suggestion of how this could be done). One such example is the research paradigm of postpositivism⁸. In postpositivism, qualitative data and analysis can be used to help interpret quantitative results (Lynch, 2003:6; 1996:12). According to Patton (2002:92-93):

[Postpositivism] recognizes that discretionary judgment is unavoidable in science, that providing causality with certainty in explaining social phenomena is problematic, that knowledge is inherently embedded in historically specific paradigms and is therefore relative rather than absolute, and that all methods are imperfect, so multiple methods, both quantitative and qualitative, are needed to generate and test theory, improve understanding over time of how the world operates, and support informed policy making and social program decision making. While modest in asserting what can be known with any certainty, postpositivists do assert that it is possible, using empirical evidence, to distinguish between more and less plausible claims [...].

⁷ The term “interpretivist paradigm” is used in this study; however, elsewhere this type of paradigm is referred to by other terms, most notably the “naturalistic paradigm” (see Lynch, 1996:14).

⁸ I am aware of criticism against the postpositivist paradigm, like that given by Holliday (2004) and Lincoln, Lynham and Guba (2011), and of even more prevalent criticism against the positivist paradigm, for example by Zuengler and Miller (2006). Yet, I agree with Lynch (2003) that this type of research is often required by stakeholders. Unlike a positivist research paradigm, a postpositivist paradigm does allow for more qualitative reflections in the evaluation of an intervention. By means of triangulation, these reflections can then either strengthen the more quantitative research, or alternatively contradict it.

This paradigm is grounded in positivism, yet makes it possible to be “open to new perspectives that could deepen our understanding” (Lynch, 2001:358). Still, it remains possible to address requirements of large-scale comparability (e.g. comparing the intervention being investigated to similar programmes at other universities), thus satisfying the demands of various stakeholders to, for example, understand how the specific intervention’s standards compare to the standards of similar interventions elsewhere (Brindley, 1998:47, 50). The greatest strength of such a mixed strategies research paradigm is that the “richness of evidence that accrues to a mixed design can be a useful validity check” (Lynch, 2003:29); something that is particularly valuable in a research design such as the one in the current study where it is not possible to use a traditional experimental design with a control group. This study uses such a postpositivist research paradigm as its point of departure.

2.5.2 Evaluating impact

At this point, the main challenge for [academic literacy] practitioners who want to improve their own practice is that there seems to be an oversupply of studies that are largely descriptions of and theoretical justifications for interventions. There are too few studies that report on the real successes or failures of such interventions. (Butler, 2013:83)

Butler’s statement highlights the need for research, and the subsequent publication of this research, on the impact that academic literacy interventions have. Van Dyk *et al.* (2011a) agree that academics must continuously critically consider, as well as measure the impact of, their courses. This, however, is not often the case.

Evaluating impact falls under the umbrella of programme evaluation, but it addresses only one facet thereof (De Vos *et al.*, 2011:453). Several other factors, including lecturers’ work satisfaction and the cost-effectiveness of the programme, could be included under the umbrella of programme evaluation. Thus, even though impact evaluation would likely fall under any programme evaluation process, it is a distinct concept that should be considered separately.

According to Cheetham *et al.* (1992:10), “effectiveness derives from a variety of perspectives and assumptions”. Thus, it would seem counterproductive to merely

consider one aspect of impact (or effectiveness). Rather, the evaluator must aim to look at a variety of facets regarding impact in order to obtain a holistic view thereof, so as to come to a valid conclusion regarding the effectiveness of an intervention. In order to do this, the current study is based on a postpositivist research paradigm (see Section 2.5.1).

2.5.2.1 Considerations when evaluating impact

Several considerations need to be taken into account when evaluating the impact of an intervention. This section considers six such aspects, namely the duration of the impact study, the number of test subjects, whether insiders or outsiders should perform the evaluation, the impact that specific teachers or lecturers might have on results, the context in which the evaluation takes place, and the amount and type of data that should be used when measuring impact.

When evaluating impact, one important aspect that needs consideration is what the duration of such an impact study should ideally be. Beretta (1992:9) rightly points out that language learning is generally seen as a long-term task, and that one cannot really make generalisations after having evaluated only a few lessons. Yet, he points out, “the longer the period under study, the greater the contamination from extraneous variables, the greater the risk of drop-out, changes in the project’s direction or content, and so on” (Beretta, 1992:9). Thus, though it is not advisable to evaluate a programme after too short a timeframe, evaluating over too long a time might also make the results unreliable. A middle-ground must ideally be found. This middle-ground might be more easily defined when considering Alderson and Beretta’s (1992:58) assertion that it is “widely agreed by now that an evaluation should above all aim to promote short-run utilisation”; so, how can the evaluation of the programme provide timeous feedback to course developers and lecturers on which aspects of the intervention are working particularly well, and which aspects need to be rethought? Although a long-term evaluation might have several advantages, specifically in terms of determining the impact of the course in students’ later years of study, one of the main objectives of an evaluation must be to give course developers and lecturers more immediate and specific feedback, at the latest at the end of the intervention (for example at the end of a semester- or a year-course). It is at this point where data are likely to be most reliable; any evaluation after this period runs the risk of variables which are unaccounted for, and which might skew the interpretation of results.

Follow-up studies can always be done afterwards to complement initial impact evaluation.

As far as the number of test subjects is concerned, similar problems apply. On the one hand, having too few subjects might make interpretations from the results of statistical tests unreliable. On the other hand, more things can go wrong (and usually do) if the scale of the evaluation is too large (Beretta, 1992:9). Therefore, a sufficient number of participants must be selected to ensure reliable and valid statistical conclusions as well as saturation in qualitative research, but excessive numbers of participants must be avoided to ensure that the study remains contained.

A question in terms of reliability is whether insiders or outsiders should ideally perform an evaluation. There are advantages and disadvantages in both cases: insiders are trusted by those involved in the programme, and thus their evaluations might more likely be used and their recommendations implemented by those involved in the programme (Alderson & Beretta, 1992:60); in contrast, outsiders can provide a fresh perspective and stakeholders might perceive them as being less biased than insiders might be (Lynch, 2003:17; Alderson, 1992:281; Alderson & Beretta, 1992:60). It is important to note that, if outsider experts are selected to provide an objective view of the intervention, this “immediately implicates a paradigm that is essentially positivist” (Lynch, 2003:21).

One variable that must be controlled as far as possible is the teacher (or lecturer) variable. Beretta (1992:11) argues: “The teacher in Program A might be more highly qualified, more enthusiastic, or may differ in any number of ways from the teacher of Program B. This could offer a rival explanation of results, detracting from claims of the treatment program”. When evaluating the impact of an intervention, it might be useful to randomly select students from a variety of classes to counter the effect that one specific lecturer might have.

What is more, the context in which the evaluation takes place must be carefully considered. South African academic literacy courses, for example, often address different aspects than similar courses in other countries. One reason could be because many South African students’ Cognitive Academic Language Proficiency (CALP) (Cummins & Swain, 1986:152) might not be developed adequately in either their mother tongues or in

English to cope with the demands of higher education (Clarence-Fincham, 2000:145)⁹. This is usually not the case in countries like the United Kingdom or Hong Kong, where students' CALP in their mother tongues is usually already well established (Parkinson *et al.*, 2008:12). Since our academic literacy courses tend to be structured differently from those in other countries, it also makes sense that an evaluation design for assessing the impact of these courses should be tailor-made for the South African context.

Amount and type of data are also factors that should be taken into consideration. Newcomer *et al.* (2010:567) warn against trying to collect too much data, since once data piles up, "a pyramiding effect [occurs] in terms of data processing and analysis effort". Evaluators must also remember that it is possible to over-rely on a specific type of evaluation data. For example, Alderson and Scott (1992:44, 54) over-rely on perception data in their evaluation of a nation-wide project in English for Specific Purposes in Brazilian Federal Universities; the authors themselves state that their evaluation would have benefitted from additional data, such as standardised test data. The types of data that can be used to triangulate information when evaluating the impact of a specific intervention are outlined in Section 2.5.3, and discussed in detail in Chapter 4.

To conclude, this advice from Alderson (1992:274) must be kept in mind:

[T]here is [not] One Best Way of conducting an evaluation. (...) Much depends upon the purposes of the evaluation, the nature of the programme or project being evaluated, the individuals involved (...) and on the timescales and resources involved. This emphatically does not mean that 'anything' goes: it is essential that evaluations be conducted in a principled, systematic and explicit manner.

2.5.2.2 Difficulties when measuring the impact of an intervention

Engstrom and Tinto (2008:47) argue that many complex forces shape student success. As a result of such variables, which are often difficult to control for, doing research on impact, specifically in education, is challenging (Howes, 2003:148). These variables include the degree to which students are urbanised, the quality of the secondary schools that they attended, students' self-efficacy, their family circumstances, employment status,

⁹ Coetzee-Van Rooy's (2010:33-34) warning that the hypotheses posited by Cummins cannot be proven in the South African context before valid and reliable academic literacy tests have been developed for all the participants' first languages must be kept in mind. The hypothesis put forward here therefore remains theoretical, yet does not contradict that an evaluation design for assessing the impact of academic literacy interventions in South Africa should best be moulded for the South African context.

workload and even their intelligence (Prinsloo & Subotzky, 2011:7; Stephen *et al.*, 2004:44; McKenzie & Schweitzer, 2001). Other factors which can influence research into the impact of a specific intervention include the exposure that students get during their studies. Archer (2008:249), who attempts to assess the impact of a writing centre intervention on student writing, agrees:

[T]here are many factors affecting student writing other than visits to the Writing Centre, and it would be artificial to attempt to construct a ‘control group’. Students write in a range of courses, get feedback, do a range of reading, and it would be difficult to ascertain the extent to which one or two visits to the Writing Centre had impacted on their writing within this larger context.

This is also the case with interventions like academic literacy courses. Although it might be easier to attribute students’ acquired academic literacy abilities to an academic literacy intervention that is usually at least a semester or even a year long, rather than just consisting of “one or two visits”, there are still various factors at play.

Furthermore, as is the case in Archer’s study, forming a control group is not possible at most South African universities, as the majority of students have some kind of academic literacy intervention built into their curricula¹⁰. Lynch (2003:24) and Alderson (1992:283-284) agree that in many, if not most language interventions, it is impossible to find suitable comparison groups which can be subjected to either no treatment, or who can be required to participate in alternative interventions. Forming control groups would not only be “artificial” (cf. Archer, 2008), but would also bring along a myriad of ethical problems. Alderson (1992:283-284) argues that “[a]lthough constraints from the real world can be resisted or adapted in some circumstances, it is inevitable that they will strongly influence the evaluation methodology that will be adopted” .

If only the intervention group is available, certain quasi-experimental designs must be used that will be weaker than traditional experimental designs, specifically in terms of making causal claims about the intervention and its impact on student success (Lynch, 1996:73-74). In such a case, one approach that could be followed would be to have a pre-test and post-test design that uses multiple measures, in addition to other tools to measure

¹⁰ There are, however, some universities where not all students are required to complete academic literacy courses, or at least not the same number of academic literacy course modules as students who have been tested as being “at-risk” of failing due to low academic literacy levels. In such cases, control groups might be used effectively. There are also exceptions where the curriculum structure has allowed researchers to effectively use control groups.

impact, such as questionnaires concerning perceptions of the programme's impact that are completed by teachers and students (Lynch, 2003:25). By combining multiple perspectives with the basic pre-test post-test design (and thus triangulating data), one diverges from the traditional positivist research design by including this characteristic of interpretivist designs, thus moving towards a mixed evaluation (or postpositivist) research design (Lynch, 2003:27).

One of the foremost weaknesses of many impact studies is that they often only look at one or two factors that might indicate the impact of an intervention; for example, Cilliers and Herman (2010:254) mention that studies are often criticised for being limited to only “determining the satisfaction of participants, rather than anything beyond that”. This seems evident from the research discussed in Chapter 3. When only one aspect of impact is considered, various unanswered questions often remain, for example what elements contributed to the impact of the intervention, how this could be measured, and what the long-term effect of the intervention might be (Van Dyk *et al.*, 2011a). There are some studies that attempt to answer as many as possible of these questions, for example a study by Archer (2008:248) that uses a “multi-faceted approach [that enables] a holistic and contextualized picture (...) to emerge”. By using a variety of research instruments and perspectives, the research design is strengthened, and qualified conclusions become possible (Lynch, 1996:74). That is also the aim of the current study.

2.5.3 Methodology for evaluating impact

In the early 1980s, it was acknowledged that there is no single way of ‘doing’ evaluation. Instead, four general standards (or principles) were generally agreed upon, namely *utility* (identifying stakeholders and providing them with necessary information on time), *feasibility* (the workability of the research design in real world settings), *propriety* (whether the evaluation process adheres to ethical principles and recognises the rights of all who might be affected by the evaluation) and *accuracy* (whether the evaluation is valid and reliable, and whether there are logical links between the data and conclusions) (Joint Committee on Standards for Educational Evaluation, 2014; Newcomer *et al.*, 2010:7). These four standards are still widely used today, and were kept in mind in the development of the evaluation design proposed in this study (see Section 1.3).

Generally, it would seem wise to include both quantitative and qualitative components when evaluating a programme (Lynch, 1992:94), since no single methodology can give a holistic overview when it comes to evaluating a language programme (Beretta, 1992:19). According to Lynch (1992:93), “[p]rogram evaluation will probably always require a traditional, quantitative component”, especially for summative evaluation. Using qualitative data in addition to quantitative data, however, can help clear up ambiguities. Dudley-Evans (1995:128) agrees, and states that quantitative data can help to answer “what” questions, (for example “what is taught”), whereas qualitative data address “how” and “why” courses work, or do not work. Thus, linking quantitative and qualitative data can be “illuminating” (Lynch, 1992:94). Ross (2009:776) concurs that it is more likely to obtain evidence that an intervention has had an effect by following a mixed mode approach. Such an approach, then, can serve as an alternative to a traditional control group experiment in situations where such experimental designs are not feasible.

As far as quantitative data collection goes, many researchers are against the use of norm-referenced tests, since these are not sensitive to the teaching and learning which occur in various classrooms (Lynch, 1996:74, 92; Lynch, 1992:66). Some educators have instead favoured criterion-referenced tests (Lynch, 1996:74; 93; Lynch, 1992:66). Both these types of tests have advantages. Norm-referenced tests (see Section 4.4.1.1) make comparisons with external courses or programmes possible, and are generally more reliable than criterion-referenced tests that are often just designed for a specific intervention (Lynch, 1992:66). Criterion-referenced tests, on the other hand, allow the evaluator to assess criteria that are specific to the particular intervention being evaluated, and make it possible to ignore criteria that are not part of the intervention content or objectives. These tests are particularly useful for discipline-specific EAP interventions (see Section 4.4.1.2).

Alderson (1992:284-285) reminds us that tests are only one part of programme evaluation, and that it is just as important to triangulate the outcomes of other methods with the results of test data. Though quantitative data are clearly valuable in any evaluation study, this type of data can also play an important role in evaluating courses. One of the greatest criticisms against qualitative data is that such data cannot always be proven to be reliable. However, one of the major strengths of qualitative data is validity (Lynch, 1992:69). A major drawback of this type of research is that it is very time-

consuming; therefore, it is recommended that there must always be an explicit purpose for any data that are collected qualitatively (Lynch, 1992:93).

Another important principle in terms of methodology to keep in mind when evaluating the impact of an intervention is triangulation.

The notion of triangulation is particularly important in evaluation. Given that there is No One Best Method for evaluation, it makes good sense to gather data from a variety of sources and with a variety of methods, so that the evaluator can confirm findings across methods. But in evaluation, outcomes and effects are often contentious, and the perception of these may vary considerably depending upon the perspective, vested interest or even personality of the participants. In addition, the views of stakeholders as to the appropriacy of particular methods for evaluation typically also differ radically. In such settings, it does not seem sensible to rely on one method for the purposes of data gathering, but rather to try to complement or neutralise one method with another. Thus, a generalisable recommendation for evaluations would be that they should wherever possible plan to triangulate in method (Alderson, 1992:285).

The value of triangulation has been acknowledged by researchers for decades (cf. Lynch, 1996:59). Denzin (1989:234) defines triangulation as “the combination of methodologies in the study of the same phenomena”. Researchers can triangulate by methodology (for example using various qualitative and quantitative tools) as well as by data sources (for example obtaining data from students, lecturers and the literature) (Lynch, 1996:59; Denzin, 1989:237). Beretta (1992:19) adds that programme evaluators should ideally “select opportunistically from a wide range of research methodologies” with the aim of providing information that would be utilisable in the short run. According to Jick (1979:602), triangulation is a “vehicle for cross validation when two or more distinct methods are found to be congruent and yield comparable data” (cf. Lynch, 1996:15). He adds that multiple independent measures which reach comparable conclusions lead to a more certain portrayal of a phenomenon (Jick, 1979:602). Lynch (1996:60) concurs, and even states that triangulation would be an “obvious strategy for strengthening the validity of evaluation findings”. He further argues that any specific technique or source could, potentially, be inherently biased. However, if various sources are used, this could assist in cancelling out such bias (Lynch, 1996:60). According to Patton (2002:93), the accuracy and credibility of findings in a postpositive framework are increased if data is triangulated by means of sources as well as analytical perspectives. In this study, the approach of triangulation will be followed to evaluate the effectiveness of a specific academic literacy intervention at a South African university.

Finally, a design for assessing the impact of an academic literacy intervention could arguably only be generalisable if other universities with similar programmes “buy into” the evaluation design. The best way for this to happen is for the process to be participatory. Alderson and Scott (1992:39-41) mention the following levels of evaluation where participation might be possible: designing the evaluation, constructing the instruments, testing out the instruments, collecting data, analysing data, drafting the final report, and reading and learning from the report. As indicated in Section 1.3, participation in the proposed evaluation design was made possible through the options discussed above.

2.5.3.1 Types of assessment in academic literacy courses

In language studies, certain theoretical concepts, called psychological constructs, are under investigation. According to Brown (2001:16), “such constructs are the underlying psychological traits or characteristics that are operating in the participants’ brains”. Since these psychological constructs cannot be directly observed, the constructs must be operationalised by using specific means of quantifying these constructs in numerical terms before they can be used in quantitative research studies; these means include tests and questionnaires (Brown, 2001:16-17).

In order to ensure triangulation by method, it would seem prudent to draw on various types of assessment. The current study proposes using a combination of assessment techniques to obtain a holistic picture of the impact of academic literacy interventions. Keeping in mind the need for triangulation of data discussed in Section 2.5.3, the three main methods identified from previous studies in the field (further discussed in Chapter 3) are now discussed in detail.

2.5.3.1.1 Valid and reliable academic literacy tests

Standardised academic literacy tests are possibly the most widely used instrument to assess students’ academic literacy levels. They are especially well-suited for assessing reading abilities and the many academic literacy abilities that go hand-in-hand with reading (cf. Grabe, 2009:450-451). Valid and reliable academic literacy tests have been proven to positively correlate with academic performance. For example, Cliff and Hanslo

(2009:272) found that students with a PTEEP score of higher than 60% were at low risk of underperforming in their studies, whereas students with a PTEEP score of under 40% were at risk of underperforming or even failing at the end of their first year. Their conclusion is that “higher test performance seems associated with higher mean academic performance”. Studies that have used standardised language proficiency tests include those by Petersen-Waughtal and Van Dyk (2011), Van Dyk *et al.* (2011a); Wait (2007) and Song (2006).

Valid and reliable academic literacy tests have numerous advantages (see, for example, Seabi, 2012:93; Nitko & Brookhart, 2011:346; Heaton, 1988:167; Beretta, 1986:434). Firstly, due to their format, it is possible to easily mark large quantities of tests (often with the help of computer software). Secondly, it becomes easier to replicate research than had non-validated and non-verified tests been used. Thirdly, the materials, scoring instruments, administration and other procedures are usually fixed and uniform, meaning that even if the assessment is done in different places and times, it will still occur in the same manner. Fourthly, these tests have been subjected to rigorous developmental cycles. Finally, because the test items do not come directly from the specific courses that are being investigated, but are rather based on predefined test constructs that are generally thought to be universal in nature, these tests can be considered to be impartial.

The current study proposes using the TALL as part of an evaluation design with the aim of assessing the impact of academic literacy interventions. There are several reasons why the TALL is a judicious option for assessing students’ academic literacy levels. Firstly, it has a well-developed test construct which considers some of the most important aspects of academic literacy, and which closely reflects students’ ability to produce and interpret academic discourse (Van Dyk, 2005:43). As recommended by Bachman and Palmer (1996:75-76), the test construct also relies on language tasks rather than language skills (see Section 2.4.2).

The TALL’s test construct is based on the following abilities:

- understand a range of academic vocabulary in context;
- interpret and use metaphor and idiom, and perceive connotation, word play and ambiguity;

- understand relations between different parts of a text, be aware of the logical development of (an academic) text, via introductions to conclusions, and know how to use language that serves to make the different parts of a text hang together;
- interpret different kinds of text type (genre), and show sensitivity for the meaning that they convey, and the audience that they are aimed at;
- interpret, use and produce information presented in graphic or visual format;
- make distinctions between essential and non-essential information, fact and opinion, propositions and arguments; distinguish between cause and effect, classify, categorise and handle data that make comparisons;
- see sequence and order, do simple numerical estimations and computations that are relevant to academic information, that allow comparisons to be made, and can be applied for the purposes of an argument;
- know what counts as evidence for an argument, extrapolate from information by making inferences, and apply the information or its implications to other cases than the one at hand;
- understand the communicative function of various ways of expression in academic language (such as defining, providing examples, arguing); and
- make meaning (e.g. of an academic text) beyond the level of the sentence.

(Weideman, 2003a:xi)

This test construct is similar to those of other academic literacy testing instruments. The PTEEP, for example, assesses abilities as represented in Table 2.1.

Furthermore, the TALL has been proven to be highly reliable. “Reliability refers to the consistency of test results” (Nitko & Brookhart, 2011:40). Davies *et al.* (1999:168) define it as the “level of agreement between the results of one test with itself or with another test”. Using Cronbach’s equation, the TALL shows an average reliability measure of 0.94 across several versions of the test between 2004 and 2010 (Van Dyk, 2010). Considering that perfect reliability would be represented as +1.0, the TALL boasts an excellent reliability level.

In addition, the TALL is viewed as a valid instrument (Van Der Walt & Steyn, 2007:146-156) which is standardised in terms of its scoring and administration (Petersen-Waughtal & Van Dyk, 2011:105; Van Dyk *et al.*, 2009:335). The TAG (Toets vir Akademiese Geletterdheid, the Afrikaans version of the TALL) has proven to have internal validity in terms of content validity, construct validity as well as face validity (including intra-test and inter-test validity, internal consistency, domain-specific validity and group

differential validity) (Van Dyk, 2010:199-261). It has also been shown to have external validity in terms of concurrent validity, predictive validity, and consequential validity (Van Dyk, 2010:261-284). Van Dyk (2010:261-284) argues that this would also be the case for the TALL – this has been corroborated by smaller, unpublished experiments. Fourthly, the TALL’s broad distribution of marks (indicated by the standard deviation and inter-quartile range) makes it easier to divide students into groups based on ability (Van Rooy & Coetzee-Van Rooy, 2015:43).

Table 2.1: Academic literacy abilities assessed in the PTEEP

Vocabulary	Students’ abilities to derive/work out word meanings from their context.
Metaphorical expression	Students’ abilities to understand and work with metaphor in language. This includes their capacity to perceive language connotation, word play, ambiguity, idiomatic expressions, and so on.
Extrapolation, application and inferencing	Students’ capacities to draw conclusions and apply insights, either on the basis of what is stated in texts or is implied by these texts.
Understanding the communicative function of sentences	Students’ abilities to ‘see’ how parts of sentences/discourse define other parts; or are examples of ideas; or are supports for arguments; or attempts to persuade.
Understanding relations between parts of text	Students’ capacities to ‘see’ the structure and organisation of discourse and argument, by paying attention – within and between paragraphs in text – to transitions in argument; superordinate and subordinate ideas; introductions and conclusions; logical development.
Understanding text genre	Students’ abilities to perceive ‘audience’ in text and purpose in writing, including an ability to understand text register (formality/informality) and tone (didactic/informative/persuasive/etc.).
Separating the essential from the non-essential	Students’ capacities to ‘see’ main ideas and supporting detail; statements and examples; facts and opinions; propositions and their arguments; being able to classify, categorise and ‘label’.
Understanding information presented visually	Students’ abilities to understand graphs, tables, diagrams, pictures, maps, flow-charts.
Understanding basic numerical concepts	Students’ abilities to make numerical estimations; comparisons; calculate percentages and fractions; make chronological references and sequence events/processes; do basic computations.

Adapted by Cliff and Hanslo (2009:269) from Bachman and Palmer (1996) and Yeld (2001)

Finally, from a logistical stance, it is a good test choice. The TALL, which is comprised exclusively of multiple choice questions, takes only one hour to write (Van Dyk, 2005:44). In addition to the relatively short time in which it is administered, its scoring is also objective, time- and cost effective (Van Dyk, 2010:17; 2005:44). Furthermore, due to the format of the test, it is possible to maintain similar conditions for pre- and post-tests; this makes it possible to control as many as possible external variables.

Many standardised academic literacy tests have limitations though. Firstly, these tests usually focus on assessing students' reading abilities at the expense of other academic literacy abilities. For example, students are not required to produce any written text in the test (mainly because of the practical constraints in administering the test) (Butler, 2009:294). Butler (2013:83) cautions that "it would be irresponsible to claim benefits for writing if writing ability is not explicitly assessed" in a test. Secondly, again due to practical constraints, but also to ensure fair and standardised marking, these tests tend to rely on multiple-choice questions. Arguments against purely relying on multiple choice questions have, however, been raised. One such argument is that this format is not appropriate as it is not able to test the normal language processing abilities of students (Scouller, 2012:469; Murphy & Yancey, 2008:366-368). Thirdly, standardised tests are not developed with a specific intervention in mind; therefore, "it is difficult to match results perfectly with instructional objectives" (Henning, 1987:8). In part due to the above limitations, but also to comprehensively and effectively assess the impact of an academic literacy intervention, it becomes necessary to consider supplementary assessment methods.

2.5.3.1.2 Assessing writing assignments

Many students have great difficulty with writing at the level required in higher education (Radloff & De La Harpe, 2000). Yet it is one of the primary means of assessment in higher education, influences their overall academic performance, and is also linked to success in their future professions (Archer, 2008:248; Weigle, 2002:4). "The teaching of writing (...) practices in Higher Education is thus inextricably linked to student success, which includes both retention and throughput" (Archer, 2008:248). Van Dyk *et al.* (2009:334) agree: "[a]s students are largely evaluated on their written work, an ability to write well plays an important role in a student's overall success in his or her academic

work”. If one acknowledges the importance of writing abilities in higher education, it then follows that in order to measure the impact of an academic literacy intervention, students’ writing abilities would have to be assessed.

According to Weigle (2002:5), “[w]riting and critical thinking are seen as closely linked, and expertise in writing is seen as an indication that students have mastered the cognitive skills required for university work”. Weigle (2002:17) further argues that “correctness in writing” is seen as more important than accuracy in speaking in academic contexts, since the academic discourse community is usually accessed through writing. Grabowski (1996:75) adds: “Writing, when compared to speaking, can be seen as a more standardized system which must be acquired through special instruction. Mastery of this standard system is an important prerequisite of cultural and educational participation”. Although the idea of one “standard” writing system might well be questioned by many (is the type of writing used in the sciences, for example, the same as that used in law?), it is probably safe to say that the system(s) of writing is (are) more standardised than those of speaking, and can therefore be assessed with greater accuracy.

Students’ writing abilities have been assessed in a variety of ways. In a study by Song (2006), for example, students were expected to submit a portfolio of various examples of essays in different genres in addition to completing an in-class essay assignment. Van Dyk *et al.* (2009) examined students’ completed pre-, mid- and post-intervention writing assignments. In a study by Carstens and Fletcher (2009b), students wrote a pre-test and a post-test in the form of a 50-minute in-class essay. The assessment of such writing assignments is rarely standardised. Each of these studies seems to use different criteria as well as measuring tools in such assessment. To determine how to best assess an improvement in student writing (and the abilities which accompany academic writing), some of the research in this field is considered below.

Writing assessments should ideally adhere to the following principles. Firstly, students should be provided with a single, clearly defined topic that will not only motivate them to write, but will also guide them in the writing process (Heaton, 1988:137, 144). A single topic will provide the marker with a “common basis for comparison and evaluation” (Heaton, 1988:138). A wide general knowledge, imagination or creativity should preferably not be required for students to be able to address the provided topic, as certain

students could otherwise be unjustly advantaged – these are, after all, not usually part of the outcomes of language or academic literacy assessment (Hughes, 2003:90). A topic about which the majority of candidates will likely have a similar background knowledge should be provided instead. To provide such a suitable topic, the evaluator should consider factors such as students’ educational background, age, and sex to ensure that the assessment will not be in favour of, or biased against, specific groups of students (Weigle, 2002:46). In cases where background knowledge is required, students could be provided the same background reading so as to ensure that students are able to draw from a similar knowledge-base. Information should also be provided about the audience of the written text (Heaton, 1988:137). Misunderstandings between the student and the evaluator’s intent can be avoided by following these guidelines on the topic, purpose for writing and audience.

Furthermore, the strict time limits that exist when students are required to write a text in a classroom situation might result in artificial and unrealistic writing assessments. In real-life situations (including scenarios such as submitting an assignment, report or essay for a mainstream subject), students engage in processes such as creating drafts and editing their work – processes which are usually also taught and promoted in academic literacy interventions (Heaton, 1988:138). Not allowing students to engage in such processes could therefore prove to be counter-effective in determining the impact of an intervention on students’ writing abilities. Allowing students the same amount of time to plan, draft and write their written text as they would have been allowed in a real-life situation (e.g. when writing an assignment for one of their subject courses) is thus advisable¹¹.

Ultimately, the assessment must be appropriate for the use for which it is intended. Bachman and Palmer (1996:17) outline six qualities that need to be kept in mind when assessing the “usefulness” of a test. These are reliability, construct validity, authenticity, interactiveness, impact, and practicality. These qualities need to be balanced against each other depending on the purpose of the test and the situation – in practise, all of them

¹¹ A disadvantage of allowing students to write assignments at home is that the researcher cannot control for independent work. However, this is the case with many continuous assessment tasks in academic literacy classrooms. If assessing the impact of an intervention on students’ academic literacy levels is the aim of such a writing task, and the sample is big enough, it might be worthwhile to risk that one or two students might not submit independently written work, if the trade-off is a more authentic end result.

cannot receive equal attention (Weigle, 2002:48). These six qualities are briefly discussed here.

Reliability “can be defined as consistency of measurement across different characteristics or facets of a testing situation (...). A test is said to be reliable if individuals receive the same score from one prompt or rater to the next, or if a group of examinees is rank-ordered in the same way on different occasions” (Weigle, 2002:49). A writing test might be made more reliable by paying special attention to the task itself (e.g. the topics and instructions) as well as variables related to scoring. As far as scoring is concerned, it should be remembered that markers may be unreliable; they are firstly inconsistent in their own marking, and secondly often disagree with colleagues on the merits of specific students’ writing (Heaton, 1988:144). One of the best ways of avoiding inconsistencies in marking is by using a marking scheme, or rubric, to assess student writing (Heaton, 1988:148). Rubrics are particularly useful for separating various features of a piece of writing for scoring purposes (Heaton, 1988:148).

Construct validity refers to “the meaningfulness and appropriateness of the interpretations that we make on the basis of test scores” (Bachman & Palmer, 1996:21). Construct validity is demonstrated if, firstly, the assessment elicits the type of writing that the assessor wants to test; secondly, all the “components of writing that are included in the definition of the construct” are reflected in the scoring criteria; and thirdly, the assessors adhere to these criteria when assessing the writing samples (Weigle, 2002:51); again, a standardised and comprehensive writing rubric would be useful in ensuring construct validity.

Authenticity is defined as “the degree of correspondence of the characteristics of a given language test task to the features of a target language use task” (Bachman & Palmer, 1996:23). This means that the type of writing task students are given in the writing assessment must be similar to writing tasks that students might be expected to complete in the academic literacy intervention in question. Authenticity should be carefully considered when setting the topic and genre for the writing task.

Interactiveness refers to the extent to which the test taker’s characteristics (in the case of language use, these may include language knowledge, being able to manage linguistic

and cognitive resources, and topical knowledge) are engaged in the assessment task (Weigle, 2002:53; Bachman & Palmer, 1996:25). An example of a task with low interactiveness would be a task in which the verbs in a paragraph must be rewritten in the past tense, whereas a task that is highly interactive would be, for example, an essay topic that interests the test-taker, and challenges him/her to set goals for the writing task, assess and analyse the various facets of the task, and plan the task (Weigle, 2002:53).

Impact refers to the effect that assessments have on individuals as well as systems (for example the course, educational system or society at large). Here, the consequences of taking the test must be kept in mind (Weigle, 2002:53-54; Bachman & Palmer, 1996:29). For example, is the test used to exclude some students from university studies, or will students be required to take additional modules as a result of the test?

Finally, practicality, namely the balance between the resources that are required for the administration and assessment of the test, and the resources that are actually available, must receive due consideration (Weigle, 2002:54; Bachman & Palmer, 1996:35). For example, although testing students by means of a portfolio might have several advantages, the appropriate resources (e.g. time, money, markers) for this type of testing might not be feasible. However, a single, extended writing task might be assessed thoroughly with the resources that are available. Furthermore, although assessing the entire student population participating in a certain intervention might seem ideal when assessing the impact of that intervention, selecting a smaller sample might still provide valid results, yet would put a significantly lower strain on resources – in fact, as Beretta (1992:9) warns (see Section 2.5.2.1), things are more likely to go wrong if the scale of the evaluation is too large.

2.5.3.1.3 Perceptions of impact

According to Carstens and Fletcher (2009b:324), it is not enough to only measure whether there was an empirical improvement in students' academic literacy marks when determining whether an academic literacy intervention was successful. The success of such interventions, they argue, is equally dependent “on students' experience, which are co-determinants of motivation and skills transfer”. It is particularly important to determine whether students consciously transfer abilities acquired in the academic

literacy intervention to their content-subjects. One way of doing that is by asking students about their behaviours. “Asking questions is widely accepted as a cost-efficient (and sometimes the only) way of gathering information about past behaviour and experiences, private actions and motives, and beliefs, values and attitudes” (Foddy, 1993:1).

Babbie (2005:253) states that surveys are “excellent vehicles for measuring attitudes and orientations”. Wentland (2011:541) agrees, and adds that surveys act as effective research tools when respondents are the best source of the information that is needed. One way of surveying participants to gain insight into their opinions and views is by means of interviews (Brown, 2001:6). However, if the researcher deems it worthwhile to explore such issues on a larger scale than is possible with interviews, questionnaires tend to be a more efficient research tool (Babbie, 2005:253; Brown, 2001:6). “Questionnaires are any written instruments that present respondents with a series of questions or statements to which they are to react either by writing out their answers or selecting from among existing answers” (Brown, 2001:6). The main objective of a questionnaire is to gather facts and opinions about a particular phenomenon from people who have knowledge about the specific issue (De Vos *et al.*, 2011:186). Therefore, when determining whether students use the abilities acquired in an academic literacy intervention, one method of getting valid results will be to ascertain students’ own opinions on whether they use academic literacy abilities in their content subjects. Lynch (1996:169) states that one can approach student attitudes towards a programme in two ways: 1) from a judgemental viewpoint, where the degree of students’ satisfaction with the programme is measured, and 2) in a description fashion, where the evaluator aims to understand the nature of students’ satisfaction and/or dissatisfaction.

Few studies (not only in the field of academic literacy, but also in the broader field of education) seem to discuss students’ perceptions of whether course outcomes were successfully achieved and transferred to other settings. Those that do touch on this topic generally make use of control groups so as to compare student perceptions (see, for example, Gibbs & Coffey, 2004). Another method that has been used in the field of education has been to use pre- and post-intervention surveys (Wallace, Shorten & Crookes, 2000) to determine students’ perceptions of improved abilities over the duration of the intervention. Although this might be effective in some fields, for example with

information literacy skills (as described by Wallace *et al* [2000]), this method would not necessarily be effective in a field like academic literacy, as this field is often so unfamiliar to many students that they are not able to express their abilities in the field before being exposed to the content and terminology of the intervention.

Some of the studies that attempt to assess students' perceptions (for example Van Dyk *et al.*, 2011a; Carstens & Fletcher, 2009b; Storch & Tapper, 2009; Van Dyk *et al.*, 2009; Archer, 2008) use in-house official feedback forms to assess students' perceptions regarding the intervention. These feedback forms are usually aimed at assessing interventions and lecturers. Although there is clearly worth in such feedback forms, they do not allow for detailed information regarding students' opinions about the value of various abilities addressed in the intervention.

One type of survey that seems particularly relevant to determining perceptions regarding transfer is called a "need-press" interaction analysis (Rao & Saxena, 2014:2; Lynch, 2003:68). In this type of survey, participants "are asked to judge the importance (need) of particular language skills or abilities, and then to judge their emphasis (press) in the teaching programme. By comparing judgements of how important something is perceived to be with how much attention it receives in the instructional setting, areas of individual learner development and programme objectives that may need improvement are identified" (Lynch, 2003:68). Students can therefore be asked to which extent they 'need' certain language abilities in other subjects by the end of their first year, and again by the end of their second year. By combining this with students' perceptions of the extent to which these abilities were emphasised in the academic literacy classroom, valuable deductions could be made regarding important abilities that were transferred (abilities which both score highly on the 'need' and 'press' aspects), abilities which are needed but not taught sufficiently (abilities which score highly on the 'need' aspect, but low on the 'press' aspect), and abilities which are taught, but not needed (abilities which score low on the 'need' aspect, but high on the 'press' aspect). The researcher must be cautious here: it is possible that students feel that certain aspects taught in an intervention are unnecessary, while lecturers disagree. Some abilities serve as building blocks for more complex learning later on, and students are not always the best judges as to which abilities will be necessary for them to achieve more complex tasks later on. In cases such

as these, additional feedback should preferably be obtained from academic literacy as well as subject experts.

Before going through the process of constructing a new questionnaire, it would be wise to first determine whether an adequate information-gathering tool already exists, or whether the necessary data already exist in some kind of format (Brown, 2001:2). However, where this is not possible, various factors should be kept in mind when constructing a questionnaire. Firstly, characteristics of effective questionnaires should be kept in mind when constructing questionnaires: items must be clear, objective and neutrally stated; double-barrelled questions should be avoided; questions should be relevant to the participants; questions should be short; negative items should be avoided; appropriate information sources (or stakeholders) for the questions should be identified; and appropriate sampling methods, data collection modes and quality assurance procedures should be determined (Newcomer & Triplett, 2010:265; Babbie, 2005:255-259; Brown, 2001:8-12).

Secondly, it should be kept in mind that questionnaires can consist of open responses (respondents create an answer using their own words) and closed responses (respondents select from pre-existing optional answers for each question) (Brown, 2001:35-36). Using scaled questions (where respondents mark a specific point on a scale) is effective when information is needed about subjective aspects of a phenomenon (De Vos *et al.*, 2011:200). Likert-scale questions are particularly effective if the researcher is interested in respondents' opinions and attitudes about language-related issues (Brown, 2001:40-41).

Thirdly, the researcher decides which stakeholders might act as the best data sources (Newcomer & Triplett, 2010:267-268). Examples of relevant stakeholders in an impact study on academic literacy interventions might be students, content-subject lecturers and academic literacy lecturers. Once relevant stakeholders have been identified, it is advisable to identify a sample; sampling is important in the research process since it saves the researcher unnecessary effort and resources without compromising the outcomes of the research (Brown, 2001:71). As far as selecting the sample is concerned, Stoker (in De Vos *et al.*, 2011:225) provides guidelines for sampling: for example, in populations of 20, 100% of participants must be sampled; in populations of 50, 64% of

participants must be sampled; in populations of 100, 45% of participants must be sampled; in populations of 200, 32% of participants must be sampled, and in populations of 500, 20% of participants must be sampled.

In a class setup, the best way of collecting data might be group-administered questionnaires, where a questionnaire is given to each member of a larger group to complete; instructions, however, are given to the group at large (De Vos *et al.*, 2011:189). If this is not possible, another option is distributing electronic questionnaires, where the survey is e-mailed to respondents or where respondents complete a web-based survey online (De Vos *et al.*, 2011b:189). Newcomer and Triplett (2010:269) state various advantages and disadvantages of surveys that are administered in person and via the Internet. Surveys collected in person, for example, require a lot of resources (for example time, money and staffing), but make it easier to collect all questionnaires that were distributed. Surveys completed via the Internet, in contrast, require fewer resources and make it possible to send out surveys to large numbers of participants, but are very difficult to collect (i.e. participants might decide not to return surveys). When sending out questionnaires, it is recommended that a description of the study accompany the questionnaire (Newcomer & Triplett, 2010:284). In the case of web surveys, reminders should be sent after short intervals (or as soon as the return rate declines significantly) (Newcomer & Triplett, 2010:284).

Data gathering goes hand in hand with data analysis. When analysing data, De Vos *et al.* (2011:196) recommend that mechanical and electronic facilities be used wherever possible. Data from closed responses need to be coded if they are not already in a coded format (Brown, 2001:93). A data coding approach, such as a rows-and-columns approach (where rows represent the respondent's answers to all the questions, and columns represent all the respondents' answers to a specific question), needs to be decided upon (Brown, 2001:95). Electronic spreadsheets such as Excel or statistical programmes such as SPSS can then be used to statistically analyse data (Brown, 2001:94-97). The most common type of statistics that are derived from survey data are descriptive statistics (Brown, 2001:114). Open responses should first be transcribed (preferably into an electronic format); thereafter, categories should be decided upon based on themes emerging from the responses (Brown, 2001:99-100). Brown (2001:213-230) describes three basic steps when analysing qualitative data: firstly, data need to be reduced in order

to make the data manageable; secondly, data should be displayed so as to analyse and synthesise information; and thirdly, conclusions based on these data should be drawn and verified.

A possible objection might be that information might not be entirely reliable if students are asked how they experienced a programme or intervention; however, “these recalled experiences may have a legitimate place in constructing an (...) account of the meaning of the programme (Lynch, 2003:131).

2.6 Conclusion

[W]e need to consider the practical implications for the successful implementation of [academic literacy interventions]. Again, the ultimate ‘success’ of such interventions depends on whether we can present conclusive evidence on their impact. (Butler, 2013:84)

In answer to Butler’s statement, this study intends to propose an evaluation design for effectively and comprehensively measuring the impact of academic literacy courses. The design will be aimed at assessing the impact of a wide variety of academic literacy interventions, and not just those aimed at English Second Language (ESL) speakers, as has been the case in much of the past research in this field (Holder *et al.*, 1999:21). The aim will be to develop an evaluation design that could be used to assess the academic literacy abilities that are needed during students’ higher education studies. As mentioned in Chapter 1, this study will use a post-positivist research paradigm, quantitative and qualitative methodology, and a variety of methods to ensure a multi-faceted approach that enables a holistic, contextualised picture to emerge.

The main contribution of the study will be the development of a comprehensive and informed design for assessing the impact of academic literacy interventions – this design will be applied to two modules of an academic literacy course to assess its usefulness and practicality. The outcomes of implementing such a design will be manifold. Firstly, academic literacy intervention coordinators will be able refine their interventions based on feedback received by applying the design. As De Klerk *et al.* (2006:149) argue, in addition to being reflective, impact studies should also provide specific guidelines as to how programmes or interventions could be restructured so as to meet identified

challenges. Where courses are shown to have a significant impact, it will become easier to justify the resources necessary to present such courses meaningfully. Cross-pollination would also be possible in that academic literacy interventions would have the opportunity to learn from both the successes and the challenges of similar programmes. Furthermore, the field of academic literacy would ultimately be strengthened if the impact of more academic literacy interventions were to be effectively measured.

While Chapter 2 has provided a general overview of the literature regarding specific key concepts to this study, students' underpreparedness for higher education studies, academic literacy and programme evaluation, Chapter 3 focuses specifically on reviewing studies that have attempted to determine the impact of various academic literacy interventions. Thereafter, a preliminary evaluation design is proposed (Chapter 4). The evaluation design is then verified in a specific context (Chapters 5 and 6) and is subsequently validated (Chapter 7).

CHAPTER 3

TOWARDS IMPACT MEASUREMENT: AN OVERVIEW OF APPROACHES FOR ASSESSING THE IMPACT OF ACADEMIC LITERACY ABILITIES¹² (ARTICLE 1)

3.1 Prelude to Chapter 3

Chapter 1 outlined the objectives and provided an overview of the current study. Chapter 2 provided a literature review of concepts that are important for the context of the current study. In particular, it focused on the concepts of impact, academic literacy and programme evaluation. Many of the concepts discussed in Chapter 2 are returned to in the following chapters.

The main aim of Chapter 3 is to provide a broad overview of both national and international studies which have attempted to determine the impact of academic literacy (or similar) interventions. The strengths as well as weaknesses of each of these studies are examined, and broad trends in the field of assessing the impact of academic literacy interventions are identified.

3.2 Introduction

It is generally acknowledged that the South African secondary education system does not sufficiently prepare students for higher education studies (Cliff, 2014:322; Van Dyk *et al.*, 2009:333; Higher Education South Africa, 2008:3). One consequence of this is poor university throughput rates, with as many as 55% of all enrolled students leaving university without graduating, and only 27% of all students graduating their 3- and 4-year qualifications in the prescribed time (Scott *et al.*, 2013:43). A prominent factor identified among students who are underprepared for higher education studies is a low level of academic literacy. Researchers almost unanimously agree: adequate academic literacy (which includes, but is not limited to, language proficiency) is crucial to students

¹² Published as: Fouché, I. 2015. Towards impact measurement: An overview of approaches for assessing the impact of academic literacy abilities. *Stellenbosch Papers in Linguistics*, 44:19-35. Minor editorial changes were made for the sake of the format of this dissertation.

being successful in their studies (Terraschke & Wahid, 2011:173; Davies, 2010:xi; Defazio *et al.*, 2010:34; Leibowitz, 2010:44; Archer, 2008:248).

Based on the aforementioned research, it would seem that a large number of students needs academic literacy support. This number grows each year, mainly due to the massification of higher education which inevitably implies more underprepared students gaining access (see Calderon [2012] and Teichler [1998] for a comprehensive discussion of this trend). Yet universities seem to have fewer and fewer resources available each year (see, for example, Hornsby & Osman, 2014:712-713; Kwiek *et al.*, 2014:6). This is possibly why “attention has shifted [in recent years] from an almost exclusive focus on access to include a concern with graduation rates and with general efficiency and quality matters” (Yeld, 2010:26). In order for the existence of academic literacy programmes to be justified in this resource-scarce higher education environment, they need to be able to show that they have a real and worthwhile impact on student success.

Impact assessment falls under the umbrella term of “programme evaluation” (De Vos *et al.*, 2011:453) – see Section 2.2.5. Situating it in an educational context, Brown (2001:15) defines programme evaluation as “the ongoing process of data gathering, analysis, and synthesis, the entire purpose of which is constantly to improve each element of a curriculum on the basis of what is known about all of the other elements, separately as well as collectively”. De Vos *et al.* (2011:449) argue that “[i]n an age of accountability, [stakeholders] demand that some evidence is provided in terms of ‘what works’, ‘how it works’ or ‘how it can be made to work better’”. This seems to be especially true in the resource-scarce South African higher education environment where the majority of students need effective academic literacy support. Academic literacy support can only be made more effective if we can determine which abilities are acquired most effectively by students and what academic literacy specialists are doing right to facilitate this acquisition, in addition to which abilities are not being acquired optimally. Only when academic literacy specialists can identify the weak points in a curriculum can they strive to responsibly improve their interventions.

When evaluating language programmes, Lynch (2003:1) points out that the areas of language assessment and programme evaluation usually overlap in that data from language assessment are often used as part of programme evaluation in order to make

decisions and judgements, reflect, and ultimately take certain actions. Bachman and Palmer (2010:21) agree: “Evaluation involves making value judgments and decisions on the basis of information, and gathering information to inform such decisions is the primary purpose for which language assessments are used”.

Two main specific purposes of programme evaluation in educational contexts are firstly to determine whether the programme is achieving its objectives, and secondly to determine which links exist between the processes of the specific programme and students’ achievement (Lynch, 2003:2). By doing this, it should be possible to determine how effective specific components of the intervention are (Lynch, 2003:7), and thus to find ways of improving the programme being evaluated (Newcomer *et al.*, 2010:6). In fact, argues Brown (2001:15), one should probably view the evaluation process as an ongoing needs assessment so as to constantly improve the programme in question (cf. Bachman & Palmer, 2010:25). These ideas are revisited in Chapter 4, where a preliminary evaluation design is proposed.

As mentioned in Section 2.5.2, impact assessment addresses a very specific, though central, facet of programme evaluation (De Vos *et al.*, 2011:453). Programme evaluation could include a myriad of factors, such as cost-effectiveness and work satisfaction of teachers and lecturers. While evaluating the impact of a programme or intervention would almost always be part of programme evaluation, impact assessment is a distinct facet that needs to be examined separately.

It should be kept in mind that there are various challenges to determining the impact of academic literacy interventions. One such challenge is that these interventions come in all shapes and sizes, for example, generic interventions, subject-specific interventions, undergraduate interventions, postgraduate interventions, reading interventions and writing centres – examples of all of these are discussed later in this chapter. This wide variety of academic literacy interventions makes it difficult, if not impossible, to assess impact by using a uniform approach. A further challenge is that the use of control groups, which would be part of traditional experimental designs, is often unfeasible in the South African context where, increasingly, all students (at least at first-year level) are required to participate in some type of academic literacy intervention. Many studies attempting to

assess the impact of academic literacy interventions must thus find other ways of providing reliable and valid results.

For the purposes of the current study, the terms “impact” and “effect” will be viewed as synonymous. De Graaff and Housen (2009:727) define these as “any observable change in learner outcome (knowledge, disposition or behavior) that can be attributed to an instructional intervention (possibly in interaction with other, contextual variables)”. An intervention’s effectiveness, then, “refers to the extent to which the actual outcomes of instruction match the intended or desired effects” (De Graaff & Housen, 2009:727-728). However, as mentioned in Section 2.2.1, it is important to keep in mind the following observation by Cheetham *et al.* (1992:9-10):

Despite much apparently straightforward use of the word, ‘effectiveness’ is not something which has an object-like reality ‘out there’ waiting to be observed and measured. Like any other data, empirical evidence about the effectiveness of [...] programmes is a product of data collection procedures and the assumptions on which they are based. The concept of effectiveness derives from particular ways of thinking and makes sense only in relation to its context. [...] The challenge is to arrive at working definitions of effectiveness in specific situations, and hence of methods of studying it, which do not permanently lose sight of its conceptual context.

Keeping the above argument in mind, and as discussed in Section 2.2.1, impact (or effect) will, for the purposes of the current study, be seen as i) the observable improvement in academic literacy abilities between the onset and the completion of an academic literacy intervention, and ii) the extent to which these abilities are necessary and applied in students’ content subjects. This definition is drawn upon in the preliminary evaluation design proposed in Chapter 4.

3.3 Previous studies on the impact of academic literacy interventions

Before considering in detail the studies that have reported on impact, it is worthwhile to distinguish between “language programmes” and “academic literacy interventions”. As the term implies, the focus of language programmes is on students’ language, and very often these programmes focus on the language abilities of second language users. Academic literacy programmes, however, include (but are not limited to) language ability (Van Dyk & Van de Poel, 2013:53). This study accepts Van Dyk and Van de Poel’s (2013:56) definition of “academic literacy” as “being able to use, manipulate, and control

language and cognitive abilities for specific purposes and in specific contexts”. Due to the dearth of studies measuring impact in either language programmes or academic literacy interventions, this chapter considers studies from both of these fields.

Studies measuring the impact of academic literacy and language courses are indeed few and far between (Mhlongo, 2014:47; Terraschke & Wahid, 2011:174; Carstens & Fletcher, 2009b:319; Storch & Tapper, 2009:208; Holder *et al.*, 1999:20). Yet, argues Butler (2013:80), in addition to having a theoretical justification for the type of intervention that is developed, the intervention’s success is ultimately determined by the impact it has on students’ learning. Some South African as well as international studies have been able to effectively measure certain aspects of such an impact. As is seen in the survey below, most of these studies focus on only one or two aspects of impact. However, as Beretta (1992:19) states, no single methodology can provide a full picture when it comes to the evaluation of language programmes (for example, academic literacy programmes). A comprehensive, validated and verified evaluation design might assist researchers in choosing a more comprehensive range of tools in order to determine the impact of academic literacy courses. An overview of studies that have attempted to measure the impact of academic literacy courses, in one form or another, follows below.

Parkinson *et al.* (2008) evaluated a course called “Communication in Science” taken by students in a science access programme at the University of KwaZulu-Natal. In their evaluation, a pre-test/post-test design was employed, using a placement test consisting of multiple-choice questions, cloze questions and writing elements. The test aimed to measure students’ ability to read for meaning, extrapolate and apply information, infer information, separate essential and non-essential information in a reading text, and extract and interpret information from texts to use in an extended writing task. This test was also taken by mainstream students at the beginning of the year, though there was no post-test for these students. Secondly, a questionnaire was given to students at the end of the year to determine their opinions of the course. Perceived improvement was assessed by asking students whether they learned “a lot”, “a little” or “nothing” with regard to several outcomes. Thirdly, students who had previously completed the course were given a questionnaire to determine whether they believed that the competencies acquired in the academic literacy course were of value in their subsequent studies. These students were asked the following question via e-mail: “Since completing the Communication course,

in what ways have the skills you have learnt in Communication in Science been useful to you?" (Parkinson *et al.*, 2008:23).

The results showed that access students improved significantly over the duration of the course, and in some cases even got scores close to those of the mainstream students at the beginning of the year. However, since no post-test was written by the mainstream students, it was impossible to determine whether they had also improved equally despite not having undergone the intervention. As far as students' perceptions of the course are concerned, findings showed that students generally enjoyed the course, and believed that they had learned a lot in the various sections. Thirdly, more senior students who had completed the course previously mostly responded that the course had been beneficial to them. Even though a control group was not available in the Parkinson *et al.* (2008) study, the research design was strengthened in that more validity was given to findings through triangulation. However, certain aspects of the evaluation design could have been addressed more comprehensively. For example, students' perceptions might have been ascertained more effectively. The question that was asked could be considered to be leading as it did not allow students to state which abilities were not useful; more detailed and extensive questions could have been asked to determine which abilities addressed in the academic literacy course were used in further studies, and to which extent they were used. The placement test could also have been analysed to determine in which areas students had improved the most over the duration of the academic literacy course, and these findings could have been correlated with students' perceptions about how much they had learned in the course. Thus, triangulation could have been strengthened in various ways.

Van Dyk *et al.* (2009) took various steps to determine whether an academic literacy intervention in the Health Sciences at Stellenbosch University had an impact on students' writing abilities. Students firstly completed pre-, mid- and post-intervention writing assignments. Results were analysed quantitatively by correlating them with each other. Assignments were also examined by lecturers who noted the difference in execution between the pre- and post-assignments and listed typical errors for both of these assignments. Writing was considered at both the micro level (including language and word choice) and macro level (including paragraph structure, cohesion and coherence, and argumentation). Students' writing at both of these levels seemed to have improved

between the pre- and post-intervention writing assignments. This qualitative feedback was the most useful feedback in this particular study; however, a weakness was that the evidence remained mainly anecdotal, consisting of lecturers' impressions. Finally, students completed feedback questionnaires on, amongst others, the relevance of material and the learning outcomes. This feedback consisted of many more positive than negative qualitative comments. A limitation might have been that the questionnaire took the form of an official student feedback form. This means that the questionnaire was designed mainly to measure students' perceptions of the course itself and the way that it was presented. Such official feedback forms are often the only tools available to lecturers to gauge perceptions on specific courses. However, they rarely allow lecturers to assess which aspects of the course students found most useful, and to what extent the course was likely to impact on their general academic success. The purpose of this study was to determine the impact that the course had on students' writing abilities, and thus a writing assignment was suitable. The different forms that the three writing assignments took, however, made it difficult to draw direct comparisons, as would be the case in a pre- and post-assignment scenario where two or more equivalent assignments with the same outcomes are used.

Van Dyk *et al.* (2011a) reported on a study conducted at Stellenbosch University that focused on the effect of an academic literacy course in the field of natural sciences on students' reading levels. The study consisted of quantitative data in the form of a pre- and post-test (aimed at assessing students' reading abilities), an online questionnaire that aimed to determine which reading abilities students believed to be important in order to be successful in their studies, as well as official student feedback forms. The Test of Academic Literacy Levels (TALL) and its Afrikaans equivalent, *Die Toets van Akademiese Geletterdheidsvlakke* (TAG), were used as pre- and post-tests. The test construct of these tests measures the following: understanding academic vocabulary; interpreting metaphor, connotation and ambiguity; understanding relations between parts of a text; interpreting and showing sensitivity to various text types; interpreting, using and producing visual information; making distinctions between various types of information; seeing sequence and order; understanding evidence used in texts; understanding the communicative functions of ways of expression in academic language; and making meaning beyond sentence level (see Weideman, 2003a:xi for a more detailed

explanation). Qualitative data consisted of open-ended questions in the official student feedback forms.

Results showed that the impact of the academic literacy course becomes clearer after a year's intervention than after a semester's intervention, indicating that long-term interventions might be more beneficial to student success than short-term interventions. Feedback from student questionnaires indicated that students believed that reading abilities were important for a student to be successful in his/her studies, that the module achieved its outcomes, and that necessary academic literacy abilities were developed. In this study, the use of a valid and reliable academic literacy test enabled conclusions based on statistical analysis that were not possible in the Van Dyk *et al.* (2009) study. However, similar limitations as in the previous study exist. For example, official feedback forms are possibly not the most effective way of assessing the impact of a course. Furthermore, the study is limited to assessing reading levels, whilst other academic literacy abilities might also have improved over the duration of the course; thus, using a wider variety of assessment instruments might have been useful.

Mhlongo (2014) assessed the impact of an academic literacy intervention at the Vaal Triangle Campus of the North-West University. He made use of the same academic literacy test that was used in the Van Dyk *et al.* (2011a) study – i.e. the TALL – but also drew on the perceptions of students as well as mainstream lecturers who taught first-year students by administering questionnaires developed for this purpose. He further drew a correlation between students' overall academic achievement and their academic literacy levels. A particularly useful aspect of this study was the use of a control group. All students who obtained below 50% for the TALL were required to participate in the academic literacy course, whereas students who obtained 50% and above were exempted – it would thus seem as though the formation of a control group might have been difficult. Mhlongo (2014), however, used two groups of students: those who obtained between 40% and 49% (and who thus participated in the intervention – the experimental group) and those who obtained between 50% and 59% (thus those who were exempted from the academic literacy course – the control group). By using two groups of students who obtained similar marks as experimental and control groups, certain statistical conclusions could be made about the impact of the academic literacy course on student success.

Mhlongo's (2014) study indicated that there was a statistically significant improvement in the experimental group's mean scores between the pre- and post-tests. Furthermore, his results indicated that there was no such improvement in the control group students' scores. Student feedback was generally positive, although some students indicated that the courses were not relevant to their studies. Some students also indicated that more time needed to be allocated to the modules. Feedback from content-subject lecturers indicated that these lecturers were largely unaware of the abilities addressed in the academic literacy course. Furthermore, they did not seem to think that the academic literacy course made a substantial difference to their students' academic literacy levels. In addition, lecturers felt that generic academic literacy courses were not ideal, as they believed their own disciplines to be very different from other disciplines. They also did not believe that it was their responsibility to help students acquire academic literacy abilities.

Carstens and Fletcher (2009b) evaluated a subject-specific essay-writing intervention for history students at the University of Pretoria. The intervention was assessed by means of a pre- and post-test (in the form of an essay) as well as student responses regarding their perceptions of the course. A seven-point scoring rubric was used for the pre- and post-test, with percentages to give the assessor a clear idea of a benchmark for each mark allocation. The scoring instrument is based on three analytical rating scales that are internationally accredited. The following four dimensions are addressed by the scoring instrument: use of source material, structure and development, academic writing style, and editing. An N/A option was given for items which are not relevant in all types of writing (for example, referencing, legibility or layout). According to Carstens and Fletcher (2009b:324), "the success of academic literacy interventions are equally dependent on students' experience, which are co-determinants of motivation and skills transfer". Therefore, a survey was conducted to determine the opinions of the participants. The questionnaire uses a standard five-point Likert scale. This type of questionnaire would seem more useful in comprehensively determining perceptions than the purely open-ended questions that were used in some of the studies discussed in this review.

Results indicated that students improved in three dimensions between their pre- and post-test essays. These dimensions included their use of source material, structure and

development, and academic writing style. Students' editing abilities did not seem to have improved over the course of the intervention. The opinion survey showed that students were generally positive about the effect of the intervention on their writing abilities. They were also in favour of the genre-specific approach that was followed in this intervention. Furthermore, the results indicated that more attention should be paid to formality and precision in academic writing, as well as developing self-confidence to challenge authority. One limitation of this study is that only ten students completed the course, making it difficult to reach statistically significant conclusions based on this small number.

Storch and Tapper (2009) assessed the impact of an English for Academic Purposes (EAP) course presented at the University of Melbourne that was aimed at developing the academic literacy abilities that are required for successful study at postgraduate level. Student writing was assessed by means of a pre- and post-test writing task. What sets this study apart from similar studies is the type of quantitative research design used in its assessment of student writing. The study measured students' fluency by looking at words per T-unit, their accuracy by counting errors in various categories, their use of academic vocabulary by comparing student lexis to Coxhead's (2000) academic wordlist, and their text structure and rhetorical quality by using a guide developed by the authors themselves. In addition to this statistical analysis, questionnaires were distributed to gather information about students' English language use and proficiency, as well as their perceptions regarding the usefulness of the course (one open-ended question was used to determine the latter).

Quantitative results from this study showed that there was no measurable effect on student fluency; however, statistically significant improvements were observed in students' grammatical accuracy and their use of academic vocabulary. Improvements were also observed in students' text structure and rhetorical quality. Qualitative outcomes indicated that the course had made students more aware of various academic writing strategies. While assessing student writing quantitatively in this manner is certainly an interesting approach that merits consideration, especially when the aim is to determine a course's strengths and weaknesses, this study might have benefitted by comparing these results to those obtained from a more traditional writing rubric. Perceptions might also

have been measured more effectively by asking more specific questions regarding the usefulness of the course.

Some studies evaluate impact by comparing the results of two or more courses with each other¹³. Harker and Koutsantoni (2005) compared the effectiveness of distance versus blended learning in a web-based EAP programme at the University of Luton. Students completed a diagnostic test that comprised a summary as well as a short essay as both pre- and post-tests. In addition, students completed formative feedback forms on what they found to be the most and least useful components of each lesson, as well as summative feedback forms on the course. Feedback forms contained both closed-ended as well as open-ended questions.

Both groups of students performed better in the short essay part of the post-test than they had in the pre-test. Blended learning students who attended more classes performed better than those who attended fewer classes. There was no significant improvement in the summary section between the pre- and the post-test. Both groups of students gave more positive than negative formative feedback. The summative feedback indicated that, when compared with the feedback from the blended learning group, the distance learning group agreed that the course addressed their needs to a greater extent, though the majority of both groups felt that they had learned valuable academic English skills. A possible weakness in this study is that the summative feedback form did not revisit the various abilities addressed during the course. There was thus no indication of how useful students considered the various abilities after having completed the entire course and having had time to reflect on these abilities. Furthermore, data from various sources were not triangulated to ultimately obtain stronger research results.

Carstens (2011) used a quasi-experimental design to compare the pre- and post-test essay ratings of students in a generic academic literacy writing course with those of students in a discipline-specific writing course. The same scoring rubric that was used in Carstens and Fletcher (2009b) was used in this study. Carstens also used surveys to determine students' opinions of the course by looking at five dimensions: staged and scaffolded teaching and the learning model, purposeful social apprenticeship, a needs-driven

¹³ Refer back to the discussion of Mhlongo (2014) for an alternative control group experiment.

syllabus, critical orientation, and skills transfer. Although both groups of students performed significantly better in the post-test than in the pre-test, students from the discipline-specific writing course outperformed those from the generic writing course. Furthermore, although both groups gave positive feedback about their respective courses, the discipline-specific group's feedback was significantly more positive than that of the generic group. According to Carstens and Fletcher (2009b), limitations of this type of quasi-experimental design include that the comparison might be jeopardised due to differences between the syllabi and presentation of the interventions, as well as differences between the two groups. Furthermore, the fact that the courses were presented consecutively rather than simultaneously might be problematic as designers as well as presenters might have learned from the first intervention, and thus applied corrective measures to the second intervention.

A selection of other studies that compare the results of two or more courses are briefly summarised here. Kasper (1997), at the Kingsborough Community College, compared the language course results of English second language students receiving content-based instruction to those who were enrolled in generic language programmes. Murie and Thomson (2001) considered the impact of an academic literacy course by comparing the retention rates of the students who participated in the course to those of a control group. Song (2006) compared the impact of content-based EAP courses with that of generic EAP courses at the City University of New York. In this study, the following aspects of students' receptive abilities were assessed: comprehension of the text; ability to identify main ideas, purpose and tone; and ability to analyse information and to draw inferences. As for productive abilities, students were expected to submit a portfolio containing various examples of essays in different genres completed during the semester. Furthermore, they completed an in-class essay assessment.

While there are certainly clear benefits in using an experimental approach with a test and a control group, this type of study is not possible at many universities where no equivalent control groups exist. Furthermore, universities are often hesitant to allow the use of true experimental designs due to ethical considerations. Thus, while using control groups might be preferable, an evaluation design – specifically in the South African context – would have to be comprehensive enough to still allow valid conclusions to be drawn by means of triangulation, despite the lack of appropriate control groups.

At the University of Cape Town, Archer (2008) attempted to assess the impact of a writing centre on students' writing. She used a multi-faceted approach in which she collected data by i) ascertaining students' perceptions with regard to writing centre work, ii) collecting writing centre consultants' comments, iii) considering students' grades, and iv) comparing independently assessed first and final student drafts (marked by looking at organisation, language use, as well as voice and register). Archer (2008:251) reminds us that students' "perception of improvement may not necessarily translate into demonstrably improved writing". It is therefore also necessary to empirically assess such an improvement. Archer triangulated data by looking at students' perceptions, their writing, the grades they obtained for their writing, and consultants' perceptions of the writing.

Students indicated that the writing centre intervention assisted them in focusing on the task, improving their voice and register, and improving macro- as well as micro-structural issues. Furthermore, students seemed to have a greater awareness of their own writing after attending writing consultations, and were more able to articulate their writing processes. All students passed the assignments on which they had consulted. Finally, between first and final drafts, students improved in all three areas, but most pronouncedly in voice and register as well as organisation. It should be noted that writing centres generally do not consider the full range of academic literacy abilities – their focus on writing is reflected in the methodology employed in this study. Possible weaknesses of this study include that the consultants were students and not necessarily qualified language experts (though they have undergone thorough training); moreover, variables such as students' educational background or the influence of other academic literacy courses were not controlled for, perhaps because this is particularly difficult in a writing-centre context.

Several studies that assess interventions look at these interventions from limited perspectives. Some studies focus mainly on quantitative data. Van Der Slik and Weideman (2008), for example, examine whether there is an improvement in the TALL scores of high risk students at the University of Pretoria after having completed a year-long academic literacy intervention. They found that all three groups of students examined, namely "extremely high risk", "high risk" and "borderline" students, showed significant improvements in their TALL scores by the conclusion of the intervention. The

group that showed the greatest improvement was the “extremely high risk” group, although this improvement was not enough to classify these students as “not at-risk” by the end of the year. The students who seemed to have gained the most of the intervention, in that they could be considered “not at-risk” by the end of the intervention, were those who were initially identified as “high risk” and “borderline”. This indicates that students who are identified as at high risk of failing their university studies due to insufficient academic literacy proficiency should receive either more intensive, or longer, academic literacy interventions. Van Wyk and Greyling (2008) assessed the impact of using graded readers for low-proficiency students at the University of the Free State. Students’ academic literacy levels were assessed by means of the Placement Test in English for Educational Purposes (PTEEP) as a pre- and a post-test. These data were not triangulated with other data sources. Carstens and Fletcher (2009a) statistically analysed the improvement in students’ writing abilities by looking at the pre- and post-test results of a cross-disciplinary essay writing intervention aimed at second-year students in the Humanities. Fouché (2009) described a writing centre intervention: a series of academic literacy workshops aimed at first-year students in UNISA’s Science Foundation Programme. In this intervention, pre- and post-test results of an academic literacy test were compared and correlated with student attendance. The problem with this type of correlation study is that it is very difficult to control variables like student motivation; more motivated students who attend more workshops (in the context of this study, or classes in other contexts) might have outperformed less motivated students who attended fewer workshops, regardless of the number of sessions attended.

In contrast, some studies rely mainly on qualitative measures to determine course impact. Thompson (2011) evaluated an “English for Tourism” intervention, aimed at fourth-year students at a Thai university. The course was assessed using a student questionnaire to determine students’ reactions to various course features, interviews with a variety of stakeholders to determine their perceptions of the programme, a teacher’s log to document and reflect on various aspects of the course, and learning materials which were analysed. Similarly, Ngoepe (2007) evaluated an academic literacy course called “English and Study Skills” at the University of Limpopo. The course was evaluated by means of lecturer interviews, student questionnaires, an analysis of materials, and a survey of similar courses. Kiely (2009) evaluated English for Academic Purposes materials at a British university by means of an ethnographic study, which included

interviews with students and teachers, an end-of-course questionnaire, field notes, and an analysis of learning materials. Butler and Van Dyk (2004) broadly looked at students' perceptions of an Engineering course at the University of Pretoria. They also mentioned anecdotal evidence from lecturers. Similarly, Bharuthram and Mckenna (2006) reported on students' perceptions (obtained by means of an evaluation questionnaire) of the success of a writer-respondent intervention at the Durban Institute of Technology. An important limitation of these studies is that no instruments were used to determine whether there was an improvement in students' academic literacy (or language) abilities between the onset and the conclusion of the various interventions. Also, in most cases, questionnaire and interview questions mainly focused on the course in general, and did not sufficiently consider various abilities addressed throughout the course.

Winberg *et al.* (2013) also took a qualitative approach to evaluating the effectiveness of four discipline-specific academic literacy case studies, each of which was based on a collaborative effort between academic literacy and content-specific specialists. In the first case study, fourth-year undergraduate Science and Technology student teams were responsible for developing product prototypes. In this case study, debriefing meetings were held in which academic literacy and content-specialists reflected on what had been learned from the collaborative effort. Subject specialists had not provided any formative feedback in this case study, a factor which the authors identify as problematic. The second case study involved a collaboration between academic literacy and subject-content specialists to develop multilingual glossaries. In this case study, participants reflected on the effectiveness of these multilingual glossaries through observations at various stages during the collaboration. Furthermore, reflective semi-structured interviews were held with the subject-content specialists. These interviews were analysed qualitatively, looking for emerging themes. In the third case study, academic literacy and subject-content specialists collaborated in co-authoring a textbook aimed at giving first-year students "linguistic access to content knowledge in an SET [Science, Engineering and Technology] discipline" (Winberg *et al.*, 2013:95). This collaboration was evaluated by conducting structured interviews with the co-authors, which were again qualitatively analysed. During this case study, regular meetings were also held between academic literacy and subject-content specialists to provide participants with a "transactional space" (Winberg *et al.*, 2013:96). The fourth case study reported on "aimed to provide linguistic access to disciplinary knowledge through interdisciplinary collaboration

involving pairs [of academic literacy and subject-content specialists in Science and Technology disciplines]”. In this case study, one academic literacy specialist was partnered with a subject-content specialist – in total, twenty lecturers participated. The collaboration “entailed dovetailing curricula, developing shared classroom materials, team teaching, and designing and co-assessing tasks” (Winberg *et al.*, 2013:97). Feedback on the success of the collaborations consisted of narrative interviews, focus group sessions, and reflective writing – these were all qualitatively analysed.

The Winberg *et al.* (2013) study highlights the importance of obtaining feedback from primary stakeholders – in this case academic literacy and content-subject specialists – when determining whether interventions could be considered effective. However, the strong focus on the working relationships between academic literacy and content-specific specialists at the expense of additional data leaves one wondering whether the students actually improved as a result of these interventions. These case studies might have been strengthened by, for example, considering feedback from students involved in the interventions as well as analysing quantitative data so as to consider more comprehensively the success of these interventions.

Another way in which impact has been measured is by investigating how language ability measures correlate to general academic success. A recent study by Van Rooy and Coetzee-Van Rooy (2015), conducted at the Vaal Triangle Campus of the North-West University, focused on the 2010 intake of first-year students and found that the Grade 12 results of students who achieved an average of below 65% for all subjects could not, with confidence, predict academic success at university; the Grade 12 results of students who achieved an average of 65% and higher, however, could be used as a predictor for academic success. The study further found that academic literacy tests are not good predictors of success at university level. However, this study found that students’ marks in academic literacy modules were good predictors of academic success. Mhlongo (2014) similarly found a significant correlation between students’ academic literacy course marks and their marks in other subjects for the 2012 intake at the same university. One question that should be raised with this type of correlation is whether the positive correlation between academic literacy course marks and content-subject marks is because higher academic literacy levels (acquired in the academic literacy course) resulted in higher marks in content subjects, or whether stronger students naturally performed better

in both measurements, and weaker students poorer in both. Thus, on its own, this measurement would not seem to be useful in assessing the impact of an academic literacy intervention. However, as part of a triangulated study (as was done in the study by Mhlongo), such a measurement could provide valuable insight into the impact of such interventions.

3.4 Discussion and conclusion

To summarise, various approaches to assessing the impact of academic literacy interventions can be identified in the literature. Two main aspects of impact stand out, namely determining whether students' academic literacy levels had improved over the duration of the intervention, and establishing whether students transferred these abilities to their other subjects.

Two main approaches have been used to assess whether there was an improvement in students' academic literacy levels between the onset and conclusion of an intervention. The first is assessing whether there is an improvement in students' writing abilities¹⁴ by using a rubric (e.g. Carstens & Fletcher, 2009b; Storch & Tapper, 2009; Van Dyk *et al.*, 2009; Archer, 2008; Parkinson *et al.*, 2008; Song, 2006) or statistically examining features of student writing (Storch & Tapper, 2009). The second is by assessing whether there is an improvement in students' academic reading abilities, often by means of a verified and validated academic literacy test (e.g. Mhlongo, 2014; Van Dyk *et al.*, 2011a; Fouché, 2009; Parkinson *et al.*, 2008; Song, 2006).

In addition to assessing whether there was an improvement in students' academic literacy abilities, it is also important to determine whether these improved abilities were effectively used in students' other subjects. A seemingly effective way of determining whether an improvement in test scores can be attributed to a specific intervention is to use appropriate control groups (consider, for example, Mhlongo, 2014; Carstens, 2011; Song, 2006; Harker & Koutsantoni, 2005; Murie & Thomson, 2001; Kasper, 1997). An

¹⁴ Although writing and reading abilities are referred to in this study, for the sake of convenience, they should be seen as broad categories that overlap, both addressing a variety of academic literacy principles. These include being able to interpret information, collaborating with the author or audience, using conventions, being aware of cultural knowledge, solving problems, and reflecting and using language appropriately (cf. Kern 2000:16-17).

additional and sometimes alternative method of determining whether the abilities acquired in a course were transferred to other subjects is by determining students' perceptions regarding the impact of a course (e.g. Mhlongo, 2014; Van Dyk *et al.*, 2011a; Carstens & Fletcher, 2009b; Kiely, 2009; Storch & Tapper, 2009; Van Dyk *et al.*, 2009; Archer, 2008; Parkinson *et al.*, 2008; Bharuthram & Mckenna, 2006; Butler & Van Dyk, 2004). One danger, however, is that it is very difficult to determine the reliability of perceptual data – just because students say that they have acquired (and transferred) abilities does not mean that this is necessarily the case. Further methods of determining whether improvement in academic literacy levels can be attributed to an intervention include interviewing stakeholders (other than students) to determine their perceptions of the intervention (e.g. Mhlongo, 2014; Winberg *et al.*, 2013; Thompson, 2011; Ngoepe, 2007), correlating student performance with class attendance (e.g. Fouché, 2009), and correlating students' performance in the academic literacy course with their performance in their content subjects (e.g. Van Rooy & Coetzee-Van Rooy, 2015; Mhlongo, 2014).

An important facet of the responsible implementation of academic literacy interventions is to assess whether these interventions have a significant impact. Merely offering academic literacy interventions to bow to national and international pressure for the establishment of such interventions is not enough. Universities, departments and units that offer academic literacy interventions are responsible for ensuring that these interventions have the highest impact possible. The studies discussed in this chapter have all attempted to do this to some extent, which indicates that some researchers are aware of the importance of assessing the impact of academic literacy interventions. However, the variety (and sometimes inconsistency) of approaches used raises the question of what the most appropriate way would be to assess the impact of academic literacy interventions.

Unfortunately, “[r]esearch on impact in education is difficult, partly because there are typically many factors involved which are difficult to control, so that the impact of any one element in the system is hard to distinguish” (Howes, 2003:148). This is certainly the case when trying to assess the impact of an academic literacy intervention, since there are many factors at play, including general exposure to academic literacy abilities in students' other subjects and possible feedback on academic literacy related issues from content-subject lecturers. Furthermore, forming a control group is not possible at many

universities, as almost all students have some kind of academic literacy intervention as part of the credit-bearing programme offering. Thus, alternative research designs must be considered if the impact of academic literacy interventions is to be meaningfully assessed.

Jick (1979) argues that a more certain portrayal of a phenomenon is provided when multiple and independent measures reach similar conclusions. Lynch (1996:60) agrees and states that “triangulation seems like an obvious strategy for strengthening the validity of evaluation findings”. He adds that the possibility of bias always exists in any particular technique or source; however, using a variety of sources of evidence could potentially cancel “the bias inherent in any one source or method” (Lynch, 1996:60). Therefore, by examining a variety of factors that may shed light on the agency of an academic literacy course in the ultimate improvement of students’ academic literacy abilities, and the extent to which these abilities were transferred to other subjects, a more valid inference can be made regarding the causal relationship between such improvement and the academic literacy intervention.

The main shortcoming identified in the studies reviewed in this chapter is that researchers generally do not focus on enough facets of impact to truly be able to make conclusions about the impact on their respective interventions. Another shortcoming is that, in many cases, the instruments that are used to determine impact have not necessarily been verified or validated. Both of these shortcomings might be due to the fact that so few guidelines exist in the literature regarding the effective measurement of impact in the field of academic literacy.

Based on the literature discussed in the previous section as well as the definition of impact that was put forward in this chapter, this study proposes that in order to determine the impact of an academic literacy intervention, two broad aspects of impact on student success must be examined, namely the improvement (if any) in students’ academic literacy levels, and the extent to which these abilities are used in and transferred to students’ content subjects. However, as Mhlongo (2014:4) points out, “each tertiary institution faces unique challenges with regard to the specific needs of its students, which makes it essential that specific academic literacy interventions [...] be assessed within the context of addressing such needs”. Since academic literacy courses vary vastly in

terms of, for example, content and purpose, any evaluation design for assessing their impact would have to be flexible. It is likely that such a design would have to include certain generic components that would address integral aspects that should be part of each academic literacy intervention¹⁵. However, the researcher would have to be able to adapt some research tools so as to most effectively assess the impact of each individual academic literacy intervention, as not all academic literacy interventions have the same foci or objectives.

The aim of this chapter was to provide an overview and critique of studies conducted thus far that have attempted to assess the effectiveness of various academic literacy interventions. Chapter 4 proposes a conceptual evaluation design that could be used for various types of academic literacy interventions, based on the literature that was reviewed in the current chapter, and the guidelines provided in this section. This design is later verified and validated by i) using it to assess the impact of an academic literacy intervention in the South African context (see Chapters 5 and 6), and ii) asking academic literacy intervention coordinators across South Africa about the extent to which the proposed evaluation design meets their needs, and how it could be refined to be applicable to their specific contexts (see Chapter 7).

¹⁵ Consider, for example, Van Dyk and Van de Poel's (2013:56) definition of "academic literacy" as "being able to use, manipulate, and control language and cognitive abilities for specific purposes and in specific contexts". Based on this definition, it would be vital that students' abilities to "use, manipulate, and control language and cognitive abilities" be assessed using methods that can be triangulated.

CHAPTER 4

IMPACT MEASUREMENT: TOWARDS CREATING A FLEXIBLE EVALUATION DESIGN FOR ACADEMIC LITERACY INTERVENTIONS¹⁶ (ARTICLE 2)

4.1 Prelude to Chapter 4

In Chapter 3, a broad overview was given of studies that have attempted to assess the impact of academic literacy (or related) interventions. Furthermore, the strengths and weaknesses in the methodologies of these studies were examined. The most common data gathering instruments that were used in these studies were also briefly considered.

The current chapter considers a range of instruments based on the studies discussed in Chapter 3. It provides guidelines for each of these instruments, discusses their limitations, and also suggests some specific instruments which could be used in the field of academic literacy. Finally, based on the information provided in Chapters 2, 3, and 4, a preliminary evaluation design for assessing the impact of academic literacy interventions is proposed.

4.2 Introduction

Due to a variety of reasons, the foremost of which is possibly inadequate secondary education, the implementation of academic literacy interventions in South African universities has become commonplace (Cliff, 2014:322; Sebolai, 2014:52; Davies, 2010:xi). What is still largely lacking, though, is evidence of the effectiveness of such interventions, and the extent of the impact they have (Sebolai, 2014:52; Butler, 2013:80; Terraschke & Wahid, 2011:174; Carstens & Fletcher, 2009a:319; Storch & Tapper, 2009:218; Holder *et al.*, 1999:20). For the purposes of the current study, impact (or effect) is defined as i) the observable improvement in academic literacy abilities between the onset and the completion of an academic literacy intervention and ii) the extent to which these abilities are necessary and applied in students' content subjects (Section

¹⁶ Published as: Fouché, I., van Dyk, T. & Butler, G. 2015. Impact measurement: Towards creating a flexible evaluation design for academic literacy interventions. *Stellenbosch Papers in Linguistics*, 45: 109-145. Minor editorial changes were made for the sake of the format of this dissertation.

2.2.1; also see Fouché, 2015:3). The terms “impact” and “effect” are used synonymously for the purposes of the current study.

According to Hatry and Newcomer (2010:678), two reasons for evaluating interventions are firstly to provide accountability to stakeholders in the intervention, and secondly to improve the effectiveness of such programmes. As discussed in Chapter 3, though there have been some attempts at assessing the impact of academic literacy interventions, the type and variety of research instruments have rarely been sufficient to validly and reliably determine firstly, whether these interventions have an impact, and secondly, what the degree of this impact might be.

There are of course various challenges to effectively assessing the impact of an academic literacy intervention. A prominent challenge is the wide variety of academic literacy interventions offered not only in South Africa, but also elsewhere. These include, but are not limited to, generic interventions (e.g. Weideman, 2003b; Van Wyk, 2002), subject-specific interventions (e.g. Carstens, 2013b; Ngwenya, 2010; Goodier & Parkinson, 2005), collaborative interventions (for instance approaches where subject-lecturers teach academic literacy abilities in the content-subject class or team-teaching approaches, e.g. Dudley-Evans, 1995)¹⁷, writing centres (e.g. Archer, 2008), and reading interventions (e.g. Van Wyk & Greyling, 2008). Each of these interventions has its own set of advantages and disadvantages – see Van de Poel and Van Dyk (2015) for a discussion of the various types of academic literacy interventions. A one-size-fits-all approach to impact evaluation seems unlikely to effectively address such a variety of interventions, each with its own purpose and set of outcomes.

A further obstacle to effectively assessing the effectiveness of academic literacy interventions in the South African context is that in many (if not most) instances, traditional experimental designs are not feasible. It is certainly preferable to have a control group against which to measure an experimental group, so as to eliminate the effect of the many variables (for example, the effect of studying other university subjects – see Archer [2008: 249]) that may also impact on students’ academic literacy levels. Yet, increasingly, South African higher education students are expected to participate in

¹⁷ See Van de Poel and Van Dyk (2015: 169-173) for a more in-depth discussion of collaborative approaches.

some type of academic literacy intervention in their first year, and in many cases, such interventions (often an academic literacy credit-bearing course) are compulsory for all students of a faculty or university. Although the triangulation of data (discussed in more detail in Section 4.4) is generally seen as a desirable method of determining the validity of findings (Lynch, 2003:152; 1996:59-61), in cases where control groups are not possible, triangulation becomes vital in reliably and validly assessing the impact of academic literacy interventions.

4.3 Aim of the chapter

This chapter aims at developing an evaluation design that could be used to assess the impact of a wide variety of academic literacy interventions. Chapter 3 provided an overview of studies that have aimed to evaluate the impact of a range of academic literacy interventions. The current chapter builds on the literature surveyed in Chapter 3, and proposes a conceptual evaluation design which could be used to assess the effectiveness of various academic literacy interventions. Chapters 5 and 6 will empirically verify this conceptual evaluation design, and the results from these chapters will further inform the final evaluation design.

Whereas Chapter 3 focused on studies that have been conducted in the field of impact assessment, Section 4.4 focuses on the instruments that have and can be used, in addition to guidelines that are given in relation to the use of these instruments, and what their respective strengths and limitations are. In Section 4.5, an evaluation design is proposed through which a combination of these instruments could be used to assess various academic literacy interventions.

4.4 Instruments for assessing the impact of academic literacy interventions

This study argues that course impact can only truly be assessed if a range of methods is used in an integrated manner. Based on the definition of impact proposed earlier in this chapter, it follows that when choosing instruments for the purposes of evaluation, it is important to keep in mind which instruments could be used to assess the observable improvement in students' academic literacy levels, and which would be appropriate when

determining the extent to which these abilities are firstly necessary, and secondly applied, in students' content subjects.

The evaluation design that will be proposed in the current chapter is based on strategies that have been followed in previous studies (discussed in Chapter 3). However, it will attempt to overcome the weaknesses that have been presented in those studies, so as to ultimately propose a protocol of valid and reliable research tools. It will further rely on the approach of triangulation to validate (and cross validate, cf. Jick [1979]) the various findings regarding the impact of academic literacy interventions.

Triangulation can be defined as “the combination of methodologies in the study of the same phenomena” (Denzin, 1989:234). Data can be triangulated in two main ways: by methodology (for example using various qualitative and quantitative tools) as well as by data sources (for example, obtaining data from students, lecturers and the literature) (Lynch, 1996:59; Denzin, 1989:237). According to Jick (1979:602), triangulation is a “vehicle for cross validation when two or more distinct methods are found to be congruent and yield comparable data” (cf. Lynch, 1996:15).

As to the relationship between evaluation and triangulation (cf. Section 2.5.3), Alderson (1992:285) argues:

The notion of triangulation is particularly important in evaluation. Given that there is No One Best Method for evaluation, it makes good sense to gather data from a variety of sources and with a variety of methods, so that the evaluator can confirm findings across methods. But in evaluation, outcomes and effects are often contentious, and the perception of these may vary considerably depending upon the perspective, vested interest or even personality of the participants. In addition, the views of stakeholders as to the appropriacy of particular methods for evaluation typically also differ radically. In such settings, it does not seem sensible to rely on one method for the purposes of data gathering, but rather to try to complement or neutralise one method with another. Thus, a generalisable recommendation for evaluations would be that they should wherever possible plan to triangulate in method.

According to Lynch (1996:60), using several sources of evidence could help the evaluator to counteract whichever bias might be inherent in any single source or method (see also Judd & Keith, 2012:40). Therefore, more valid inferences with regards to the causal relationship between an academic literacy intervention and an improvement in

students' academic literacy abilities can be made by considering various sources of evidence that may shed light on the agency of a specific academic literacy intervention.

4.4.1 Assessing improvement in students' academic literacy levels

In determining whether an academic literacy intervention might have an impact (as defined in the introduction of this chapter) on student success, the first necessity would be to determine whether there was indeed an improvement in students' academic literacy levels between the onset and the completion of the intervention. Furthermore, one would have to ascertain which aspects of academic literacy showed a marked improvement, and which seemed to be unchanged after the completion of the intervention. This section thus addresses the first aspect of the definition of impact as given in Section 4.2.

According to Henning (1987:2), a common use of tests is to determine how effective an intervention was. Tests also provide feedback (or a backwash effect) about which course outcomes would seem to have been successfully taught, and which outcomes might need to be approached differently (see, for example, Hughes, 2003:1-2) as the course is refined. This is a process which should surely be continuous in all courses, especially those in higher education, in which knowledge can never be considered a constant, and is always in flux. Furthermore, by analysing tests statistically (for example, by means of t-tests, regression analyses, correlations, effect sizes and analyses of variance), various external variables can be accounted for.

Assessments must be reliable, adequate and appropriate. Reliability refers to “the actual level of agreement between the results of one test with itself or with another test” – a high level of agreement would indicate that there were not measurement errors in the assessment (Davies *et al.*, 1999:168). The focus of reliability is on the empirical aspects of measurement; validity, in contrast, “focuses on the theoretical aspects and seeks to interweave these concepts with the empirical ones” (Davies *et al.*, 1999:169). An assessment can be considered valid if it measures the concept(s) that it intended to measure (Davies *et al.*, 1999:221). Another important aspect of validity, namely multiple sources of evidence, is discussed in more detail in Section 4.5. The adequacy and appropriateness of a test both speak to its validity. Where a test's adequacy is concerned, the researcher must determine how broad the scope of the assessment needs to be to

obtain reliable results. For example, if an individual student's level of competency is to be measured, a portfolio might be necessary (Hughes, 2003:87), as this would adequately show the student's full range of competency, since no individual student performs consistently over a range of assignments. As a student is likely to receive fluctuating marks for various assignments, a portfolio could assist in assigning a valid overall mark to an individual student. While it is true that tests usually focus on individual students, in the context of course evaluation, the focus moves to the programme of instruction (Henning, 1987:2-3). In such cases, individual score fluctuations should even out if the average improvement of a group is taken into consideration; therefore, a single writing assignment might be as effective as a portfolio for the intended purpose. It is also important that pre- and post-assignment conditions are the same, as this contributes to the reliability of the assessment (Carstens & Fletcher, 2009a:322; Storch & Tapper, 2009:210; Shaw & Liu, 1998:228-230). Further, the assessment must be relevant to its context to appropriately assess a set of outcomes. For example, an argumentative essay on the rise and fall of the Roman Empire might not be appropriate for a discipline-specific academic literacy intervention for science students. It is important to keep in mind that in the end, trade-offs will have to be made when ensuring the adequacy and appropriateness of assessment, depending on the test purpose and situation (Weigle, 2002:48). It is the responsibility of the evaluator to ensure that the assessment is as adequate and appropriate as possible within the limitations of the assessment context. The principles discussed above were kept in mind in the assessment tools proposed below.

This chapter proposes using a pre-test/post-test design as part of any evaluation design. Various possibilities exist within such designs, three of which stand out: using widely-used, reliable and validated academic literacy tests, language for specific academic purposes tests, and/or extended writing activities. By statistically analysing an improvement between pre- and post-tests, it should be possible to control for external variables such as students being taught by various lecturers, students' age and so forth.

4.4.1.1 Instrument 1: Reliable, validated generic academic literacy tests

Several reliable, validated academic literacy tests are widely used to assess students' academic literacy abilities, both before and after interventions. These norm-referenced

tests (see Section 2.5.3) are particularly appropriate for testing reading abilities (and the wide range of academic literacy abilities that accompany reading) (cf. Grabe, 2009:450-451). Reliable and validated academic literacy tests have shown to be positively correlated with academic performance (Cliff & Hanslo, 2009:272). However, this positive correlation fails to reach medium effect size in cases where academic literacy interventions are implemented after the writing of the academic literacy test – an outcome which is to be expected if the assumption is made that the academic literacy intervention would have a significant impact on student success (Van Dyk, 2015:180; Van Rooy & Coetzee-Van Rooy, 2015:40; Van Dyk, 2010:272).

Such tests have several advantages. These include that they make it possible to easily mark large quantities of tests and make replicating research easier than would be the case with assessments that had not been validated and verified. In addition, such tests generally have fixed and uniform procedures, materials, scoring instruments and administration processes that ensure that assessment will occur in the same manner at different places and times. Furthermore, these tests have undergone rigorous developmental cycles (Seabi, 2012:93; Nitko & Brookhart, 2011:346; Heaton, 1988:167; Beretta, 1986:434). Moreover, these tests are generally considered to be impartial, in that test items are not based on specific courses being investigated; instead, test questions are based on a predefined test construct that is seen to be fairly universal in nature (Beretta, 1986:434). Finally, these tests are ideal to determine the initial academic literacy levels of first-time students, as students entering tertiary education do not yet have the necessary subject-specific knowledge to write a language for specific academic purposes (LSAP) test (see Section 4.4.1.2 for a further discussion of this type of test).

Within the evaluation design proposed in this chapter, it would be suggested that students write a reliable and valid academic literacy test with an appropriate construct as a pre- and a post-test. The pre- and post-tests would then statistically be compared with each other to determine whether there had been a statistically significant improvement in students' academic literacy levels over the duration of the academic literacy intervention being investigated.

The Test of Academic Literacy Levels (TALL) is proposed as an example of an adequate and appropriate test¹⁸ in the South African context for various reasons. Firstly, it has a theoretically sound test construct that takes into consideration the most important aspects of academic literacy, and that closely reflects students' ability to interpret academic discourse (Van Dyk, 2005:43-44). The test construct, which is described in Van Dyk and Weideman (2004:16-17), also takes into consideration Bachman and Palmer's (1996:75-76) suggestion that tests should be constructed around language tasks instead of language skills, since in real-world situations, language is used in the completion of complex tasks, rather than being divided into the distinct skills of speaking, listening, reading and writing. Secondly, the TALL has been proven to be a very reliable test. Using Cronbach's equation, the TALL has shown an average reliability measure of 0.94 across several versions of the test between 2004 and 2010 (Van Dyk, 2010:155). Thirdly, the TALL is viewed as a valid instrument (Van Dyk, 2010:285-286; Van Der Walt & Steyn, 2007:146-156) which is standardised in terms of its scoring and administration (Petersen-Waughtal & Van Dyk, 2011:105). The TAG (Toets vir Akademiese Geletterdheid, the Afrikaans version of the TALL) has been proven to have internal validity in terms of face validity, content validity, and construct validity (including internal consistency, intra-test and inter-test validity, group differential validity and domain-specific validity) (Van Dyk, 2010:199-261). It has also been proven to have external validity in terms of predictive validity, concurrent validity, and consequential validity (Van Dyk, 2010:261-284). The same author argues that the same would be the case for the English equivalent of the TAG (i.e. the TALL) - smaller, unpublished experiments have corroborated this argument. Fourthly, the TALL has a broad distribution of marks (indicated by the standard deviation and inter-quartile range), making it easier to partition students into groups based on ability (Van Rooy & Coetzee-Van Rooy, 2015:43). Finally, it is an efficient test choice from a logistical stance. The TALL is a one-hour long test (Van Dyk, 2005:42) comprising solely of multiple choice questions. The test is therefore administered in a relatively short period of time, and its scoring is objective, time- and cost effective (Van Dyk, 2010:17; 2005:44). What is more, the format of the test makes it possible to have similar conditions for both the pre- and post-tests, which ensures that as many external variables as possible are controlled for.

¹⁸ Also see Section 2.5.3.1.1.

The limitations of this type of test must be kept in mind though. Firstly, there should ideally be some alignment between the test construct and the intervention's design. If both the test construct and the intervention's design are based on a similar comprehensive definition of academic literacy, it would follow that such an alignment would exist, thus strengthening the deductions that can be made based on the results of the test. However, as Green (2005:58) states, this type of test is not always designed with a specific intervention in mind, and vice versa. If the test construct and the intervention's design are based on different theoretical underpinnings, the evaluator should consider whether such a test would be able to say anything about the impact of the intervention. A second limitation is that it would be difficult to make deductions regarding the transfer of abilities to students' content subjects based on the results of a generic test. Thirdly, these tests usually do not expect students to, for example, produce language, build an argument, or communicate research. The TALL, for instance, does not require that students produce written text due to practical constraints in the administration of the test (Butler, 2009:294)¹⁹. Butler (2013:83) cautions that an improvement in writing might not necessarily be reflected in such tests. A fourth limitation is that these tests tend to rely on multiple-choice questions, again often due to practical constraints, but also to ensure consistent and objective scoring. Yet, the appropriateness of solely relying on this format has been questioned. One of the arguments that has been raised has been that multiple choice questions are not able to test students' normal language processing abilities – thus, the processes that students engage in when processing information and producing language, for example in written form (Scouller, 2012:469; Murphy & Yancey, 2008:366-368).

The above limitations can be addressed by triangulating data through multiple sources of evidence. One such an additional source of evidence that could be used to more validly make inferences with regards to the possible transfer of abilities to students' content subjects is language for specific academic purposes (LSAP) tests.

¹⁹ Though inter-item correlations have indicated that the multiple choice questions of the test correlated well with writing assessment; a decision was further taken not to include writing in the TALL due to logistical considerations as well as the difficulty in achieving inter-marker reliability (Van Dyk 2010).

4.4.1.2 Instrument 2: Language for specific academic purposes (LSAP) tests

Despite the advantages of generic academic literacy tests, evaluators might feel the need to use tests which reflect how academic literacy abilities are contextually applied in students' disciplines. Language for special academic purposes (LSAP) tests deal with relevant academic subjects (Krekeler, 2006:104). Such tests, which can often be classified as criterion-referenced tests (see Section 2.5.3), might use texts from various subjects such as business studies, engineering and medicine (Krekeler, 2006; Fulcher, 1999). LSAP tests are usually differentiated from general academic purposes tests by two characteristics, namely authenticity and the interaction of language and content – thus, the impact of students' background knowledge of a specific field on their ability to carry out the task in the same way that would be required in the target situation (Douglas, 2000:2). Two main reasons for using LSAP tests, according to Douglas (2000:6-8), are that specific purpose language is precise (i.e., technical language has very specific characteristics that must be controlled by people who work in the field), and that context influences language performance. It should also be added that LSAP tests might more reliably be able to indicate whether students are able to transfer academic literacy abilities to their content subjects than might be the case with generic academic literacy tests.

Though there are exceptions (e.g. the International English Language Testing System [O'Sullivan, 2012:72]), individual LSAP tests are generally not used in multiple countries or even institutions (as is the case with general academic literacy tests as discussed in section 4.4.1.1), since by their nature, LSAP tests are usually shaped to the needs of specific student groups at specific institutions. This means that there is a risk that they might not have been validated and verified to the same extent as more widely-used generic tests. Furthermore, these tests are difficult to create, since they demand more resources and skills than do generic tests (Hamp-Lyons & Lumley, 2001:129). Examples of these resources and skills include more manpower (specifically staff members with ESP experience and knowledge) and time (with concomitant financial implications), as several discipline-specific tests need to be created instead of one generic test.

It should be noted that some researchers (Davies, 2001:144; Elder, 2001:164)²⁰ have questioned whether ESP tests should be used at all, and rather argue that general academic literacy tests be used to assess ESP students' language abilities. This is, in part, due to the issues regarding verification and validation mentioned above. Some research has even indicated that ESP tests are no more effective than more general EAP tests in assessing students' command of academic language competency (see, for example, Davies, 2001:144). Criticisms such as the ones above could once again be addressed by using multiple sources of evidence (for example writing assessments and generic academic literacy tests) when evaluating the effectiveness of an academic literacy intervention. Such multiple sources of evidence could provide a much richer insight into how and why an intervention has, or does not have, an impact. A further criticism could be that an LSAP test might not be appropriate as a pre-test for first-year students, as they do not enter universities with subject-specific knowledge (also see Section 4.4.1.1). This could be addressed if the researcher can justify, empirically or theoretically, that the pre- and post-tests are equivalent.

If the evaluator does decide that using an LSAP test is necessary and justifiable, it is very important that guidelines from the literature be taken into account. Some guidelines for setting such tests include that the assessor must:

- ensure that experts in the content-subject field are involved in setting the test (Gnutzman, 2009:530);
- ensure that the test is reliable; thus, that “individuals receive the same score from one prompt or rater to the next, and if a group of examinees is rank-ordered in the same way on different occasions, different versions of a test, or by different raters” (Weigle, 2002:49) – reliability can be affected by variables related to the task and variables related to the scoring process (Bachman & Palmer, 1996:19-20);
- have an appropriate test construct that does not leave out important areas that were addressed in the intervention, or add aspects that were not the focus of the intervention (Hamp-Lyons & Lumley, 2001:131; Fulcher, 1999:225-226);
- use test items that are representative of and relevant to the domain being tested, for example by doing a needs- and content-analysis of students' other subjects

²⁰ See Weideman (2013) for similar arguments with regards to English for Specific Purposes courses.

(Douglas, 2001:176; Hamp-Lyons & Lumley, 2001:131; Fulcher, 1999:222; Bachman & Palmer, 1996:21);

- take into consideration the purpose of the test (Hamp-Lyons & Lumley, 2001:130);
- ensure that there is an appropriate perception of authenticity (i.e. face validity, or more specifically, response validity [Henning, 1987:92], so as to encourage students to complete the language tasks to the best of their ability [Bachman & Palmer, 1996: 23-24]);
- use authentic texts, authentic task types, authentic assessment criteria and authentic rater orientations (by making sure raters have a suitable specific purposes background with regard to situations in which specific purposes language will be used in tests) (Hamp-Lyons & Lumley, 2001:131; Douglas, 2000:16-18);
- ensure that the test is interactive; i.e., that it involves students' language knowledge, strategic competence, topical knowledge, and affective schemata so as to ensure that students not only display knowledge of the language, but that they also show that they can use the language in context (Bachman & Palmer, 1996:25);
- consider the impact of the test at the micro level (i.e. on the test-takers and other individuals who are affected by the test) and at the macro level (i.e. on society and the education system) (Bachman & Palmer, 1996:29);
- keep practicality in mind – thus, consider how many resources are required by the test, and how many are available in reality (Bachman & Palmer, 1996:35);
- not assume previous knowledge in the specific field and take steps to ensure that students' background knowledge of the field does not unduly skew the test results in that background knowledge is measured rather than academic literacy levels; one way of avoiding this is by explaining key terminology and concepts (Krekeler, 2006:107-108); and
- pilot the test (Krekeler, 2006:106; Douglas, 2000:5-6; Fulcher, 1999:226).

If the test has not been thoroughly verified by adhering to the above guidelines, and validated by, for example, piloting the test, data gained from LSAP tests should be used with extreme caution. Regardless, data should preferably be used as part of a larger

triangulated study in which additional sources of evidence are used to verify and validate results.

LSAP tests share many of the limitations of generic academic literacy tests, as discussed in Section 4.4.1.1. These include that the test might not be considered valid if the test construct and the intervention's design do not align, and that LSAP tests often do not require students to produce written texts, but rather rely on multiple-choice questions. A further argument for not solely using an academic literacy test when determining course impact is that, as is the case for all research, triangulating data by means of source can assist in cross-validating findings. A combination of assessment tools can also assist the researcher in viewing impact from various angles, which could lead to a more holistic impact measurement. It is therefore worthwhile to consider additional methods of assessment if the impact of an academic literacy intervention is to be measured comprehensively.

4.4.1.3 Instrument 3: Writing assignments

Being able to write well is closely connected to success in both academic and professional spheres (Weigle, 2002:4). In fact, states Weigle (2002:5), "[w]riting and critical thinking are seen as closely linked, and expertise in writing is seen as an indication that students have mastered the cognitive skills required for university work" (also see Butler's [2013:83] concern discussed in the previous section regarding relying solely on multiple choice tests). It is often also through writing that one is able to assess whether students are able to apply abilities addressed in an academic literacy intervention – such abilities might include the ability to structure information and the ability to conduct and integrate research appropriately. To determine how an improvement in writing (and the accompanying research, argumentation and other academic literacy abilities) might best be assessed, it is useful to consider some of the literature in the field.

Researchers recommend that writing assessments adhere to various principles (summarised in Figure 4.1). Firstly, students should be given a clearly defined topic which will motivate them to write and guide them in their writing endeavour (Weigle, 2002:53, 93; Heaton, 1988:137, 144). Such a topic can be adapted to suit the content area of various discipline-specific courses. Where appropriate, such a writing assessment

topic could even be shared by a content subject and the academic literacy intervention. By giving only a single topic (be that a generic topic or a subject-specific topic in the case of subject-specific academic literacy interventions), the evaluator will have a "common basis for comparison and evaluation" (Heaton, 1988:138; also see Shaw & Weir, 2007:247). The topic should be of such a nature that creativity, imagination or wide general knowledge should not play a significant role in the final writing task, and thus unduly advantage some students, since these are usually not the outcomes of general language assessment (Hughes, 2003:90). Rather, a topic about which most candidates will have similar background knowledge should be provided. Students could also be provided with the same background reading about the topic (should background reading be required), to ensure that the playing field is equal. It is advisable to take into account students' age, gender, and educational background so as to create assessments which will not be biased against, or in favour of, certain groups of students (Weigle, 2002:46). Moreover, students should be made aware of the audience they are writing for (Weigle, 2002:52; Heaton, 1988:137). Being given clear guidelines regarding the topic, purpose for writing and audience will help to avoid misunderstandings between the student and the test-givers' intent.

Furthermore, strict time limits (which are a logistical reality in most test situations) might make writing assessments more artificial and unrealistic, since writers are not able to sufficiently engage in processes like drafting and editing (Weigle, 2002:52; Heaton, 1988:138) - processes which are usually encouraged in academic literacy interventions, specifically in the widely-used process approach to writing (e.g. Jeffery, 2009:5). Allowing students sufficient time to plan, conduct research for, draft and write their academic text is therefore advisable. However, practicality must be kept in mind. Due to time-constraints, it is not necessarily feasible to give students weeks to complete an assignment. A compromise must be reached where the assessment affords students sufficient time to draft and complete a writing assignment, yet stays practical and feasible.

Marking rubrics are invaluable when assessing writing tasks, as they are one of the best ways of limiting (though not eliminating) inconsistencies in marking (Weigle, 2002:108-139; Heaton, 1988:148). Rubrics attempt to separate various features of written text for scoring purposes (Weigle, 2002:122; Heaton, 1988:148). This type of marking is ideal

for the purposes of this study, which aims to assess whether there is an improvement in specific academic literacy abilities.

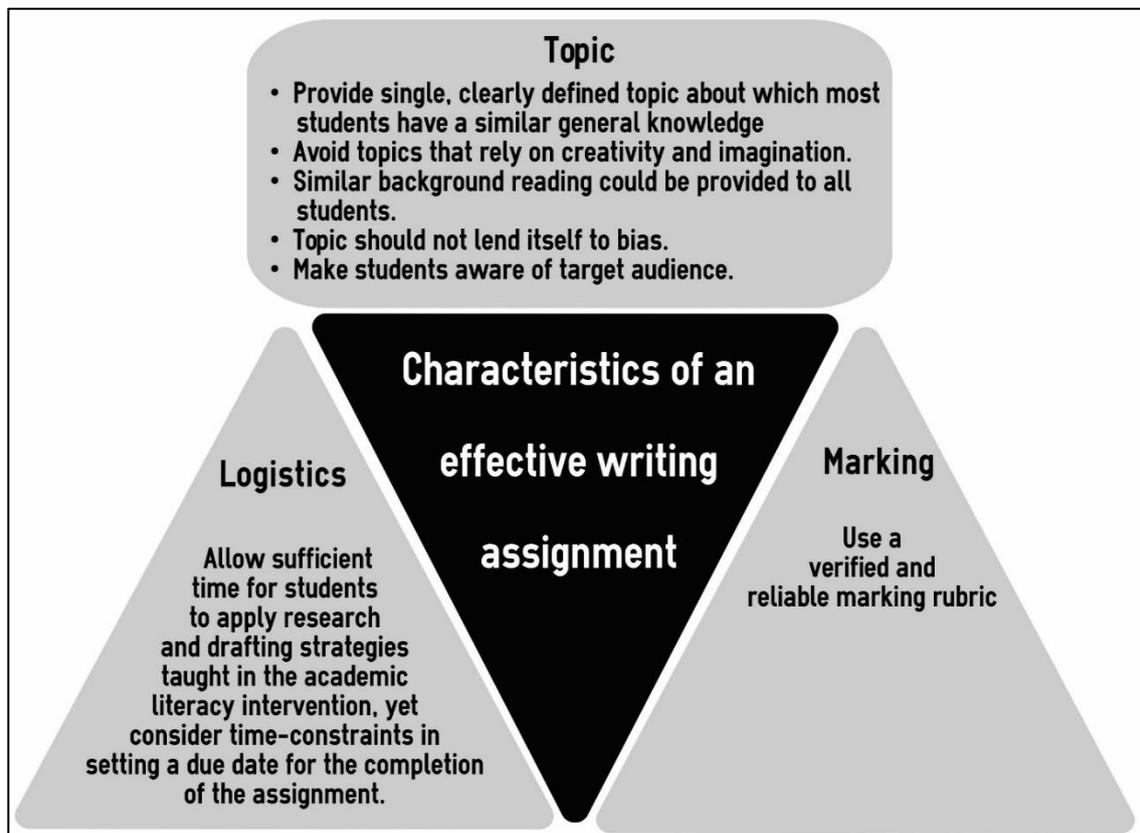


Figure 4.1: Characteristics of an effective writing assessment

Keeping the principles discussed in this section in mind, this study proposes that an extended²¹ writing assessment be completed. This assessment could be generic or subject-specific in nature, depending on the type of academic literacy intervention in question (i.e. a generic or subject-specific academic literacy intervention). For example, in the case of a generic academic literacy intervention, a pre- and post-assignment essay-type question could be used. In such a case, a 300-word text of argumentative writing on the same or on a similar topic could be written at the beginning and at the end of the intervention. The conditions of the pre- and post-assignments should be similar, so as to exclude as many external variables as possible. Students could be provided with a topic as well as sources which they could use to support their argument. As mentioned earlier though, a framework for assessing the impact of an academic literacy intervention should

²¹ “Extended” here refers to an assignment that is longer than a typical paragraph answer. For the purposes of this study, a 300-word assignment is considered “extended” at first-year level. Longer assignments might, however, be required for more advanced students.

be flexible. Although argumentative writing is a commonly used genre in academic literacy courses and university writing in general (Zhu, 2009:34; Hyland, 2004:18, 122; Paltridge, 2004:87; Yeh, 1998:49-51), academic literacy interventions may also focus on other types of writing, for example laboratory reports, discursive writing or business report writing. The main reason for using other types of writing would be to determine whether students are able to apply the writing conventions addressed in the academic literacy intervention within their respective subject fields – thus, whether transfer is likely to have taken place. Also, some academic literacy interventions might not focus on specific abilities like research, referencing and synthesising. Thus, depending on the nature of writing required in the academic literacy intervention, the genre, topic and structure of the assignment itself should be adapted to suit the outcomes of the intervention.

An example of a comprehensive rubric that could be used to assess such an extended writing assignment is given in Appendix A²². This rubric, which was adapted from Carstens and Fletcher (2009b:59-60), is based on three analytic rating scales that are internationally accredited and comprehensively covers a wide variety of general academic literacy outcomes under the headings "structure and development", "academic writing style", "editing" and "use of source material". The combined essay marks, as well as the marks of various sections in the rubric, can be statistically compared with each other for both the pre- and post-assessment writing pieces.

An alternative to this rubric is Renkema's (2001:40-44; 1998:29-31) C3 model. This model considers five text levels at the hand of three criteria, namely "correspondence", "consistency" and "correctness". The five text levels are "text type", "content", "structure", "formulation" and "presentation". A summary of this model can be found in Appendix B, and an in-depth discussion thereof can be found in Carstens and Van de Poel (2012:58-79).

An additional option when assessing student writing is to follow the quantitative strategies employed by Storch and Tapper (2009:210-212) so as to cross-validate and triangulate findings. Their study quantitatively measures students' fluency (by looking at

²² This rubric was adapted after a piloting process in which it was given to several colleagues to determine whether it covered all the aspects that such a rubric would typically need to cover.

words per T-unit), their accuracy (by counting errors in various categories), and their use of academic vocabulary (by comparing student lexis to the Coxhead [2000] academic wordlist). Appendix C provides more information on their methods.

4.4.2 Assessing the extent to which academic literacy abilities are necessary and applied in students' content-subjects

The second aspect of the definition of impact given in Section 4.2, namely the extent to which academic literacy abilities are necessary and applied in students' content subjects, is yet to be addressed. That is the aim of this section. Butler (2013:82) asks a critical question: does an improvement in generic test scores necessarily imply that these improved academic literacy abilities will be transferred to students' content subjects? Based on the literature surveyed in Chapter 3, the most common solution to this problem seems to be measuring students' perceptions regarding such transfer by means of questionnaires. The main objective of a questionnaire is to gather facts and opinions about a particular phenomenon from people who have knowledge about the specific issue (De Vos *et al.*, 2011:186). Moreover, "[a]sking questions is widely accepted as a cost-efficient (and sometimes the only) way of gathering information about past behaviour and experiences, private actions and motives, and beliefs, values and attitudes" (Foddy, 1993:1)²³. Two aspects that can be determined by means of questionnaires are 1) students' perceptions regarding which academic literacy abilities they acquired in the academic literacy intervention (and the extent to which these were acquired), and 2) students' academic literacy needs in content subjects according to students and lecturers²⁴.

Although surveys can use a variety of rating scales, the Likert scale is widely used. This five-point scale is commonly used to determine the extent to which participants agree with a statement of attitude or opinion (Henning, 1987:23). Likert scale questionnaires "are particularly effective in that they elicit information in a manner that permits quantification and comparison with other programs or with other features of the same

²³ Such information should however be combined with empirical evidence regarding an improvement in academic literacy abilities, as discussed in Section 4.4.1. Only when considering multiple sources of evidence can valid inferences be drawn.

²⁴ Using additional sources of evidence, such as an analysis of study guides or diagnostic test data, would be valuable in corroborating such results.

program" (Henning, 1987:143). It would thus seem appropriate to use the Likert scale in the questionnaires suggested below. In addition to Likert scale questions, surveys can also contain open-ended questions through which qualitative data could be obtained. This qualitative data would be another source of evidence. Other forms of qualitative data are discussed in more detail in Section 4.4.2.5.

In addition to using questionnaires (discussed in more detail in Sections 4.4.2.1 and 4.4.2.2), some valuable information regarding which academic literacy abilities are necessary and applied in students' content-subjects could be obtained by doing an analysis of students' content-subject study guides, tests and examinations (see Section 4.4.2.3), correlating academic literacy results to other variables (see Section 4.4.2.4) and obtaining qualitative feedback from primary stakeholders (see Section 4.4.2.5). It is important that an appropriate balance be reached in the use of these instruments. One instrument cannot necessarily replace another – rather, by using a combination of several of these instruments, different aspects will be illuminated in the evaluation process. These instruments are discussed in more detail in the remainder of this section.

4.4.2.1 Instrument 4: Determining academic literacy ability acquisition and academic literacy needs in content subjects – student questionnaire

According to Carstens (2009:324), it is not enough to only measure whether there was an improvement in students' academic literacy marks when determining whether an academic literacy intervention was successful. The success of such interventions, they argue, is equally dependent on how students perceive the interventions and the abilities addressed therein, as this determines, at least in part, students' motivation and the extent to which abilities are transferred. It is particularly important to determine whether students consciously transfer abilities acquired in the academic literacy classroom to their content-subjects²⁵. One way of doing that is by asking students about their behaviours. Lynch (1996:169) states that one can approach student attitudes towards a programme in two ways: 1) from a judgemental viewpoint, where the degree of students' satisfaction with the programme is measured, and 2) in a descriptive fashion, where the evaluator

²⁵ It is, however, important to keep in mind that there might be a discrepancy between students' belief that they are transferring abilities, and them truly transferring these abilities in practice. That is a shortcoming of the instrument proposed here.

aims to understand the nature of students' satisfaction and/or dissatisfaction. It is also possible to address both of these by including both closed-ended Likert-scale-type questions as well as open-ended descriptive questions in a student questionnaire; by doing this, both quantitative and qualitative data can be collected.

Some studies use in-house official feedback forms to assess students' perceptions regarding academic literacy interventions (e.g. Van Dyk *et al.*, 2011a:499-500); these forms are usually aimed at assessing the course and the lecturer. Such forms are often the only tools available to lecturers to gauge students' perceptions of specific courses. Although there is clearly value in such feedback forms, they rarely allow for detailed information regarding students' opinions about the value of various abilities addressed in the intervention. They also do not allow students to indicate which academic literacy abilities they believe would be of value to them. One type of survey that seems particularly relevant in determining perceptions regarding transfer is a "need-press" questionnaire (Rao & Saxena, 2014:22-25; Lynch, 2003:68). In this type of survey,

[participants] are asked to judge the importance (need) of particular ... skills or abilities, and then to judge their emphasis (press) in the teaching programme. By comparing judgements of how important something is perceived to be with how much attention it receives in the instructional setting, areas of individual learner development and programme objectives that may need improvement are identified (Lynch, 2003:68).

Students could be asked to which extent they "need" certain academic literacy abilities in content subjects by the end of their first year, and again by the end of their second year. By combining this with students' perceptions of the extent to which these abilities were emphasised in a specific instructional setting (i.e. the academic literacy classroom), valuable deductions could be made regarding important abilities that were perceived to have been transferred (those that score highly on both the "need" and "press" aspects), abilities which are needed but not taught sufficiently (those that score highly on the "need" aspect, but low on the "press" aspect), and abilities which are taught, but not needed (those that score low on the "need" aspect, but high on the "press" aspect). By addressing both of these aspects, Kiely's (2009:108) conception of a learner questionnaire fulfilling the purpose of understanding the intervention in question in terms of the quality of the learning experience of students would probably be successfully attained.

The current study suggests that Van Dyk's (2014) Questionnaire on Academic Literacy be used as the basis of such a need-press questionnaire. The Questionnaire on Academic Literacy addresses a wide range of abilities that could potentially be addressed in an academic literacy intervention. This questionnaire has been revised to meet the needs of the current study, and can be found in Appendix D as the Need-Press Questionnaire on Academic Literacy Abilities.

One method in which data from such a questionnaire could be triangulated is by comparing the results from the questionnaire to empirical data. Van Dyk, Van de Poel and Van der Slik (2013), for example, compared students' self-reported reading preparedness to their reading profiles as measured by the TALL and the TAG. One of the findings of this study, namely that perceived preparedness does not reflect actual student preparedness (Van Dyk *et al.*, 2013:363)²⁶, should serve as a warning that students' perceptions are not necessarily always valid, and thus it is advisable that a questionnaire be used as part of a triangulated study.

As can be inferred from the Van Dyk *et al.* (2013) study, a possible objection to using students' perceptions in determining the impact of an academic literacy intervention might be that information might not be entirely reliable; yet, argues Lynch (2003:131), "these recalled experiences may have a legitimate place in constructing an (...) account of the meaning of the programme". It must further be remembered that although students might feel that certain aspects addressed in an intervention are unnecessary, lecturers might disagree. Some abilities serve as building blocks for more complex learning later on, and students are not always the best judges as to which abilities will be necessary for them to ultimately achieve more complex tasks. In cases such as these, additional feedback should preferably be obtained from subject experts. A further reason for obtaining additional feedback is that this type of data is subjective. Judd and Keith (2012:40) argue that "[t]his source of data, while not ideal, can still contribute to our assessment of a given outcome through the method of triangulation, in which multiple sources of evidence used together measure a construct in corroboration". Additional sources of feedback are addressed below.

²⁶ It would have been interesting to repeat the same procedure as a post-test, to determine whether students' perceptions and their reading ability became more aligned after a year of university study.

4.4.2.2 Instrument 5: Determining academic literacy needs in content subjects – lecturer questionnaire

Although determining students' perceptions regarding the academic literacy abilities which are needed for them to succeed in their studies is certainly a valuable aspect in determining the relevance and ultimate impact of an academic literacy intervention, it is also important to cross-validate this information with that from additional sources.

One way in which this could be done is by asking content-subject lecturers which abilities they deem most vital for students to succeed in their respective subjects. In addition, academic literacy lecturers could also be consulted on which abilities they believe students need to master for them to succeed in their studies. Although lecturers' opinions on which abilities students need to acquire for them to be successful in their content subjects is not a sufficient source of information for an evaluator to make deductions about the transfer of abilities, the information obtained from lecturers (as is the case with instrument 6 discussed below) could indicate whether the academic literacy intervention in question focuses on appropriate abilities. If lecturers mention abilities they deem to be vital, but these abilities are not addressed in the academic literacy intervention, this would speak to the impact that the intervention could potentially have.

This study proposes that a revised version of the Questionnaire on Academic Literacy (Van Dyk, 2014) that was suggested in the previous section be modified into a questionnaire that can be given to content-subject and academic literacy lecturers to determine which abilities they believe an academic literacy intervention should address in supporting students to achieve content-subject outcomes. This questionnaire can be found in Appendix E. This questionnaire also contains Likert scale as well as open-ended questions, so as to ensure that empirical as well as descriptive data²⁷ are gathered (see Lynch [1996] discussed in Section 4.4.2.1). In fact, this questionnaire is very similar to the student questionnaire discussed in Section 4.4.2.1, and purposefully so – it is very important that these two questionnaires correspond with one another, so that findings can easily and validly be compared in a triangulated study.

²⁷ Section 4.4.2.4 discusses further methods of gathering qualitative data from stakeholders.

If students show an improvement in certain academic literacy abilities, students and lecturers believe that these abilities are necessary in students' content subjects, and students believe that they use these abilities in their content subjects, this would be a fairly good indication that transfer of abilities is likely. Conclusions regarding transfer could, however, be strengthened by using some of the instruments discussed below.

4.4.2.3 Instrument 6: Content analysis of study material

Another option of determining which academic literacy abilities students need to utilise in their content subjects, in addition to the instruments above, would be to do a content analysis of students' study material, specifically their assignments, tests and examinations, as was done in Fouché (2009:83-101). In this study, the assignments of students in Unisa's Science Foundation Programme were analysed. First, appropriate subjects (i.e. that required some writing as part of their assignments) were identified. Assignments were then categorised based on Bloom's Taxonomy (Fouché, 2009:95-100). An advantage of considering Bloom's Taxonomy as part of a content analysis is that the researcher can determine whether the cognitive levels addressed in the academic literacy intervention correspond to those that are required in students' content subjects at, for example, first-year level. Carstens (2008) similarly analysed study guides from a representative sample of departments in the Faculty of Humanities at a major South African university. This was done in order to establish an empirical base for designing an undergraduate academic literacy course. She found that essays and critical analyses were the preferred modes of writing in the humanities, while project reports and essays were favoured by the social sciences. The most frequently used rhetorical modes at undergraduate level were identified as discussion, analysis, argumentation, explanation and description.

This method might be best employed in conjunction with one (or both) of the tools discussed above, namely determining the perceptions of students and lecturers. Though valuable information might be obtained from a content analysis of a variety of assignments and tests, it is likely that certain aspects of academic literacy might be overlooked. For example, would such a content-analysis indicate whether students need assistance with note-taking strategies, cohesive devices or academic vocabulary? Furthermore, such an analysis might highlight certain abilities that are indeed necessary

for students to complete assignments and tests, but it would be very difficult to determine whether students actually experience difficulty with these abilities; for example, the content analysis might indicate that students need to use subject-specific vocabulary, but content-subject lecturers might feel that students experience no difficulties with subject-specific vocabulary, and do not need additional guidance in this ability. Conversely, lecturers and students might not always consider the full range of academic literacy abilities that might be required in their content-subjects - a content analysis of assignments might prove valuable in such cases. Furthermore, conducting a content analysis or a lecturer survey could indicate how a course could be refined so as to have a greater impact in future. By using a combination of these methods, one instrument could act as confirmation for or enrichment of the other, leading to more reliable results.

This instrument has various limitations. A content analysis of study material might not be feasible or useful in all academic literacy intervention evaluations. Specifically, if the intervention services students from a broad range of fields or subjects, it becomes almost impossible to consider the academic literacy requirements of individual subjects that would result from such a content analysis²⁸. This instrument might therefore be more useful for discipline-specific academic literacy interventions where there should be a close correlation between the outcomes of the academic literacy intervention and the content-subjects it services. A further limitation is that a content-analysis in itself cannot indicate that academic literacy abilities were transferred to students' other subjects. At most, it can indicate whether the abilities addressed in the academic literacy intervention correspond to those that students need to be successful in their studies. It can thus indicate gaps that exist in an academic literacy intervention's design, which might in turn limit the impact of such an intervention.

4.4.2.4 Instrument 7: Correlating academic literacy results with other variables

Another way of gathering information regarding possible transfer of academic literacy abilities is by correlating academic literacy results with other variables. For example, Van Rooy and Coetzee-Van Rooy (2015:37-38) correlate, amongst others, the results of

²⁸ Such a content analysis could, however, assist in confirming that the general academic literacy abilities that are addressed in a generic course are, in fact, drawn on in students' mainstream subjects.

an academic literacy course with students' general academic achievement (also see Van Dyk, 2015:178-180; Mhlongo, 2014:80-82; Van Dyk, 2010:262-274). In their study, there was a positive correlation (with large effect sizes of between 0.56 and 0.77 for the two semesters of the academic literacy intervention) between students' academic literacy results and their marks in content subjects. Caution should be used before relying on such a correlation to make assumptions about the transfer of abilities (as is the case with all the tools measuring transfer, as discussed in this chapter). Although it is possible that improved academic literacy abilities could result in better marks in students' content subjects, it is also possible that students who are academically strong would naturally perform better in both measurements, whilst students who are academically weaker would naturally perform poorer in both. Despite this limitation, this measure could provide valuable insight into the effect of an academic literacy intervention as part of a triangulated study.

Another correlation that could be drawn is between students' performance in an academic literacy intervention and their class attendance (e.g. Fouché, 2009:133). In theory, if more frequent academic literacy class attendance is correlated with higher academic literacy levels, this should indicate a causal relationship between the intervention and improved academic literacy levels. However, as in the above measure, this should be used with caution and only as part of a triangulated study, as variables such as student motivation could skew the results of such a correlation. It should also be noted that this correlation would not address the transference of abilities as such – at most, it might provide information regarding the causality of the course with regard to an improvement in the academic literacy scores obtained from the instruments proposed in Section 4.4.1.

4.4.2.5 Instrument 8: Qualitative feedback from primary stakeholders

A further approach to determining whether a specific academic literacy intervention met the needs of students' content subjects, and whether the academic literacy intervention needs to be adapted, is to obtain qualitative feedback from content-subject lecturers and academic literacy specialists (two types of primary stakeholders with regard to academic literacy interventions), as was done in the Winberg *et al.* (2013) study. This instrument lends itself particularly well to collaborative academic literacy interventions. Based on

the Winberg *et al.* (2013) study, the following approaches are cited as examples of obtaining such qualitative feedback:

- regular meetings between content-subject and academic literacy specialists could be held throughout the intervention with the aim of exploring the successes and shortcomings of specific interventions;
- debriefing meetings between academic literacy and content-subject specialists could be held so as to reflect on an academic literacy intervention;
- participants could reflect on the effectiveness of an intervention by means of observations at various stages of a collaborative intervention;
- semi-structured interviews could be held with content-subject specialists, which would in turn be qualitatively analysed by looking for emerging themes; and
- feedback on the success of interventions could be obtained through focus group sessions, narrative interviews or reflective writing – any of which would again be qualitatively analysed.

An additional qualitative instrument that could provide insight into the likelihood of transfer in interventions where academic literacy specialists collaborate with content-subject specialists is by describing the level of collaboration in terms of eight dimensions, as proposed by Carstens (2013a:118-119). This instrument also obtains feedback from two types of primary stakeholders, namely academic literacy specialists as well as content-specialists. The eight dimensions considered in this tool are authorship, autonomy, collaboration, teaching staff, content integration, curriculum for academic literacy activities, materials, and assessment. These eight dimensions are represented in a table where the level of collaboration for each is indicated on a sliding scale, from being “most collaborative / most integrated” to being “least collaborative / most autonomous”. If there were a high level of collaboration across each of these eight dimensions, it could indicate that transfer might have been more likely. However, this method would ideally have to be used in conjunction with others, for example some of the approaches used in the Winberg *et al.* (2013) study. At most, Carstens’ (2013a) tool could shed additional light on possible transfer of abilities; used in isolation, this instrument is unlikely to allow valid deductions regarding transfer.

A further approach that could be followed is to obtain qualitative feedback from academic literacy lecturers involved in the intervention. In Van Dyk *et al.* (2009:334-335), for example, lecturers were asked to note the difference in execution between a pre- and post- writing assignment. As part of a triangulated study, such an approach could provide valuable insight into an improvement in student writing.

4.5 An evaluation design for academic literacy interventions

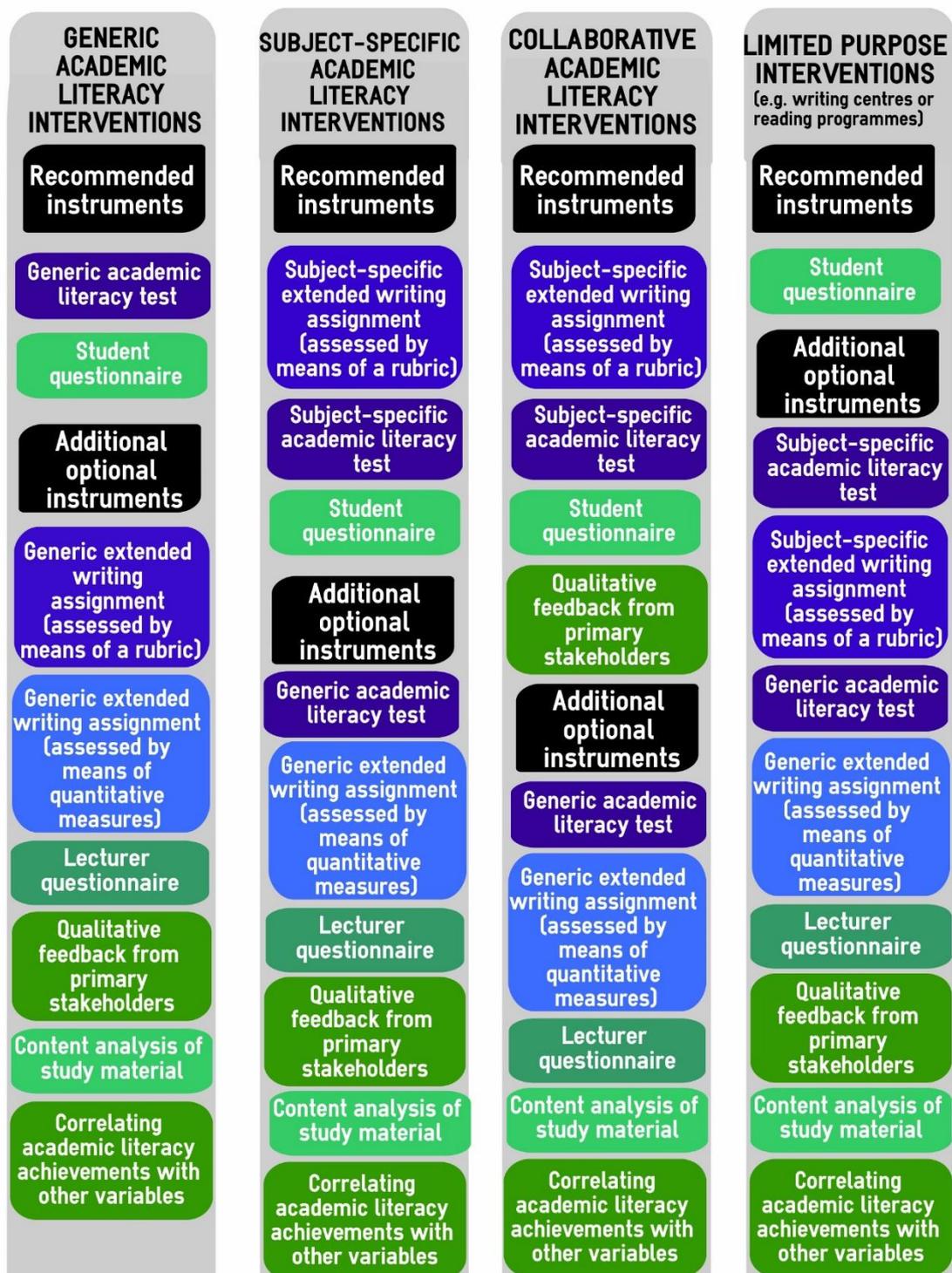
According to Kiely (2009:99), “[p]rogramme evaluation is a form of enquiry which describes the achievements of a given programme, provides explanations for these, and sets out ways in which further development might be realized”. The instruments discussed above could all contribute towards this aim.

Based on a) the definition of impact set forth in Section 4.2 and b) the approaches discussed in the previous section, this study proposes that two broad facets of impact must be considered to determine the impact of an academic literacy interventions on student success. Firstly, the evaluator must determine whether there was an improvement in students’ academic literacy levels, and the extent of such an improvement. Secondly, the evaluator must determine whether, and to which extent, these abilities are used in, and transferred to, students’ mainstream subjects.

Mhlongo (2014:4) raises an important point in this regard. He states that academic literacy interventions must be assessed in the context of addressing the specific needs of students at various higher education institutions, as different institutions have students with unique needs. An evaluation design for assessing the impact of academic literacy interventions would therefore have to be flexible, so as to accommodate a wide range of academic literacy interventions that vary vastly in terms of, for instance, their content and purpose. Such a design would probably have to include specific core components which would address aspects that should form an integral part of any academic literacy intervention. However, the researcher should also have the freedom to choose between various research instruments, and even to adapt these instruments to his/her needs, in order to assess the impact of a specific academic literacy intervention as effectively as possible.

As argued by Judd and Keith (2012:36), using a valid framework (or evaluation design) is possibly the best way of ensuring that change (for example in students' academic literacy levels over a period of time, and the transference thereof) is indeed due to a particular agent (for example an academic literacy intervention). By using an appropriate combination of instruments in such a design, Lynch's (2003:152) suggestion that multiple measures be used, and data be triangulated not only by method, but also by source, can be followed (Lynch, 1996:59-60; Jick, 1979:602). Moreover, using multiple measures contributes to the validity of findings. According to Lynch (2003:148), validity in the context of evaluation refers to "the extent that our evaluation findings provide credible answers and solutions for our evaluation questions and goals". Using multiple sources of evidence (as would be the case in the proposed evaluation design) would contribute to the validity of the findings in the evaluation of an academic literacy intervention (Judd & Keith, 2012:40; Lynch, 1996:59-61).

Based on the variety of potential evaluation instruments discussed above, the evaluation design depicted in Figure 4.2 is proposed as a tool which evaluators could use as a guideline for choosing appropriate evaluation instruments for a variety of contexts. It is suggested that findings are validated wherever possible by means of triangulating various options provided for the four main groups of academic literacy interventions, namely generic interventions, subject-specific interventions, collaborative interventions and limited purpose interventions (see Van de Poel & Van Dyk, 2015:169-173). The current study proposes using at least two of the suggested instruments in cases where control groups can be used. When the use of control groups is not feasible, it is recommended that at least three different instruments are triangulated so as to draw valid conclusions. Of these instruments, at least one should ideally measure an improvement in students' academic literacy abilities, and at least one should measure the transfer of these abilities to students' other subjects. By using such a combination of instruments, the evaluator will be in a position to triangulate by both method and source (cf. Lynch, 1996:59-60; Jick, 1979:602). Note that these are the minimum guidelines. If the context and resources allow for more instruments to be used, this should be done, as it would strengthen any conclusions drawn based on the triangulated data from these instruments.



- Where control group studies are possible to strengthen the research design, these should be used in combination with at least two instruments above.
- Where control group studies are not possible, findings should be validated by means of triangulation by choosing at least three of the options above when evaluating any given academic literacy intervention.
- In both scenarios above, one of the instruments should ideally be an instrument that measures the improvement in academic literacy abilities (coloured blue), and at least one should be an instrument that measures transfer (coloured green).

Figure 4.2: Proposed evaluation design for academic literacy interventions

It is important to note that there are likely to remain limitations, regardless of combination of instruments used, depending on the type of academic literacy intervention being assessed. For instance, even if an assessment of a generic course indicates that students improved over a range of academic literacy abilities (by means of, for example, a pre- and post-test design using a standardised academic literacy test and a writing assessment), and even if it can be shown that the abilities addressed in the academic literacy intervention are necessary in students' studies (by, for example, drawing on a content analysis and student and lecturer questionnaires), and students believe that they applied the abilities they acquired in the academic literacy intervention in subsequent years (by, for example, a need-press questionnaire), there would still be no undisputable empirical proof that students transferred the abilities acquired in the academic literacy intervention to their content-subjects. Though triangulating data would go far in strengthening inferences made with regard to transfer, the remaining limitations would have to be acknowledged.

Some interventions, such as those using collaborative approaches, might lend themselves better to claims of transfer of abilities (for example by using qualitative feedback from primary stakeholders), yet have their own set of limitations that make their implementation difficult (see Van de Poel & Van Dyk, 2015:169-173). Ultimately, the researcher remains responsible for ensuring that as comprehensive a combination of instruments as possible is used in order to strengthen any inferences that are made about the intervention's impact.

4.6 Conclusion

Weideman (2013:20) states that "as a technical artefact, a language course undoubtedly has to be effective or valid, consistent, differentiated, appealing, theoretically defensible, yield meaningful results, be accessible, efficient, accountable, and so forth". Kiely (2009:99) adds that programme evaluation is a "socially-situated cycle of enquiry, dialogue and action". This chapter has aimed to contribute to the discussion regarding the impact and accountability of academic literacy interventions by considering eight instruments which could be employed to determine the impact of such interventions. The instruments were divided into two broad categories, namely instruments that determine whether there was an improvement in students' academic literacy levels, and instruments

which determine which academic literacy abilities were needed in students' content subjects and whether these abilities were transferred to students' content subjects. Finally, an evaluation design was proposed which could be used to validly assess the impact of a range of academic literacy interventions.

In the chapters that follow, this proposed evaluation design will be verified and validated by 1) using it to assess the impact of a specific academic literacy intervention at a South African university (Chapters 5 and 6), and 2) asking academic literacy intervention coordinators across South Africa whether, and to which extent, the proposed evaluation design meets their needs and how it could be refined so as to be effectively applied to their specific contexts (Chapter 7). A final evaluation design will then be proposed (Chapter 7). Ultimately, this study hopes to provide an effective and flexible structure with which evaluators can “[describe] the achievements of a given programme, [provide] explanations for these, and [set] out ways in which further development might be realized” (Kiely, 2009:99).

CHAPTER 5

IMPACT MEASUREMENT: QUANTITATIVELY DETERMINING THE IMPROVEMENT IN STUDENTS' ACADEMIC LITERACY LEVELS AT A SOUTH AFRICAN UNIVERSITY²⁹ (ARTICLE 3)

5.1 Prelude to Chapter 5

In Chapter 4 of this dissertation, an evaluation design was proposed which could be used in the assessment of a variety of academic literacy interventions. This flexible design consisted of a range of instruments from which an evaluator could select. It was recommended that at least three different instruments be used in cases where control groups were not feasible, and at least two different instruments in cases where control groups were present. In each case, it was recommended that at least one of these instruments measure the possible improvement in students' academic literacy abilities, and at least one measure the extent to which these abilities are necessary in, or transferred to, students' content subjects (see the definition of impact given in Section 5.2).

The instruments from which an evaluator could select an appropriate combination are as follows:

- Generic academic literacy test
- Subject-specific (LSA) academic literacy test
- Generic extended writing assignment (assessed by means of a rubric)
- Subject-specific extended writing assignment (assessed by means of a rubric)
- Quantitatively assessing writing assignment
- Student questionnaire
- Lecturer questionnaire
- Content analysis of study material
- Correlating academic literacy achievements with other variables
- Qualitative feedback from primary stakeholders

²⁹ This article has been submitted for publication. The remaining two articles (articles 4 and 5) will be submitted once article 3 has been accepted for publication.

In this chapter, as part of an attempt to verify the design, a selection of instruments is used to determine whether there was an improvement in students' academic literacy levels in an academic literacy course at the North-West University's Potchefstroom Campus. This chapter considers the first semester of this course³⁰.

The instruments that are used in this stage of the study and which were considered relevant to the specific context of the academic literacy course being investigated are a generic academic literacy test, a generic extended writing assignment assessed by means of a rubric, a generic extended writing assignment assessed by means of quantitative measures, and correlating academic literacy results with other variables³¹.

5.2 Introduction

Academic literacy interventions are commonplace in most South African universities (Sebolai, 2014:52; Fouché, 2009:41-42), largely due to inadequate secondary education in the country (Cliff, 2014:322; Davies, 2010:xi) and the trend towards massification of tertiary education (Calderon, 2012; Teichler, 1998). However, few studies provide evidence of the impact of these interventions (Sebolai, 2014:52; Butler, 2013:80; Terraschke & Wahid, 2011:174; Carstens & Fletcher, 2009b:319; Storch & Tapper, 2009:218; Holder *et al.*, 1999:20). The current study defines impact (or effect, which is used synonymously in this chapter) as “i) the observable improvement in academic literacy abilities between the onset and the completion of an academic literacy intervention and ii) the extent to which these abilities are necessary and applied in students' content subjects” (see Section 2.2.1; Fouché, 2015:21). This chapter only considers the first aspect, namely measuring an observable improvement in students' academic literacy levels. This aspect, however, is investigated by using several methods.

Yeld (2010: 26) states that in recent years, there has been an increasing focus on graduation rates, as well as the effectiveness and quality of higher education. Academic literacy interventions thus firstly have the responsibility of indicating how they contribute to effectively preparing students for the literacy demands of higher education. Secondly, they

³⁰ Chapter 6 considers both semesters of the course, and also investigates to which extent the various abilities addressed in the course are needed in, and transferred to, students' other subjects.

³¹ Chapter 7 discusses the problems encountered with the various instruments in detail, and were applicable suggests improved research instruments or additional guidelines in using existing instruments.

have the responsibility to show what measurable effect they have on students' academic literacy levels. It is also important to determine areas in which they do not have a sufficient impact so that re-orientation can occur, thus ensuring that academic literacy interventions are optimally effective (Newcomer & Triplett, 2010:6). In fact, the evaluation process should ideally be seen as a continuing needs assessment that constantly works towards improving the academic literacy intervention in question (Bachman & Palmer, 2010:25; Brown, 2001:15).

This chapter forms part of a larger study in which an evaluation design is proposed which can be used to assess the impact of a variety of academic literacy interventions. Two previous chapters have, firstly, considered national and international studies that have aimed to assess the impact of academic literacy interventions in the past (Chapter 3), and secondly, proposed a flexible, comprehensive design that could be used to validly and reliably assess the effectiveness of a variety of academic literacy interventions, including discipline-specific interventions, generic interventions, writing centre interventions as well as reading programmes (Chapter 4). This is the first of two chapters which will use a variety of instruments proposed in Chapter 4 so as to ultimately assess the impact of a specific academic literacy intervention at the Potchefstroom Campus of North-West University. Thereafter, the evaluation design will be implemented and further refined, based on feedback received from stakeholders and the researcher's practical experience after having implemented the design. Ultimately, the study aims to contribute to addressing the research gap regarding the impact of academic literacy interventions, thereby strengthening the field of academic literacy in South Africa.

5.3 Background

At the North-West University's Potchefstroom Campus, all first-year students are required to write the Test of Academic Literacy Levels (TALL) or its Afrikaans equivalent, the *Toets van Akademiese Geletterdheidsvlakke* (TAG)³². Students who are identified as 'at-risk' by the TALL are required to complete a semester-long academic literacy course. After completing this course, they join the students who were not identified as at-risk for a

³² The current study only focuses on the students who study through the medium of English, and will thus refer to only the TALL and the concomitant English academic literacy courses from this point onward.

second semester-long academic literacy course (which is compulsory for all first-year students). Both of these courses are taught in the students' first year of studies.

The academic literacy intervention which is the focus of the current study is the first semester of a course called 'An introduction to academic literacy' (AGLE111). A total of 624 students registered for this course in 2015. The following outcomes were taken from the AGLE111 weekly lesson plan and serve to illustrate the nature of the course:

1. Identifying word meaning from context;
2. Paraphrasing text;
3. Making effective notes from presentations and reworking these notes to paragraphs and mind maps;
4. Including references in a text;
5. Understanding academic genres and identifying and finding reliable academic sources;
6. Explaining the concepts of active reading, skimming and scanning;
7. Using skimming and scanning to obtain information from texts;
8. Identifying the qualities of, and being able to write good introductions and conclusions;
9. Creating a table of contents, and using it to plan and structure text;
10. Writing paragraphs with clear topic sentences, one main idea and applicable support;
11. Identifying action words and content words in examination questions and assignments, and planning well-structured responses to examination questions;
12. Identifying reasons for using the passive voice;
13. Identifying inaccurate information;
14. Writing correct sentences;
15. Calculating basic percentages;
16. Explaining and being able to identify visual manipulation;
17. Referring correctly to different parts of graphs and tables;
18. Identifying reasons for using graphic information, analysing graphics and discussing graphics appropriately;
19. Being aware of the structure of a seminar, being able to ask effective questions, and being able to answer questions effectively; and
20. Distinguishing between open, closed and hypothetical questions.

This generic academic literacy course, which is worth 12 credits, services all the faculties at the Potchefstroom Campus of the North-West University (thus, students belong to faculties as diverse as Arts, Natural Sciences, Theology, Education Sciences, Economic and Management Sciences, Law, Engineering, and Health Sciences). Students attend two one-hour classes per week which are embedded in the curriculum. The duration of the course is from February to May.

The impact of AGLE111 has not yet been formally assessed. That, then, is the aim of the remainder of this chapter. Firstly, the design and method of the study will be discussed in Section 5.4. The results of the study are thereafter discussed in Section 5.5.

5.4 Design and method

Four instruments were identified from the evaluation design proposed in Chapter 4. These instruments were considered appropriate as they suit the generic nature of the subject and were considered feasible, taking into account the wide range of faculties and large number of students serviced by this subject. In addition, they could be easily integrated into the subject's existing assessment plan. These instruments are used in tandem, and results are ultimately triangulated, to determine whether there was an improvement in students' academic literacy levels between the onset and the completion of the AGLE111 academic literacy course. These were:

1. Using a generic academic literacy test;
2. Using a generic extended writing assignment (assessed by means of a rubric);
3. Quantitatively assessing an extended writing assignment; and
4. Correlating academic literacy achievements with students' results in their other subjects.

A sample of 173 (of the 624 registered students) was used for the first and the fourth instruments. A sample of 139 students was considered for the second instrument, while a smaller sample of 50 students was considered for the third instrument, since this instrument was very work-intensive.

Although a random sample was initially identified for the second and third instruments, this random sample had to be abandoned due to difficulty in obtaining electronic versions of students' post-writing assignments. Ultimately, a sample of convenience was identified based on students who had submitted both pre- and post-writing assignments electronically. The pre-writing assignment results of this sample of convenience were compared to the pre-assignment marks of the initial random sample to determine whether these groups could be considered comparable (and thus, whether the sample of convenience could be said to be representative of the entire population). The rubric (instrument 2 of this chapter) was used, and effect sizes were calculated for the difference in performance across the four main sections addressed in this rubric. A very small effect size (Cohen's $d = 0.08$) was found between the two groups' results, indicating that there was no practically significant difference between the groups. Thus, based on the results for the pre-writing assignment using the writing rubric, these two groups seem to be comparable, which means that it is likely that the sample of convenience can be said to be representative of the entire population.

The following section reports on the results of the study by considering each of the four instruments selected for the current study.

5.5 Results and discussion

5.5.1 Instrument 1: Using a generic academic literacy test

Students wrote the 2015 version of the TALL in February as a pre-test, and the AGLE111 examination in June as a post-test. The TALL and the AGLE111 examination are not the same test. In this course's context, it was not practical to write the TALL as an examination as it does not sufficiently cover the abilities focused on in the AGLE111 course. Although the course (and its examination) is based on the TALL's theoretical construct, it does contain some elements not covered by the TALL, for example referencing abilities. A statistical item analysis could not be conducted to determine whether the TALL and the AGLE111 examination were statistically equivalent, since both tests were not written by the same population at the same point in time (Davies *et al.*, 1999:199). However, according to Steyn (Forthcoming), tests could be considered equivalent when the same test construct is measured, when degrees of reliability are consistent, and when tests are

administered under similar conditions (Steyn, 2012; Van Dyk, Van Rensburg & Marais, 2011b). As is shown below, if these guidelines are taken into consideration, one can accept that there is a large degree of construct equivalence between the TALL and the AGLE111 examination.

Both assessments include the following sub-sections: understanding texts, vocabulary, grammar and text relations, scrambled text, and graphic and visual information. However, the first test has a 5-mark section called “text types” that is not included in the AGLE111 examination. Questions were analysed based on functions/abilities addressed by the questions in each test. The various headings are described more fully below:

- **Classifying / categorising:** Classifying and categorising information
- **Grammar:** Using parts of speech and concord appropriately. Understanding sentence structure.
- **Inferencing:** Understanding and making inferences from information
- **Main ideas:** Identifying topics and main ideas; distinguishing between essential or non-essential information
- **Metaphor:** Using language metaphorically
- **Quantities:** Working with quantities / proportions / percentages without a calculator
- **Text genres:** Text types and text genres
- **Text relations:** Text relations and cohesion
- **Vocabulary:** Ability to make use of academic vocabulary in context

When questions are categorised based on the functions addressed by the questions in each test, the following is seen (see Figure 5.1). Both tests contain a similar percentage of questions (based on function type) that addresses vocabulary. There are fairly large discrepancies in the percentages of questions that address the remaining functions.

If, however, questions addressing the understanding of texts (namely those that address “Classifying / categorising”, “Inferencing”, “Main ideas”, “Metaphor” and “Quantities”) are grouped together, this discrepancy is reduced (see Figure 5.2). It could thus be said that the two tests address the functional categories of “Understanding texts” and “Vocabulary” to a similar extent. These two categories also constitute a large part of each of the respective tests.

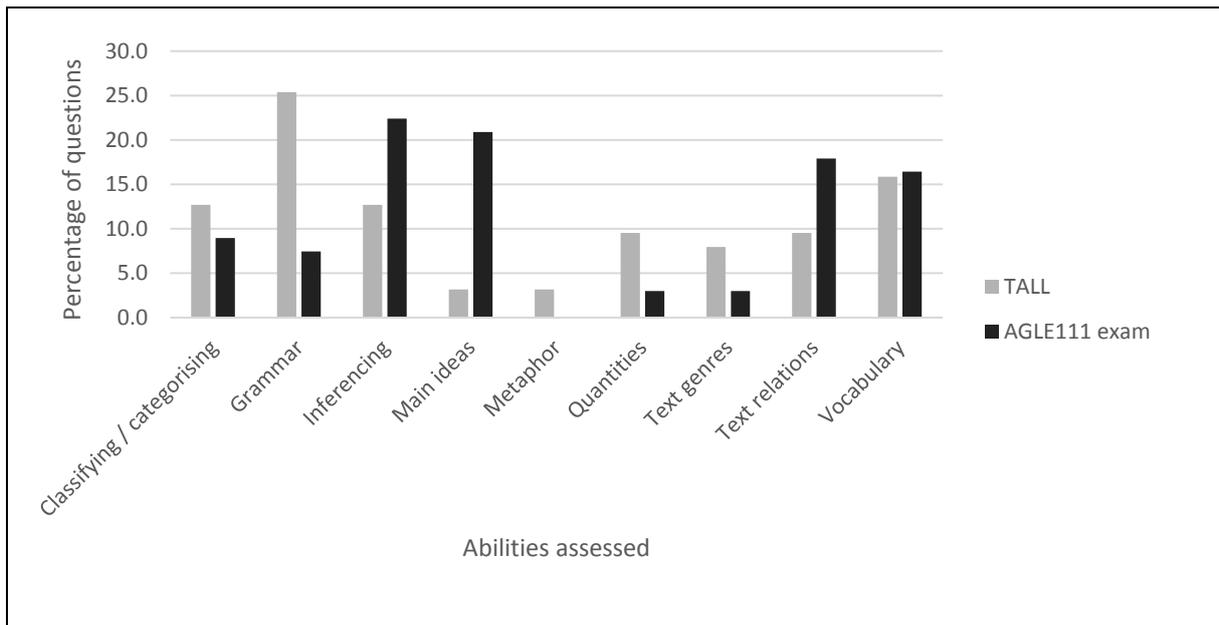


Figure 5.1: Questions categorised based on abilities assessed

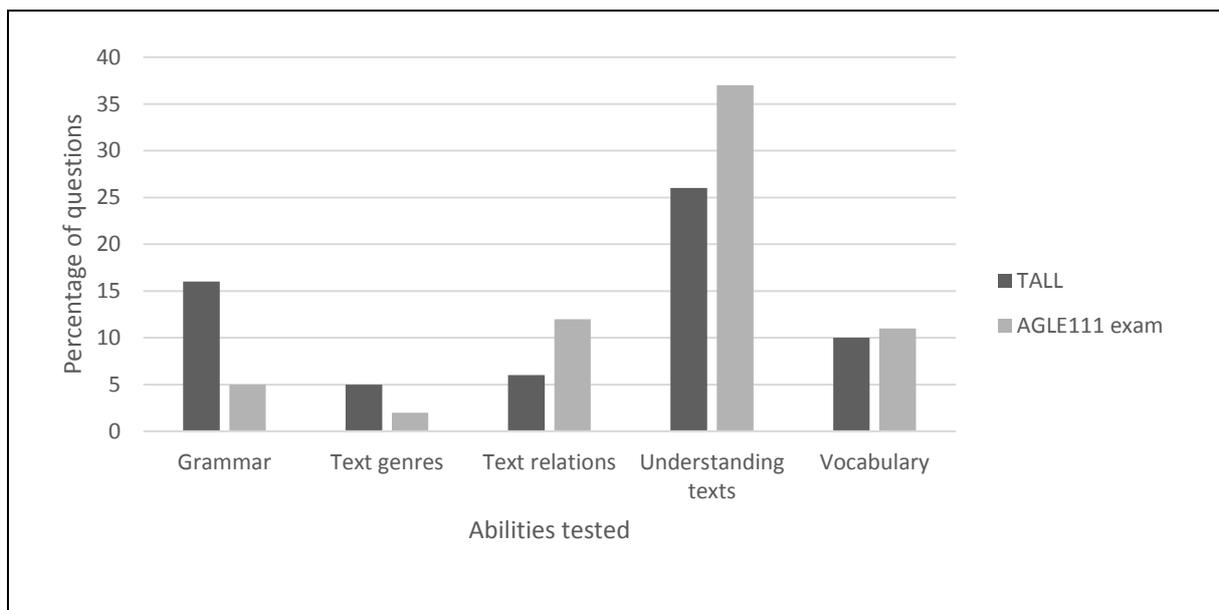


Figure 5.2: Questions categorised based on abilities assessed – re-categorised

When the reliability of the two tests is examined, the following is found. The TALL boasts a high internal reliability measure (using Cronbach’s Alpha) of 0.82³³, while the AGLE examination has a fair internal reliability rate of 0.73. Considering that both tests use a very similar theoretical construct, and that both have more than acceptable reliability measures (of more than 0.5; see Gliem & Gliem, 2003:87), an improvement between pre- and post-test can be considered valid. However, most of the sub-sections in the two tests cannot be

³³ All TALL sub-sections have fair reliability rates of between 0.47 and 0.88 (using Cronbach’s Alpha) when outlier questions are removed.

similarly compared due to low internal reliability rates in three of the five AGLE sections, namely “Grammar and text relations”, “Scrambled text”, and “Graphic and visual information”, all of which have Cronbach Alpha scores of between 0.23 and 0.44. One reason for the low reliability rates for the “Grammar and text relations” as well as “Graphic and visual information” sections might be the order in which these were written in the two tests. In the TALL, the “Grammar and text relations” section appeared at the end of the test, whereas the “Graphic and visual information” section appeared at the end of the AGLE111 examination. It seems as though several students were unable to complete the tests in the required time, and it can be assumed that several others were rushing to complete each test, which might affect the results in these two sections. Another possible reason for low reliability scores might be the number of questions addressed in each section. “Scrambled text”, for example, consisted of only five questions – such a low number of questions in a section often has a low reliability coefficient (Wells & Wollack, 2003:5). Two sections which do, however, show valid internal reliability rates are “Understanding texts” and “Vocabulary”, with Cronbach Alpha scores of 0.59 and 0.54 respectively. One question in the “Vocabulary section” proved to be particularly problematic – if this question is removed, the Cronbach Alpha score improves from 0.54 to 0.60.

Both the TALL and the AGLE111 examinations were administered under similar conditions. Both are in a multiple-choice question format. Furthermore, both are written under examination conditions.

5.5.1.1 Discussion

Students showed a slight improvement between writing the pre- and the post-test (see Table 5.1). The mean for the TALL was 46.16%, while the mean for the AGLE111 examination was 50.14, thus indicating an improvement of approximately 4% (Cohen’s $d = 0.27$; $p = < .001$). The medium effect size of 0.27 might be due to reliability concerns within the individual sections of the AGLE111 examination. Alternatively, it could be that the duration of the semester-long course was too short for a more practically significant improvement. When the improvement of the various sections that were present in both tests is considered, it would seem that students showed the greatest improvement in the questions relating to “Grammar and text relations”, “Understanding texts” as well as “Vocabulary”. However, due to the low internal reliability rate, the data obtained from

“Grammar and text relations” will not be considered here. Nor will the data from the “Scrambled text” and “Graphic and visual information” sections be considered, for the same reason. Both of these sections showed a sharp decrease in marks (of approximately 10%) between the pre- and the post-test, which could be explained by the low reliability rates of these sections; thus, these sections seem not to have reliably measured the constructs they were meant to measure.

Of those sections that had an acceptable internal reliability rate, the greatest improvement could be seen in the “Understanding texts” section, in which students improved by 9.37%. The “Vocabulary” section also showed a large improvement of 8.46%. Paired t-tests were done to determine the significance of these improvements³⁴. Though both of these sections were statistically significant ($p < .001$), the effect sizes were small in both cases (0.21 and 0.37 respectively).

Table 5.1: Improvement between the pre-intervention and post-intervention academic literacy tests (1st semester)

	Mean TALL (%)	Mean AGLE111 examination (%)	Improvement (%)	Standard deviation (pre-test; post-test)	Sig. (2 tailed) (p value)	Effect size *
TOTAL	46.16	50.14	3.98	14.95; 11.96	<.001	0.27
Understanding texts	45.35	54.72	9.37	20.91; 14.72	<.001	0.21
Vocabulary	43.99	52.45	8.46	17.61; 20.15	<.001	0.19
Grammar and text relations	35.04	47.95	12.91	27.56; 12.65	<.001	0.37
Scrambled text	67.86	49.13	-18.73	33.66; 26.28	<.001	-0.28
Graphic and visual information	51.52	43.38	-8.14	21.79; 15.02	<.001	-0.16
* Effect sizes: <.10: trivial; .10 - .30: small to medium; .30 - .50: medium to large; >.50: large to very large						

A clear limitation to this instrument is that the same test was not used as both pre- and post-test, and that the AGLE111 examination had not been tested as valid and reliable by piloting it before students wrote it. However, in most contexts it would be impossible to do

³⁴ Paired t-tests were performed to take into account the dependency of measurements on the same person. These tests are the equivalent of the mixed design ANOVA in cases where there are only two repeated measurements per person.

so for an examination that needs to be created anew each year. The validity of a test is usually only determined after the test has been written, using multiple sources of evidence. This is not practically possible in the case of most examinations; nor is piloting such an examination beforehand. In this context, using the TALL (which had been proven to be valid and reliable) as both pre- and post-test was not viable (as it did not cover all of the outcomes of the module), but future studies should aim to use existing valid and reliable tests as both pre- and post-tests. Where this is not possible, the researcher must keep in mind these limitations to using academic literacy tests in a pre-test / post-test scenario.

5.5.2 Instrument 2: Using a generic extended writing assignment (assessed by means of a rubric)

Students completed a writing assignment at the beginning of the semester, and one at the end of the semester. These assignments were similar in the following aspects, as recommended in Section 4.4.1.3. Both were approximately two pages³⁵ in length. Students were given a single topic as a pre-test topic, and a single topic as a post-test topic (Shaw & Weir, 2007:247; Heaton, 1988:138). Both pre- and post-assignment topics were of a generic nature, falling within the genre of the argumentative essay (thus, students had to argue a specific point), and were clearly defined in terms of the scope and content that had to be covered (see Weigle, 2002: 53, 93). Students were made aware of the audience for which they were writing for both of these assignments (see Weigle, 2002: 46). Furthermore, students' educational background, age and gender were taken into account in the choice of these topics (see Weigle, 2002: 46). Both topics required students to conduct some research, and called for similar background knowledge (see Hughes, 2003: 90): the first topic concerned the reasons for teenage smoking, and the second topic concerned gender equality. Thus, students would be able to respond to both topics using the level of background knowledge that a first-year student is likely to have, in addition to some literature to support their arguments. Students had time to complete both writing assignments at home, giving them enough time to go through necessary drafting processes and to conduct necessary research (see Weigle, 2002:52; Heaton, 1988:138).

35 Or roughly between 550 and 700 words – if the actual student assignments are analysed, the pre-assignment had an average word count of 631, and the post-assignment an average of 640 words.

The main difference between the two assignments was that some structure was provided in the first assignment. An introductory paragraph as well as four further introductory phrases were provided in this assignment, whilst no similar structure was provided in the post-assignment. This scaffolding provided in the first assignment was not considered necessary in the second as students had been taught to provide this structure in the AGL111 course they had just completed.

Both sets of assignments were assessed by two marking assistants. A marking session was held at the beginning of the process to ensure that marking assistants marked at the same level. Furthermore, the first five assignments marked by each marking assistant were sent back to the researcher and returned with detailed feedback. Thereafter, the researcher moderated a sample of all marked assignments to ensure that the marking level stayed consistent throughout the marking process. To further limit inconsistencies in marking, the same rubric (based on Carstens & Fletcher, 2009b:59–60) was used for both pre- and post-assignments (see Weigle, 2002:22; Heaton, 1988:148).

The original rubric was adapted after asking various academic literacy specialists to comment on it, and to recommend additional areas that were not yet covered by the rubric. The adapted rubric can be found in Appendix A. For the purposes of the current study, some sub-categories from that rubric were left out as they were not relevant to this marking assignment³⁶. These headings were “Introduction” (as an introduction was provided in the pre-assignment, and no comparison could be drawn), “Thesis statement” (for the same reason), “Technical vocabulary” (as this was a generic topic), and “Integration of visual data” (as no visual data were required by either of these topics). A decision was taken to approach “Paraphrasing information from source texts” differently – rather than giving students a mark out of seven for this section, all student assignments were analysed by means of Turnitin (LLC, 2015) to obtain a percentage that indicated how much plagiarism (if any) was committed. All texts were cleaned for this purpose by deleting original instructions (including the provided introduction and phrases of the pre-assignment), direct quotations as well as reference lists. Only five or more words in a row were considered plagiarism for the purposes of this assessment. This guideline was decided upon in consultation with the University’s Turnitin liaison as a phrase length of five words or more

³⁶ As argued in Section 4.2, instruments will likely need to be adapted for each unique context, as was the case for the current context.

was deemed suitable to identify actual plagiarism rather than common word clusters used by students.

The sections of the rubric have a high internal reliability. The various sections have the following Cronbach's Alpha scores: "Structure and development" = 0.91; "Academic writing style" = 0.86; "Editing" = 0.66; and "Use of source material" = 0.85.

With the exception of the "Paraphrasing information from source texts" section (which is indicated as a plagiarism percentage), students received a score out of seven. The following qualifiers were given for each of the seven options: 7 = Excellent; 6 = Very good; 5 = Good 4 = Average; 3 = Below average 2 = Poor 1 = Very poor. A sample of 139 students (out of an original 624) was assessed using this instrument.

5.5.2.1 Discussion

As can be seen in Table 5.2, students improved in all areas covered by the marking rubric. The sub-categories that showed the greatest improvement were those related to the use of source material, namely "Referencing technique", "Relevance of source data", "Appropriately citing quotations" and "Integration of source data with text". It is however notable that more plagiarism was found in the post-test than in the pre-test (11.9% and 9% respectively, though this difference was not found to be statistically significant, with a *p* value of .140).

A paired t-test was done to determine whether there were practically significant improvements in the various categories and sub-categories on which students were assessed. The improvement in all of the categories and sub-categories (with the exception of "Paraphrasing information from source texts") was statistically significant (thus, *p* < .050). If Cohen's (1988) guidelines (of considering an effect size of 0.2 small, an effect size of 0.5 medium and an effect size of 0.8 large) are followed, then all of the categories and sub-categories were close to having either a medium or large effect size, with the greatest effect sizes being evident in the "Use of source material", "Academic writing style" and "Editing" categories.

Table 5.2: Improvement between the pre- and post-intervention writing assignments (assessed by means of rubric) (1st semester)

	Mean pre-assignment / 7	Mean post-assignment / 7	Improvement / 7	Standard deviation (pre-assignment; post-assignment)	Sig. (2 tailed) (<i>p</i> value)	Effect size *
STRUCTURE AND DEVELOPMENT	4.33	4.79	0.46	0.88; 0.75	<.001	0.53
1. Development of main argument	4.4	4.8	0.4	0.99; 0.98	.001	0.39
2. Paragraph development (topic sentences, main ideas, supporting information)	4.6	5	0.4	1.00; 0.93	<.001	0.41
3. Relevance of content to topic	4.7	5.3	0.6	1.05; 0.86	<.001	0.64
4. Linking devices (structuring thought with discourse markers, pronouns etc)	3.9	4.4	0.5	1.01; 0.91	<.001	0.43
5. Conclusion	4.1	4.6	0.5	0.99; 1.09	<.001	0.41
ACADEMIC WRITING STYLE	3.95	4.75	0.79	0.88; 0.70	<.001	0.90
6. Syntax: phrase and clause structure	3.6	4.2	0.6	1.03; 0.97	<.001	0.60
7. Academic vocabulary	3.9	4.8	0.9	1.08; 0.93	<.001	0.81
8. Style (formality; rhetorical mode)	4.4	5.2	0.8	1.03; 0.78	<.001	0.86
EDITING	3.99	4.69	0.70	0.82; 0.76	<.001	0.85
9. Spelling, capitalisation and punctuation	3.8	4.3	0.5	1.17; 1.02	<.001	0.38
10. Concord and tense	4.1	4.9	0.8	0.94; 0.95	<.001	0.95
11. Layout and format	4.1	4.9	0.8	0.98; 1.10	<.001	0.79
USE OF SOURCE MATERIAL	2.91	4.53	1.62	0.87; 0.94	<.001	1.86
12. Referencing technique	2.5	4.1	1.6	1.08; 1.16	<.001	1.45
13. Appropriately citing quotations	2.6	4.5	1.9	0.92; 1.22	<.001	2.07
14. Relevance of source data	4.0	5.2	1.2	1.20; 0.99	<.001	1.00
15. Integration of source data with text (synthesising)	2.5	4.4	1.9	1.00; 1.25	<.001	1.82
PLAGIARISM						
16. Paraphrasing information from source texts (expressed as a plagiarism percentage)	9%	11.9%	-(2.4%)	17.12; 16.58	.140	0.14
* Effect sizes: <.10: trivial; .10 - .30: small to medium; .30 - .50: medium to large; >.50: large to very large						

Students are usually not required to reference source material at secondary school level. It therefore makes sense that such a great improvement would be seen in the “Use of source material” section between the pre-assignment (at which time students had received little or no instruction on this topic) and post-assignment (at which time students had received instruction on this topic).

In addition to the use of source material in texts, the academic literacy course in question also seems to have impacted all other categories on this rubric, in particular “Editing” and “Academic Writing Style” (both of which showed large improvements as well as large effect sizes of 0.85 and 0.90 respectively) – all of which form part of the English syllabus at secondary school level. Thus, the academic literacy course does seem to be fulfilling its purpose in addressing the poor academic literacy levels with which students enter universities.

The section with the lowest level of improvement (of 0.46 marks, or approximately 7%), namely “Structure and Development”, still shows a clear statistically significant improvement with a medium effect size. It is possible that the abilities addressed in this section (namely “Development of main argument”, “Paragraph development”, “Relevance of content to topic”, “Linking devices” and “Conclusion”) are complex and thus difficult to acquire – an even greater improvement might be evident after another semester of academic literacy study (see Section 6.5.2.1).

5.5.3 Instrument 3: Quantitatively assessing an extended writing assignment

Students’ writing assignments were further analysed by means of measures suggested by Storch and Tapper (2009) – this analysis is summarised in Appendix C. Storch and Tapper (2009) use the following measures: fluency, accuracy, use of academic vocabulary, as well as text structure and rhetorical quality. This study chose to only use two of these measures, namely accuracy and use of academic vocabulary (the reasons for this are indicated below). Due to the labour intensity of this instrument, it was not feasible to mark the same sample of 139 students that was used in the previous section. A smaller random sample of 50 pre- and 50 post-assignments was thus used. The sections below indicate why two measures, namely fluency and text structure, and rhetorical quality, were left out, and how the remaining two measures, namely accuracy and academic vocabulary, were measured.

Fluency

Storch and Tapper (2009:211) measured fluency by looking at the number of words written as well as the words per T-unit. For the purposes of this study, considering students' total number of words was not thought useful as students were given guidelines regarding length, and as a result their pre- and post-assignments were of similar length (on average 537 and 568 words respectively). Furthermore, considering words per T-unit was also not considered sensible for this study. In the case of the Storch and Tapper (2009) study, the authors believed that a longer T-unit would equate to more fluent writing, which may well be the case for postgraduate students. However, in the case of first-year students, the aim is rather to encourage students to write shorter sentences (which may or may not be reflected in shorter T-units). Thus, T-unit length was not considered suitable for this study. It should be noted that even in the Storch and Tapper (2009) study, no statistically significant difference could be found between pre- and post-scores with regards to words per T-unit. It is possible that fluency is best measured using a more traditional rubric, and that a quantitative approach to measuring fluency is not necessarily effective for the purpose of most academic writing. An alternative option that was not explored for the purposes of the current chapter would be to do frequency counts of words that are typically used to enhance the fluency of a text (such as conjunctions or pronouns), to determine whether more such words are used after the conclusion of an intervention than was the case before the commencement of the intervention.

Rhetorical quality

Storch and Tapper (2009:212) assessed rhetorical quality by considering the overall structure, coherence and cohesion, and content of students' writing. A traditional marking rubric was used for this purpose. The subsections on this marking rubric were very similar to those under the "Structure and development" section used in the current study (discussed in Section 5.5.2). It was thus considered unnecessary to replicate this assessment using an instrument that was so similar to the one already used.

Accuracy

Storch and Tapper (2009:219-220) coded assignments using 18 categories of errors. For the current study, a decision was taken to merge some of these categories to ultimately use only ten categories of errors. After consulting with various colleagues, it was decided that some of the distinctions seemed unnecessary for first-year students (for example, it was considered sufficient to categorise word choice, collocation and derivational errors under the same heading, namely “word choice”). In addition, decreasing the number of categories reduced the complexity of the marking, and facilitated increased conceptual clarity, as many of the previous categories were difficult to distinguish from each other in practice. Figure 5.3 indicates how categories were merged.

Current study		Storch and Tapper (2009) study
Article / determiner errors		Determiners
Plural errors		Articles
Word choice errors		Plural
Agreement errors		Word choice
Tense errors		Collocation
Preposition errors (note: unlike in the Storch and Tapper [2009] study, incorrect prepositions were coded as preposition errors, and not as word choice errors)		Derivational (word form) errors
Syntax errors		Agreement (noun or pronoun with verb)
Linking errors		Agreement of verb with subject.
Possessive errors		Tense and verb form
Punctuation errors (note: unlike in the Storch and Tapper [2009] study, all punctuation errors were counted, and not just those that obstructed meaning)		Prepositions
	Word order	
	Absence of major constituent	
	Absence of minor constituent	
	Passive form	
	Errors in linking ideas	
	Possessive	
	Punctuation	
	Capitalisation	

Figure 5.3: Categories used to calculate accuracy in written assignments

To compensate for the slight difference in structure between the pre- and post-assignments (as discussed earlier), the paragraph and phrases that were provided in the pre-assignment were deleted from the assignments assessed in this sample. Furthermore, students’

plagiarism counts were subtracted from the total word count, as sections plagiarised from the Internet and academic articles were unlikely to contain any mistakes. The word count of assignments was thus adapted to only reflect the words that the students wrote themselves. Raw error counts were converted to reflect an error count per 100 words so as to compensate for the different lengths of students' assignments.

Due to the high frequency of mistakes in the pre- and post-assignments of the current study (almost double the mistakes that were recorded in the Storch and Tapper [2009] study), it was decided not to calculate a ratio of error free T-units per total T-units or error free clauses per total clauses as was done in the Storch and Tapper [2009] study. The main concern with these measurements is that these ratio scores cannot distinguish between T-units and clauses which have only one error, and those with more than one error (see, for example, Storch & Tapper, 2009:211; Bardovi-Harlig & Bofman, 1989:22). Rather, a measurement of errors per 100 words (the third measurement used by Storch and Tapper [2009]) was used as a more useful representation of accuracy.

Academic vocabulary

Storch and Tapper (2009:212) assessed students' use of academic vocabulary by using Coxhead's (2000) Academic Word List (AWL). They counted a) the number of AWL types, b) the number of AWL tokens and c) the numbers of tokens as a percentage of words written. The current study used the same approach, but also added the number of AWL word families. Further, in addition to the number of tokens, the number of AWL families and types were also calculated as a count per 100 words (thus, as a percentage of words written) to compensate for different essay lengths.

A word family is a collection of variations on a word. For example, the words "hypothesis", "hypotheses", and "hypothesise" would belong to one word family. The AWL contains 570 such word families – this accounts for roughly ten percent of all words in academic texts (while these 570 families only account for 1.4% of words in fiction) (Coxhead, 2000: 213). A word type count indicates how many unique words from the AWL were used. Thus, if the student used the word "hypothesis" twenty times, this would only count as one word type, whereas it would count as twenty AWL tokens.

Cobb's (2015) Vocabprofile, which was adapted from Heatley, Nation and Coxhead (2002), was used for all of these word frequency counts. As was the case in Section 5.5.2 and the "Accuracy" sub-section of this section, all assignments were cleaned by deleting any provided paragraphs, phrases and instructions. Plagiarism was compensated for by deleting any plagiarised sections. The larger sample of 139 students was again used for this section, due to the relative ease of performing frequency counts with the help Cobb's (2015) Online Vocabprofile.

5.5.3.1 Discussion

Accuracy

Although all error categories showed a reduction in mistakes between the pre- and the post-assignment, very few showed a statistically significant reduction in errors with *p* values below .050 (see Table 5.3). There was a substantial reduction in total errors per 100

Table 5.3: Reduction of language mistakes between pre- and post-assignments (1st semester)

	Mean pre-assignment (errors per 100 words)	Mean post-assignment (errors per 100 words)	Reduction in mistakes between pre- and post-assignments	Standard deviation (pre-assignment; post-assignment)	Sig. (2 tailed) (<i>p</i> value)	Effect size *
Total errors	16.93	8.48	8.45	36.32; 3.47	.097	-0.23
Article / determiner errors	0.82	0.33	0.49	1.71; 0.34	.039	-0.28
Plural errors	0.61	0.28	0.34	0.87; 0.35	.008	-0.39
Word choice errors	1.7	1.48	0.21	1.58; 0.77	.334	-0.13
Agreement errors	0.87	0.38	0.49	1.89; 0.35	.085	-0.26
Tense errors	0.32	0.18	0.14	0.46; 0.21	.086	-0.30
Preposition errors	0.72	0.4	0.32	0.76; 0.35	.010	-0.43
Syntax errors	2.54	1.86	0.68	2.47; 0.86	.036	-0.28
Linking errors	0.32	0.26	0.06	0.45; 0.26	.399	-0.13
Possessive errors	0.29	0.11	0.17	0.79; 0.19	.799	-0.22
Punctuation	3.27	2.03	1.24	6.24; 1.23	.142	-0.20
* Effect sizes: <.10: trivial; .10 - .30: small to medium; .30 - .50: medium to large; >.50: large to very large						

words (this number dropped by almost half), but the reduction is not significant due to the large standard deviation (36.32) in the pre-test scores. The categories that did show a statistically significant reduction in errors are “Article / determiner errors”, “Plural errors”, “Preposition errors” and “Syntax errors”. All of these categories also show medium effect sizes. This information corresponds to the related categories discussed in Section 5.5.2.1, namely “Spelling, capitalisation and punctuation”, “concord and tense” as well as “syntax: phrase and clause structure” that also showed only medium effect sizes and moderate improvement. It is interesting to note that a large improvement was seen in a related category on the TALL/AGLE pre- and post-tests, in contrast to the average improvement observed using the accuracy score in the current section. The “Grammar and text relations” section on these tests showed a 12.91% (Cohen’s $d = 0.37$; $p = < .001$) improvement.

Academic vocabulary

Students used almost double as many words from the AWL in the post-assignment when compared to the pre-assignment (see “Tokens per 100 words”). As is evident from Table 5.4, this improvement is reflected in the increased number of word families per 100 words and word types per 100 words, all of which increased significantly with medium to large effect sizes. The most practically significant indicator of increased academic vocabulary usage is the tokens per 100 words. The most useful indicators might, however, be AWL families and AWL types, as these two indicators demonstrate how many new academic words each student used in the post-assignment in contrast to the pre-assignment. A

Table 5.4: Increase in academic vocabulary usage between pre- and post-assignments (1st semester)

	Mean pre-assignment	Mean post-assignment	Improvement between pre- and post-assignments	Standard deviation (pre-assignment; post-assignment)	Sig. (2 tailed) (p value)	Effect size *
Academic Word List Families per 100 words	5.3	7.3	1.9	6.10; 4.65	.008	0.26
Academic Word List Types per 100 words	5.9	7.6	1.8	7.10; 4.57	.008	0.24
Academic Word List Tokens per 100 words	8.3	13.3	5.0	10.44; 6.70	<.001	0.45
* Effect sizes: <.10: trivial; .10 - .30: small to medium; .30 - .50: medium to large; >.50: large to very large						

limitation of this instrument is that it merely measures the usage frequency of words on the AWL, and does not measure whether students used these words correctly. This limitation can be (and was) addressed by considering feedback received by means of other instruments (see Section 5.5.2)

The results obtained from counting the frequency of AWL words in students' writing correspond to those obtained from the rubric used in Section 5.5.2. Results from the rubric indicate that there was a practically significant improvement of approximately 13% (Cohen's $d = 0.81$; $p = < .001$) in students' use of academic vocabulary. Another perspective on students' vocabulary use can be found by looking at the accuracy scores obtained from using Storch and Tapper's (2009) error categories. One of these categories is "word choice errors". This category does not consider increased academic vocabulary usage, but rather looks at whether students use (all) words correctly (i.e. accuracy). There was a small decrease in students' word choice errors, though this was not statistically (neither theoretically nor practically) significant. Even though students might have used more academic words in the post-assignment than in the pre-assignment, the level of accuracy of their vocabulary usage therefore seems to remain largely unchanged.

5.5.4 Instrument 4: Correlating academic literacy achievements with other variables

A final instrument that was used for the current study was correlating students' academic literacy achievements with other variables. Firstly, students' AGLE111 examination results were correlated with their final AGLE111 marks. Secondly, students' AGLE111 examination marks, their AGLE111 final marks as well as their TALL marks were correlated with their results in their other subjects.

5.5.4.1 Discussion

Non-parametric correlations were drawn using Spearman's rho (see Table 5.5)³⁷. There was a significant and strong correlation ($r_s = 0.65$; $p = < .001$) between students' AGLE111 examination marks and their final AGLE111 marks, indicating that the module's

³⁷ Spearman rank order correlations were performed to indicate whether there is a relationship between students' various academic literacy test results as well as their results in their other modules. The purpose was not to make a prediction of any variable and therefore no regression analyses were performed.

continuous assessment was well in line with the examination. There was also a significant and strong correlation ($r_s = 0.63$; $p = < .001$) between the TALL and the AGLE111 examination, further strengthening the assumption that these tests (in their entirety, though not necessarily all of their separate sections) can be considered equivalent.

Table 5.5: Non-parametric correlations using Spearman’s rho (1st semester)

		AGLE111 exam	TALL	AGLE module	Marks in all other subjects
AGLE111 exam	Correlation Coefficient (r_s)	1.00	0.63	0.65	0.24
	Sig. (2-tailed)		<.001	<.001	.003
	N	173	169	161	161
TALL	Correlation Coefficient (r_s)	0.63*	1.00	0.51	0.21
	Sig. (2-tailed)	<.001		.000	.008
	N	169	169	157	157
AGLE module	Correlation Coefficient (r_s)	0.65	0.51	1.00	0.35
	Sig. (2-tailed)	<.001	<.001		<.001
	N	161	157	161	161
Marks in all other subjects	Correlation Coefficient (r_s)	0.24	0.21	0.35	1.00
	Sig. (2-tailed)	.003	.008	<.001	
	N	161	157	161	161

Unlike previous studies (e.g. Van Dyk, 2015:180; Van Rooy & Coetzee-Van Rooy, 2015:10-11; Mhlongo, 2014:80-82) which found strong correlations between the results of academic literacy modules and students’ results in their other subjects, the current study found only a moderate correlation (of 0.35) between the final AGLE111 mark and students’ marks in their other subjects. Thus, whereas one could infer from the studies mentioned above that better performance in the academic literacy interventions under discussion in these studies might lead to improved performance in students’ other subjects, this study cannot safely make the same assumption. It must be kept in mind, however, that students had only completed one academic literacy semester at this point, in contrast to the studies mentioned above. A stronger correlation might be found after students have completed the second AGLE module (AGLE121) (see Section 6.5.5.1 for a discussion of this research). A further fact which could influence this correlation is that the current study only focuses on those students who were identified as at-risk at the beginning of the year. The weak correlation (of 0.21) between the TALL and students’ marks in other subjects is to be expected, as one would assume that if the academic literacy intervention had any impact,

this impact would weaken any stronger correlation that would have existed had no intervention occurred.

5.6 Summary of findings

Students seem to have improved across a wide spectrum of academic literacy abilities after taking a one-semester long (from February to May) academic literacy course. Some notable improvements between the beginning and the end of the semester are students' ability to use source material in written assignments, an increase in students' academic vocabulary range and usage, students' ability to edit their work (including a reduction in language errors as well as an improvement in the layout and format of their assignments) and their ability to structure and develop their writing.

From the data obtained in this study, it may be inferred that the course has successfully addressed the following outcomes (see Section 5.3 for a complete list of the course's outcomes):

- Identifying word meaning from context (see Section 5.5.1.1, "Vocabulary" and "Understanding texts"³⁸; Section 5.5.2.1, "Academic vocabulary"; and section 5.5.3.1, "Word choice errors" in the "Accuracy" sub-section as well as word frequency counts in the "Vocabulary" sub-section);
- Including references in a text (see Section 5.5.2.1, "Use of source material");
- Understanding academic genres and identifying and finding reliable academic sources (see Section 5.5.2.1, "Relevance of source data");
- Identifying the qualities of, and being able to write good introductions and conclusions (see the "Conclusion" category in Section 5.5.2.1);
- Creating a table of contents, and using it to plan and structure text (see Section 5.5.2.1, "Layout and format");
- Writing paragraphs with clear topic sentences, one main idea and applicable support (see Section 5.5.2.1, "Paragraph development"); and

³⁸ The low effect sizes of these sections should be kept in mind though.

- Writing correct sentences (see Section 5.5.1.1, “Grammar and text relations”³⁹; Section 5.5.2.1, “Syntax”, “Spelling, capitalisation and punctuation”; and “Concord and tense” as well as the accuracy section of Section 5.5.3.1).

Some outcomes did not seem to have improved between the beginning and the end of the semester. One such outcome was “paraphrasing text”. It should however be kept in mind that the only measure of this in the instruments used for this study was students’ plagiarism counts, which increased slightly (though not statistically significantly) between the beginning and the end of the semester. It is possible that the way in which they paraphrased information had indeed improved – students’ improved accuracy counts (Section 5.5.3) as well as improved “Syntax”, “Spelling, capitalisation and punctuation” as well as “Concord and tense” scores might be an indication that paraphrased sections were dealt with more effectively. Another outcome where students seemed to have performed more poorly in the post-test than in the pre-test is “Identify reasons for using graphic information, analyse graphics and discuss graphics appropriately”, although this might be explained by the low internal reliability of this section in the AGLE111 examination. More information could have been collected had this outcome also been assessed in the writing assignment.

Some outcomes could not be assessed by means of the instruments that were used. In some cases, that is because appropriate instruments were not available; for example, no appropriate instrument could be found with which to assess students’ study strategies or time management abilities. In other cases, it is because it is not necessarily possible to directly measure some outcomes. Outcomes where this might have been the case include “Using skimming and scanning to obtain information from texts”, “Identifying action words and content words in examination questions and assignments, and planning well-structured responses to examination questions”, and “Being aware of the structure of a seminar, being able to ask effective questions, and being able to answer questions effectively”. Several outcomes which were not directly measured were, however, implicitly measured by measuring other outcomes; for example, students’ ability to read effectively was indirectly measured in that they had to conduct research and understand this research so as to report on it in their written work. Similarly, students would presumably be able to

³⁹ The low internal reliability of this section in the AGLE111 examination should once again be considered.

identify the reasons for using the passive voice if this ability is used appropriately in their writing.

In some cases, outcomes were not assessed by means of the instruments that were used, but could easily have been assessed had the instruments been adapted slightly. For example, the writing assignment could have required students to refer “correctly to different parts of graphs and tables”. It is therefore advisable to first consider all course outcomes before setting up assignments to be used as assessment instruments (see also the comment above about the “Identify reasons for using graphic information, analyse graphics and discuss graphics appropriately” outcome)⁴⁰.

Strong correlations could be found between the TALL and the AGLE111 examination, as well as between the AGLE111 examination and the total AGLE111 module, indicating that the pre- and post-test, as well as the examination and the continuous assessment of the AGLE111 module are all well aligned. However, as only a moderate correlation could be found between the AGLE111 module mark and students’ other marks, no certain claims can be made indicating that the AGLE111 module on its own is likely to lead to improved marks in students’ other subjects. Such an effect, however, might be more apparent after the second semester, when students have been able to implement all the abilities acquired throughout the AGLE111 as well as the AGLE121 modules.

5.7 Conclusion

As can be seen from the previous section, as far as the assessment of this particular course is concerned, the course seems to effectively address several academic literacy abilities, but could potentially be improved by focusing more on abilities that showed no, or only a moderate, improvement, for example students’ visual literacy abilities, their ability to paraphrase information, and even their ability to structure assignments appropriately (the moderate improvement in this ability might have been influenced by the scaffolded approach that was followed in the pre-assignment). It should be kept in mind though that

⁴⁰ Considering that the sections on “Graphic and visual information” in the TALL and AGLE111 examination were not statistically comparable, it would have been particularly helpful to measure outcomes regarding visual information by means of alternative instruments.

these abilities do receive more attention in the second semester module (AGLE121), and a clear improvement might only be visible after the completion of that module.

Despite these encouraging findings, it is important to remember the limitations of a quasi-experimental design as employed in this study. Because no control group was used, it is impossible to attribute students' improvement to the academic literacy intervention without any doubt. However, the use of triangulated data sources does strengthen the assumption that the intervention did influence students' academic literacy levels.

As far as the evaluation of academic literacy interventions in general is concerned, it would seem that using a variety of instruments is indeed worthwhile, as richer information is consequently provided, thus assisting the researcher in correctly identifying the intervention's strengths and weaknesses. Using a variety of instruments made it possible to either corroborate data (for example considering the improvement in students' language usage by using instruments 2 and 3) or to provide a richer context for deductions made from data. An example of this is when one considers students' vocabulary usage. While it seems clear that students use a wider range of academic vocabulary, and use it more frequently, in the post-assignment than in the pre-assignment (taking into consideration instrument 2 and the vocabulary section of instrument 3), they still make a similar number of vocabulary errors after a semester-long academic literacy course.

Chapter 6 will assess the increase and possible transfer of academic literacy abilities at the end of a year of academic literacy study. Pre- and post-tests and assignments will once again be used, in addition to perceptual data from students. Once this has been done, it is hoped that a complete picture of the impact of a year of academic literacy study at a South African university will emerge.

CHAPTER 6

AN “ENLIGHTENING COURSE THAT EMPOWERS FIRST YEARS”?: A HOLISTIC ASSESSMENT OF THE IMPACT OF A FIRST-YEAR ACADEMIC LITERACY COURSE (ARTICLE 4)

6.1 Prelude to Chapter 6

Chapter 6 reports on a continuation of the impact assessment discussed in Chapter 5. Whereas Chapter 5 only considered the improvement in students’ academic literacy levels in the first semester of the academic literacy course, Chapter 6 examines the improvement in students’ academic literacy levels over the full two semesters of the course, and also examines to which extent the abilities addressed in the course are needed in, and transferred to, students’ content subjects (see Chapter 2.2.1 for the definition of impact used for the current study).

For this part of the study, the following instruments were selected from the evaluation design proposed in Chapter 4:

- Generic academic literacy test
- Generic extended writing assignment
- Subject-specific extended writing assignment
- Quantitatively assessing writing assignment
- Student questionnaire
- Correlating academic literacy results with other variables

One goal in this process of verifying the evaluation design is to determine where various instruments proposed in Chapter 4 could be improved, or approached differently. While Chapters 5 and 6 touch on the challenges experienced, Chapter 7 expands on these challenges and suggests improved research instruments, or improved guidelines for using these instruments.

6.2 Introduction

Academic literacy interventions have become commonplace at South African universities to address the low academic literacy levels with which students enter higher education institutions (Cliff, 2014:322; Sebolai, 2014:52; Davies, 2010:xi). However, very little research has been conducted to determine the impact of these interventions (Sebolai, 2014:52; Butler, 2013:80; Terraschke & Wahid, 2011:174; Carstens & Fletcher, 2009a:319; Storch & Tapper, 2009:218; Holder *et al.*, 1999:20). The current chapter is a continuation of a study reported on in Chapter 5, which aims to address this research gap by evaluating the impact of an academic literacy course at a South African university, using instruments suggested in an evaluation design for academic literacy interventions, as proposed in Chapter 4. Chapter 5 reported the improvement in students' academic literacy levels in the first semester of an academic literacy course. The current chapter considers the improvement in students' academic literacy abilities over the entire two-semester period. It further reports on the importance of various abilities for students to be successful in their studies, and to which extent those abilities were addressed in the academic literacy course, by considering student questionnaires.

Several researchers agree that students' academic literacy and academic language abilities are unlikely to improve significantly without some type of intervention. Rosenthal (1996:24), for example, argues that language proficiency improves over time through practice. Eskey (1983:322) agrees that students need to be explicitly guided for their academic language usage to improve. According to De Graaff and Housen (2009:729), several studies have confirmed that "instructed learners ultimately reach higher stages of inter-language development and higher levels of proficiency than uninstructed L2 learners". Holder *et al.* (1999:27) show that "the predictive validity of [students'] literacy ratings will be maintained" unless academic literacy levels are directly addressed (cf. Farnill & Hayes, 1996:264; Thompson, 1990:101). It thus seems likely that some form of academic literacy intervention is necessary for students' academic literacy levels to improve significantly, and that a significant improvement in academic literacy levels can likely be attributed to an academic literacy intervention. If it has been shown that there was in fact a significant improvement in students' academic literacy levels between the onset and the completion of an academic literacy intervention, the assumption that this improvement can be attributed to the academic literacy course

(as indicated by the literature above) can be strengthened by asking students to which extent the academic literacy intervention addressed the academic literacy needs they have in other subjects. That was done in the current study, and is discussed in Section 6.5.4. The methodology employed in this part of the study is further discussed in Section 6.4. The next section considers the background of the current study.

6.3 Background

To determine whether students are at risk of being unsuccessful in their studies due to low academic literacy levels, the North-West University's Potchefstroom Campus requires all first-year students to write the Test of Academic Literacy Levels (TALL) or its Afrikaans equivalent, *Die Toets van Akademiese Geletterdheid* (TAG)⁴¹, before commencing their studies. At-risk students complete the first (one semester long) module of an academic literacy course. Thereafter, a second module (again of one semester) is presented and must be completed by all students (both at-risk at not at-risk students). Both of these modules are presented in students' first year of studies. In 2015, a total of 624 students registered for the first module, and 843 registered for the second module. The academic literacy course (called AGLE) is a generic course which services all of the Potchefstroom Campus' faculties (including Arts, Natural Sciences, Theology, Education Sciences, Economic and Management Sciences, Law, Engineering, and Health Sciences). Students attend two 50-minute classes per week, and each module (or semester) of 12 credits runs over approximately four months.

Chapter 5 discussed the improvement in students' academic literacy levels after having completed the first semester-long module. This was done by using the following instruments: an academic literacy test, a writing assignment assessed by means of both a rubric and quantitative measures, and correlating students results with other variables. The impact of the AGLE course in its entirety has, however, not yet been formally assessed. The remainder of this chapter aims to do precisely this by firstly discussing the design and method of the study in Section 6.4, and thereafter discussing the results of the study in Section 6.5.

⁴¹ The current study only focuses on the students who study through the medium of English, and will thus refer to only the TALL and the concomitant English academic literacy courses from this point onward.

6.4 Design and method

For the design of the current evaluation, five appropriate instruments were chosen from the evaluation design proposed in Chapter 4. These instruments are used in conjunction and in many instances serve as validity checks for each other in that data are triangulated. The instruments are:

1. a generic academic literacy test;
2. a generic extended writing assignment (assessed by means of a rubric);
3. a quantitative assessment of an extended writing assignment;
4. a student questionnaire; and
5. a correlation of academic literacy achievements with students' results in their other subjects.

For the sake of consistency, the sample which was drawn from for instruments 1, 2, 3 and 5 is the same sample that was used in Chapter 5. A limitation of this study is therefore that the results of students who completed only the second module are not examined. A sample of convenience was used for the fourth and fifth instruments. The fourth instrument's responses include responses from students who only completed the second academic literacy module. Though a sample of convenience was used for all these instruments, the sample was spread across all the classes, meaning that the effect that a specific (good or poor) lecturer had could not significantly affect the outcomes (cf. Beretta, 1992:11)

The following section reports on the results of the study by considering each of the instruments selected for the current study.

6.5 Results and discussion

6.5.1 Instrument 1: Using a generic academic literacy test

As was the case in the first half of this study reported on in Chapter 5, all students again wrote an examination that was very similar to the TALL that was written at the beginning of the year. The reliability of this examination was determined, and sections in which students seemed to have improved were identified.

The first step was to determine the reliability rates of the various tests used in this study. Using Chronbach's Alpha, the TALL (which was used as a pre-test at the beginning of the year) had a reliability rate of 0.82. The June examination (i.e. the AGLE111 examination) had a reliability rate of 0.73, and the November examination (i.e. the AGLE121 examination) had a reliability rate of 0.79.

Steyn (Forthcoming) states that when tests 1) measure the same test construct, 2) have consistent degrees of reliability, and 3) are administered under similar conditions, such tests can be considered equivalent (also see Steyn, 2012; Van Dyk *et al.*, 2011b). Based on these criteria, the TALL and both AGLE examinations could be considered equivalent.

However, it becomes more problematic to compare the various sub-sections due to low internal reliability rates in various sections of the two AGLE examinations. These low reliability rates might be due to the small number of questions in some sections (for example, one section consisted of only five questions). Another reason could be because the two examinations could not be piloted (new examinations need to be set each year, and pilot groups are not practically or logistically possible). It is also possible that the constructs used within the specific sections were not always consistent (e.g. that the visual literacy section also tested vocabulary or text comprehension). Due to low internal reliability rates in various sections across the three tests employed for this study, these sections of the tests were not analysed further, as no conclusive statistical conclusions can be made on the improvement in the various subsections. The fact that it was not possible to determine in which areas of academic literacy students had improved the most is a limitation of the current study.

6.5.1.1 Discussion

Two-tailed t-tests were used (through SPSS) to determine whether the improvement between pre-, mid- and post-tests were significant⁴². As can be seen in Table 6.1, there was a significant improvement between the TALL (pre-test) and the AGLE111 June

⁴² This was done to take into account the dependency of measurements on the same person. Paired t-tests are the equivalent of the mixed design ANOVA in cases where there are only two repeated measurements per person.

examination (mid-test), with students' marks improving from 46.16% to 50.14% ($p = <.001$). There was a further significant improvement from 50.14% to 56.36% between the AGLE111 June examination (mid-test) and the AGLE121 November examination (post-test) ($p = <.001$). Medium effect sizes can be seen between pre-test and mid-test (Cohen's $d = 0.30$) and between mid-test and post-test (Cohen's $d = 0.33$). However, a significant improvement of 7.63% with a large effect size (Cohen's $d = 0.55$, $p = <.001$) can only be seen if the entire year's intervention is taken into account (thus, considering the pre- and the post-test). This would imply that a year-long intervention is more advantageous to students' academic literacy acquisition than is the case with a semester-long intervention.

Table 6.1: Improvement between the pre- and post-intervention academic literacy tests⁴³ (2nd semester)

	Mean (%)	Improvement (%)	N	Standard deviation	Sig. (2 tailed) (p value)	Effect size*
TALL (pre-test)	46.33	4.14	292	13.945	<.001	0.30
AGLE111 examination (mid-test)	50.46			10.657		
AGLE111 examination (mid-test)	50.46	3.49	292	10.657	<.001	0.33
AGLE121 examination (post-test)	53.95			8.296		
TALL (pre-test)	46.33	7.63	292	13.945	<.001	0.55
AGLE121 examination (post-test)	53.95			8.296		

* Effect sizes: <.10: trivial; .10 - .30: small to medium; .30 - .50: medium to large; >.50: large to very large

6.5.2 Instrument 2: Using a generic extended writing assignment (assessed by means of a rubric)

To provide additional information about which academic literacy abilities had been most improved after the conclusion of the intervention, a pre- and post- written assignment was assessed by means of a rubric (see Appendix A). It proved to be a great challenge to obtain writing assignments from students at the end of the second semester. The academic literacy course itself required students to complete a group assignment at the

⁴³ Note that the figures for the TALL and AGLE111 examination differ slightly from those reported on in Chapter 5 due to different sample sizes used. The sample used for this chapter consist of students for whom TALL, AGLE111 and AGLE121 data were available.

end of the year, but not an individual assignment. This was, in part, due to large student numbers and limited resources, but also because effective group work was one of the outcomes of the course. It was, however, decided that an assignment which was the result of group work could not be used to validly determine whether there was an improvement in individuals' writing abilities. It also was not possible to add an additional writing assignment to the curriculum, as the course had already used all the credits (and concomitant study hours) that it had been allocated. Consequently, an alternative solution had to be found.

It was decided to ask students who had been in the original convenience sample, and for whom there were available pre-intervention data regarding their writing abilities (refer to Section 5.4), to submit any individual writing assignment that they had completed in any of their mainstream subjects. However, it soon became clear that very little writing is done in students' first year of study. Numerous messages (via e-mail and sms) were sent to all students; in addition, all students were called two times to ask them to submit written work. It was found that only five subjects (out of a total of 98 subjects) had required an individual written assignment from students in their first year – this included one subject that set a laboratory report as an assignment, and another that required students to write a letter. That meant that only 53 students of the 330 for whom pre-intervention writing assignment data were available had completed a written assignment in their second semester. Of these assignments, the researcher was able to collect a convenience sample of 38 – some students no longer had a copy of their assignments, and some promised to send their assignments but never did⁴⁴. The fact that students do so little individual writing in their first year, and that so few assignments could consequently be collected, is a limitation of the current study. A further limitation is that the due dates for these assignments were from August to October. If all assignments had been due at the end of the semester, a greater improvement in students' writing abilities might have been observed.

⁴⁴ Although it is possible that results might have been skewed by asking students to voluntarily submit assignments, in that stronger students might have been more likely to submit, this is unlikely to have been very problematic, since 72% (38 out of 53) of assignments were collected. After having marked the assignments, it seemed as though the assignments collected came from a broad range of ability levels, and that they were representative of the ability levels of all students. The low number of assignments, however, remains a limitation of the study.

An advantage of collecting assignments that had been written for students' content subjects is that transfer can practically and effectively be measured. With these assignments, it can clearly be seen whether students apply the abilities addressed in the academic literacy course to their content subjects. Moreover, students are not overburdened (and over-assessed) due to additional writing assignments being required.

6.5.2.1 Discussion

As is the case with Section 6.5.1, the pre-assignment results discussed in this section and Section 6.5.3.1 differ slightly from those discussed in Chapter 5 due to the different sample used (out of necessity – see Sections 6.4 and 6.5.2). However, pre-assignment results are comparable to those of the sample used in Chapter 5.

In Table 6.2, the pre-assignment refers to the assignment done at the start of the intervention in February, and the post-assignment refers to the assignments done at the end of the year in students' content subjects (with due dates from August to October). Chapter 5 discusses the difference between pre-assignments and assignments completed at the end of the first semester (thus, in May). As two different samples were used for Chapters 5 and 6, it was not possible to include these mid-assignments in the analysis below.

The internal reliability of the various sections of the rubric was high, with Chronbach's Alpha values of between 0.63 and 0.85 for the sections "Structure and development", "Academic writing style", "Editing" and "Use of source material". This would imply that the instrument used can be considered to be internally reliable.

As can be seen from Table 6.2, there was a significant and very large improvement for all of the aspects measured by means of the rubric, with the exception of "Paraphrasing information from source texts". Though this aspect did improve by 3.39 percent, the improvement was not found to be significant. Although all areas of student writing improved significantly, the greatest improvements can be seen in "Use of source material" (maybe not surprising, as students are rarely required to use source material at secondary education level) and "Academic writing style". It would also seem as though students effectively transfer all academic literacy abilities measured in the rubric used for

Table 6.2: Improvement between the pre- and post-intervention writing assignments (assessed by means of rubric) (2nd semester)

	Mean (pre-assignment; post-assignment) / 7	Improvement (%)	N	Std. Deviation (pre-assignment; post-assignment)	Sig. (2-tailed) (<i>p</i> value)	Effect Size*
STRUCTURE AND DEVELOPMENT	4.52; 6.04	21.80	38	0.72; 0.90	<.001	2.14
1. Development of main argument	4.74; 5.97	17.67	38	0.80; 1.15	<.001	1.56
2. Paragraph development (topic sentences, main ideas, supporting information)	4.63; 6.18	22.18	38	0.88; 0.96	<.001	1.76
3. Relevance of content to topic	4.79; 6.63	26.32	38	0.94; 0.63	<.001	1.97
4. Linking devices (structuring thought with discourse markers, pronouns etc.)	4.24; 5.87	23.31	38	0.97; 1.02	<.001	1.68
5. Conclusion	4.18; 5.55	19.55	38	0.90; 1.83	<.001	1.53
ACADEMIC WRITING STYLE	4.23; 5.90	23.81	38	0.78; 0.76	<.001	2.12
6. Syntax: phrase and clause structure	3.79; 4.84	15.04	38	1.12; 1.08	<.001	0.94
7. Academic vocabulary	4.29; 6.18	27.07	38	0.84; 0.98	<.001	2.27
8. Style (formality; rhetorical mode)	4.61; 6.66	29.32	38	0.92; 0.67	<.001	2.24
EDITING	4.16; 5.47	18.80	38	0.83; 0.73	<.001	1.58
9. Spelling, capitalisation and punctuation	4.11; 5.26	16.54	38	1.11; 1.01	<.001	1.04
10. Concord and tense	4.21; 5.68	21.05	38	0.84; 0.66	<.001	1.75
11. Layout and format	3.79; 6.37	36.84	38	1.21; 0.79	<.001	2.13
USE OF SOURCE MATERIAL ⁴⁵	2.86; 6.29	47.17	29	0.78; 0.97	<.001	4.22
12. Referencing technique	2.39; 6.14	53.57	28	1.07; 1.49	<.001	3.52
13. Appropriately citing quotations	2.60; 6.12	50.29	25	0.91; 1.56	<.001	3.86
14. Relevance of source data	4.19; 6.81	37.36	26	1.10; 0.57	<.001	2.39
15. Integration of source data with text (synthesising)	2.76; 5.81	43.54	21	0.94; 1.57	<.001	3.23
PLAGIARISM						
16. Paraphrasing information from source texts (expressed as a plagiarism percentage)	10.42%; 7.03%	3.39	38	15.46; 13.87	.310	0.26
* Effect sizes: <.10: trivial; .10 - .30: small to medium; .30 - .50: medium to large; >.50: large to very large						

⁴⁵ The sample sizes in this section are smaller than in previous samples as not all writing assignments from the various content subjects required source material as part of the assignment. In some cases, only some aspects of the use of source material (e.g. including a bibliography) were required.

this study to their content-subjects, considering that the post-assignments analysed for the current study were all writing assignments prescribed in students' content subjects, with no connection to the academic literacy subject.

In Chapter 5, it is shown that there were moderate to large improvements in students' marks between the pre-assignment and mid-assignment. It is not possible to make conclusive inferences by comparing those data to the data discussed in the current chapter, due to the different samples used. However, considering that pre-test results for the two samples are comparable, it would seem as though the improvements seen after a semester-long academic literacy course are much more pronounced after a year-long academic literacy course – often with effect sizes of three to four times higher than those found after only a semester-long academic literacy course. There could be several reasons for this. One reason might be that the second semester is more effective in addressing students' academic literacy abilities. Another reason might be that academic literacy abilities require an extended period of time to be ingrained, and that students are only able to fully utilise the academic literacy abilities after having been actively engaged with them for a relatively long period. One way of determining which of these two cases are more likely would be to do a comparative study in which writing results of students who were considered at-risk (and who therefore completed two semesters of the academic literacy course) are compared to students who were not considered to be at-risk (and who therefore only completed the second semester of the academic literacy course – a group that was excluded for the current study).

6.5.3 Instrument 3: Quantitatively assessing an extended writing assignment

The same writing assessments used in conjunction with instrument 2 (Section 6.5.2) were used in conjunction with this quantitative instrument, and the same limitations exist. As was done in Section 5.5.3, Storch and Tapper's (2009) guidelines were followed in quantitatively assessing students' writing assignments – the adapted guidelines, limited to students' accuracy and academic vocabulary, are summarised in Appendix C, and the reasons for these adaptations are elaborated on in Section 5.5.3. Cobb's (2015) Vocabprofile, which was adapted from Heatley *et al.* (2002), was again used for word frequency counts based on Coxhead's (2000) Academic Word List (AWL). The entire sample of 38 writing assignments was used for this instrument.

6.5.3.1 Discussion

Significant improvements could be seen in students' accuracy scores (see Table 6.3). As far as students' language accuracy is concerned, total errors per 100 words decreased by more than half between pre- and post-assignments – a significant reduction with a large practical effect size (Cohen's $d = -1.60$, $p = < .001$). The greatest reduction in errors could be seen in students' use of the plural, word choice, agreement (i.e. concord agreement and pronoun agreement), tense, prepositions, syntax and punctuation – all of which showed p values below .010 and large to very large effect sizes. Students also showed significant improvements with p values of at least below .050 and medium to large effect sizes in their use of articles and determiners, linking devices (for example discourse markers) and the possessive. These large and significant improvements in students' accuracy usage are in keeping with the significant improvements seen in their academic literacy test (Section 6.5.1) and in their writing when marked with a marking rubric (Section 6.5.2).

Table 6.3: Reduction in language mistakes between pre- and post-assignments (2nd semester)

	Mean (pre-assignment; post-assignment) / 100 words	Reduction in mistakes between pre- and post-assignments / 100 words	N	Std. Deviation (pre-assignment; post-assignment)	Sig. (2-tailed) (p value)	Effect Size *
Errors per 100 words						
Total errors	9.48; 4.22	5.26	38	3.28; 1.42	<.001	-1.60
Article / determiner errors	0.54; 0.32	0.22	38	0.53; 0.33	.030	-0.41
Plural errors	0.42; 0.14	0.28	38	0.54; 0.15	.001	-0.52
Word choice errors	2.01; 0.62	1.40	38	0.98; 0.46	<.001	-1.43
Agreement errors	0.64; 0.26	0.38	38	0.48; 0.27	<.001	-0.79
Tense errors	0.30; 0.10	0.19	38	0.28; 0.16	.000	-0.70
Preposition errors	0.57; 0.31	0.26	38	0.39; 0.27	.001	-0.68
Syntax errors	2.49; 1.03	1.46	38	1.12; 0.55	<.001	-1.30
Linking errors	0.29; 0.16	0.14	38	0.38; 0.19	.037	-0.36
Possessive errors	0.13; 0.04	0.08	38	0.17; 0.08	.009	-0.49
Punctuation	2.09; 1.24	0.85	38	0.91; 0.69	.000	-0.93
* Effect sizes: < -.10: trivial; -.10 to -.30: small to medium; -.30 to -.50: medium to large; > -.50: large to very large						

Similar improvements could not be seen in students' use of academic vocabulary (see Table 6.4). Although students seem to use more AWL families ("hypothesis",

“hypotheses” and “hypothesise”, for example, would count as one family) and types (“hypothesis” and “hypothesise”, for example, would count as two types), this increased usage is not significant ($p = .131$ and $.560$ respectively). The only significant increase in academic vocabulary usage was in the number of tokens used (for example, the word “hypothesis” might have been used five times in the text and thus be counted as five tokens). Therefore, although certain academic words are used more frequently in students’ writing, the range of words has not significantly improved in students’ post-assignments.

Table 6.4: Increase in academic vocabulary usage between pre- and post-assignments (2nd semester)

	Mean (pre-assignment; post-assignment)	Increased usage pre- and post-assignments	N	Standard deviation (pre-assignment; post-assignment)	Sig. (2 tailed) (p value)	Effect size *
Families / Types / Tokens per 100 words						
AWL Families	4.18; 4.80	0.62	38	1.24; 2.23	.131	0.50
AWL Types	4.58; 5.46	0.88	38	1.46; 2.36	.056	0.60
AWL Tokens	5.82; 8.46	2.64	38	2.03; 3.76	<.001	1.31
* Effect sizes: <.10: trivial; .10 - .30: small to medium; .30 - .50: medium to large; >.50: large to very large						

This is in contrast to what was found in Section 5.5.3.1, where significant improvements with medium effect sizes could be seen in both the use of AWL families (Cohen’s $d = 0.26$; $p = .080$) and AWL types (Cohen’s $d = 0.24$; $p = .080$) between pre-assignment (written at the beginning of the intervention) and mid-assignment (written at the end of the first-semester). This might indicate that students did not transfer the ability to use a wider range of vocabulary to the writing of their content-subject assignments (which were used as post-assignments), and that more should be done to consciously facilitate such a transfer.

6.5.4 Instrument 4: Student questionnaire

A questionnaire (Appendix F) was given to students at the end of the second semester. The aim of the questionnaire was to determine a) how important students believed various academic literacy abilities to be for them to succeed in their other subjects and b)

to which extent they believed these abilities were addressed in the academic literacy course.

According to De Vos *et al.* (2011:189), collecting data from questionnaires works best if the questionnaires are physically given to a group with instructions given to the group at large. They mention that an alternative is asking respondents to complete a web-based survey online, which was the route that the current study originally followed. However, there was an extremely low response rate, with only two out of approximately 900 students responding to the questionnaire, despite several reminders. Indeed, several authors warn that electronic questionnaires tend to have a very low response rate (e.g. Newcomer *et al.*, 2010:269; Michaelidou & Dibb, 2006:290). Because of the low response rate, questionnaires were handed out in the last week of classes. Due to various factors, such as the South African student protests in October 2015, very few students came to class in the last week, thus only 84 questionnaires could be collected. This is however considered a sufficient sample, though a limitation is that it is a sample of convenience rather than a random sample. Questionnaires were given to students who were required to do both academic literacy modules, and to those who were only required to do one academic literacy module. Of the responses received, 26 students had completed both semesters, and 58 had completed only the second semester. A further limitation is that a larger sample of students who had completed both semesters of the AGLE course could not be collected.

Due to concerns that the questionnaire proposed in Chapter 4 was too long, it was shortened by combining various categories. For example, “Understand subject terminology” and “Use subject terminology” was combined to read “Understand and use subject terminology”. Questions regarding computer literacy were also excluded, as it was decided that the questionnaire would focus purely on academic literacy. Ultimately, 48 questions were reduced to 33 questions (including four open-ended questions).

Guidelines by De Vos *et al.* (2011:196) and Brown (2001:93-100, 213-230) were followed in that questionnaire data were coded using Excel. Open responses were then analysed by means of ATLAS.ti.

6.5.4.1 Discussion

The responses of students who had completed both semesters, and those who completed only the second semester, are presented separately in Table 6.5. As can be seen from the data, the vast majority of students considered all of the listed abilities as being either “very important”, or as being “important” for them to successfully complete their content subjects. It is interesting to note that the students who had initially been identified as at-risk (and who therefore completed both modules) were more likely to rate various abilities as “very important”, while the students had not initially been identified as at-risk (and who therefore only completed the second module) were more likely to identify various abilities as being “important”. For example, 80.8% of students who had completed both semesters considered it to be “very important” to “listen effectively in class”, while 59.6% of students who had completed only the second semester considered this ability to be “very important”. However, when the responses for “very important” and “important” are added, the result is very similar (100% for students who had completed both modules versus 94.7% for students who had completed only the second module). This pattern can be seen in the responses for most abilities. This might indicate that the at-risk students are aware that they have a greater need to become proficient in various abilities than would be the case for the students who were not identified as being at-risk.

The abilities that both groups rated slightly lower (thus, they were more likely to rate these abilities as being “not very important” or “not important at all” were “Do oral presentations in class and participate in academic discussions” and “Understand and apply underlying concepts of empirical research”. Yet even these abilities were considered important by the majority of students, and thus cannot be ignored by an academic literacy intervention.

The second section of the questionnaire asked students whether various abilities were addressed “Far too much”, “Too much”, “About right”, “Too little”, or “Far too little” – a summary of student responses can be found in Table 6.6. The consensus of the majority of students was that most of the abilities were addressed “About right”, indicating that the AGLE course addresses the abilities students need to be successful in their other subjects. One exception was the ability to “Paraphrase and summarise information” –

Table 6.5: Importance of abilities in content subjects, according to students (figures expressed as percentages)

	Both modules				Second module only			
	Very important	Important	Not very important	Not important at all	Very important	Important	Not very important	Not important at all
Listen effectively in class	80.8	19.2	0.0	0.0	59.6	38.6	1.8	0.0
Take effective notes during class and from reading material	50.0	50.0	0.0	0.0	52.6	40.4	7.0	0.0
Paraphrase and summarise information	61.5	34.6	3.8	0.0	50.9	38.6	10.5	0.0
Have an appropriate reading speed	57.7	34.6	7.7	0.0	51.9	35.2	11.1	1.9
Use appropriate reading strategies for different goals	50.0	34.6	15.4	0.0	42.1	43.9	14.0	0.0
Understand assigned reading	57.7	38.5	3.8	0.0	59.6	38.6	1.8	0.0
Understand and use academic vocabulary	53.8	46.2	0.0	0.0	67.9	30.4	1.8	0.0
Understand and use subject terminology	56.0	40.0	4.0	0.0	67.9	28.6	3.6	0.0
Use the conventions of academic language (formality, vocabulary, exact language, objective language etc.)	61.5	38.5	0.0	0.0	51.8	42.9	5.4	0.0
Use subject-specific conventions	26.9	57.7	7.7	7.7	26.3	63.2	8.8	1.8
Do oral presentations in class and participate in academic discussions (during and outside of class, with students and lectures, in spoken or written form)	30.8	34.6	26.9	7.7	33.3	38.6	21.1	7.0
Analyse and comprehend (know what to do) assignment and exam questions	76.9	19.2	3.8	0.0	64.9	29.8	5.3	0.0
Structure writing (for exams, tests or assignments)	69.2	30.8	0.0	0.0	58.9	35.7	5.4	0.0
Write short coherent pieces of text	36.0	52.0	12.0	0.0	24.6	59.6	14.0	1.8
Write long coherent pieces of text	28.0	48.0	24.0	0.0	30.4	50.0	16.1	3.6
Apply relevant processes involved in academic argumentation (fact/opinion, ir/relevant information)	30.8	61.5	7.7	0.0	50.9	38.6	7.0	3.5
Develop a main argument or thesis	42.3	46.2	11.5	0.0	43.9	47.4	8.8	0.0
Interpret visual data	46.2	42.3	11.5	0.0	36.8	49.1	12.3	1.8
Create visual data and integrate it with written work	30.8	57.7	11.5	0.0	24.6	50.9	24.6	0.0
Understand and apply underlying concepts of empirical research (including methodologies)	46.2	34.6	15.4	3.8	40.4	36.8	21.1	1.8
Use different sources for research (databases, books, scientific journals, the Internet, etc.)	53.8	42.3	3.8	0.0	56.1	31.6	12.3	0.0
Process, interpret and report on gathered data	50.0	50.0	0.0	0.0	45.5	49.1	5.5	0.0
Reference a variety of sources (in-text [direct and indirect quoting and bibliography)	68.0	28.0	4.0	0.0	51.8	41.1	7.1	0.0
Use evidence from texts to support and challenge ideas	65.4	30.8	3.8	0.0	50.0	46.4	3.6	0.0
Identify relevant and reliable information	76.0	24.0	0.0	0.0	50.9	45.5	3.6	0.0
Synthesise (integrate) information from various sources	40.0	52.0	8.0	0.0	44.6	48.2	7.1	0.0
Refer to different points of view appropriately	44.0	52.0	4.0	0.0	41.1	53.6	5.4	0.0
Use appropriate time-management strategies	62.5	37.5	0.0	0.0	54.5	30.9	14.5	0.0
Use appropriate learning strategies (such as using various learning styles)	64.0	36.0	0.0	0.0	56.4	30.9	12.7	0.0

52.1% of at-risk students and 54.7% of students who were not at-risk felt that this ability was addressed either “Far too much” or “Too much”. This response should, however, be weighed up against the fact that students’ plagiarism scores were not significantly lower at the end of the year than they had been at the beginning of the year (see Section 6.5.2). It might thus be advisable to reconsider how this ability is taught rather than to reduce the time spent on it.

Students who completed both modules indicated five other abilities that more than 50% of students felt were addressed “Far too much” or “Too much” in the course. These are “Understand and use academic vocabulary”, “Understand and use subject terminology”, “Analyse and comprehend assignment and exam questions”, “Use different sources for research” and “Reference a variety of sources”. It is thus possible that some of these abilities should be focused on less in the first semester of this course; in other words, the semester that is completed only by at-risk students.

These responses must, however, again be considered in the context of other data gathered during this evaluation process. For example, many students felt that academic vocabulary acquisition was focused on too much during the year. This focus seems to have borne fruit in the first semester though, with students showing a significant increase in the number of word families, word types and word tokens used between the beginning and the end of the semester (see Section 5.5.3.1). However, this increased academic vocabulary range does not seem to have been significantly transferred to students’ other subjects; although students used significantly more word tokens in the subject-specific end-of-year assignment, neither the word families nor the word types showed a similar significant improvement (see Section 6.5.3.1). Considering the importance of vocabulary for academic success (see, for example, Gardner & Davies, 2013:1; Nagy & Townsend, 2012:101-103; Snow & Kim, 2007:123), instead of focusing less on this ability, it might be advisable to retain this focus and rather revisit the way in which academic vocabulary is approached in the subject, with the focus on assisting students to transfer this ability to their other subjects. Considering that all the abilities listed were considered “important” or “very important” by the majority of students in Table 6.5, a similar reconceptualization might be worthwhile for other abilities that did not show a satisfactory improvement.

Table 6.6: Extent to which abilities are addressed in academic literacy course (figures expressed as percentages)

	Students who only completed second semester (not at risk)					Students who completed both semesters (at risk)				
	Far too much	Too much	About right	Too little	Far too little	Far too much	Too much	About right	Too little	Far too little
Listen effectively in class	27.3	22.7	40.9	4.5	4.5	14.5	16.4	58.2	5.5	5.5
Take effective notes during class and from reading material	21.7	26.1	34.8	13.0	4.3	16.4	18.2	40.0	18.2	7.3
Paraphrase and summarise information	39.1	13.0	39.1	4.3	4.3	18.9	35.8	30.2	9.4	5.7
Have an appropriate reading speed	28.6	19.0	52.4	0.0	0.0	12.7	14.5	67.3	5.5	0.0
Use appropriate reading strategies for different goals	26.1	17.4	43.5	8.7	4.3	10.9	27.3	47.3	14.5	0.0
Understand assigned reading	31.8	18.2	50.0	0.0	0.0	10.9	29.1	56.4	3.6	0.0
Understand and use academic vocabulary	34.8	21.7	34.8	8.7	0.0	18.5	24.1	48.1	9.3	0.0
Understand and use subject terminology	30.4	26.1	39.1	4.3	0.0	12.7	27.3	43.6	12.7	3.6
Use the conventions of academic language	34.8	13.0	47.8	4.3	0.0	9.3	29.6	53.7	7.4	0.0
Use subject-specific conventions	13.0	17.4	47.8	13.0	8.7	5.7	22.6	54.7	15.1	1.9
Do oral presentations in class and participate in academic discussions	13.0	17.4	43.5	17.4	8.7	11.1	11.1	61.1	9.3	7.4
Analyse and comprehend assignment and exam questions	26.1	26.1	39.1	8.7	0.0	14.8	22.2	51.9	5.6	5.6
Structure writing (for exams, tests or assignments)	17.4	26.1	43.5	13.0	0.0	17.0	24.5	41.5	11.3	5.7
Write short coherent pieces of text	26.1	17.4	43.5	13.0	0.0	12.7	20.0	52.7	12.7	1.8
Write long coherent pieces of text	26.1	8.7	60.9	4.3	0.0	12.7	21.8	52.7	12.7	0.0
Apply relevant processes involved in academic argumentation (fact/opinion, ir/relevant information)	21.7	21.7	43.5	13.0	0.0	16.4	21.8	56.4	5.5	0.0
Develop a main argument or thesis	21.7	21.7	47.8	8.7	0.0	25.5	18.2	47.3	9.1	0.0
Interpret visual data	17.4	30.4	47.8	4.3	0.0	14.8	11.1	51.9	18.5	3.7
Create visual data and integrate it with written work	13.0	26.1	52.2	4.3	4.3	7.3	14.5	52.7	18.2	7.3
Understand and apply underlying concepts of empirical research	13.0	13.0	65.2	4.3	4.3	12.7	20.0	49.1	14.5	3.6
Use different sources for research	39.1	17.4	34.8	8.7	0.0	27.3	23.6	41.8	3.6	3.6
Process, interpret and report on gathered data	18.2	13.6	63.6	4.5	0.0	16.7	11.1	64.8	7.4	0.0
Reference a variety of sources	47.8	30.4	17.4	0.0	4.3	31.5	22.2	42.6	3.7	0.0
Use evidence from texts to support and challenge ideas	31.8	9.1	54.5	4.5	0.0	22.2	20.4	50.0	7.4	0.0
Identify relevant and reliable information	30.4	13.0	52.2	0.0	4.3	22.2	20.4	51.9	3.7	1.9
Synthesise information from various sources	13.0	21.7	52.2	8.7	4.3	13.0	33.3	48.1	3.7	1.9
Refer to different points of view appropriately	17.4	21.7	52.2	4.3	4.3	13.0	18.5	51.9	16.7	0.0
Use appropriate time-management strategies	26.1	13.0	47.8	8.7	4.3	13.2	20.8	39.6	17.0	9.4
Use appropriate learning strategies	34.8	8.7	43.5	8.7	4.3	22.6	11.3	47.2	11.3	7.5

The fact that at-risk students' feel that too much time is spent on using different sources for research and referencing a variety of sources, might indicate that content that is addressed in the first semester is being duplicated in the second semester, since fewer students who were not at risk felt the same way. Still, even if to varying degrees, both groups seemed to feel as though too much time was spent on this ability. Conversely, this was the ability that received the most positive comments under the qualitative question "Which abilities or strategies taught in your academic literacy module(s) did you find most helpful in your other subjects?". Thus, if less attention is paid to this ability in future, the course developer will have to carefully consider how to approach this so as not to lose the positive impact it seems to have in terms of transfer to other subjects.

The results from the quantitative data are reflected in the answers provided for the two open-ended questions, namely "Which of these abilities do you think need to be focused on more in your academic literacy course" and "Which of these abilities do you think need to be focused on less in your academic literacy course". One inconsistency between the qualitative and quantitative data is that nineteen students commented that they would want the academic literacy course to focus more on referencing (by far the ability which was commented on most), while the majority of responses in the quantitative data indicated that students thought this ability was focused on too much.

The abilities that students felt were most helpful in their content subjects were referencing (38 comments), writing assignments and structuring text (11 comments), paraphrasing and summarising (9 comments), learning styles (6 comments), academic language and identifying relevant (and a variety of) sources (5 comments each). As mentioned earlier, there is a discrepancy between the large number of positive comments regarding how helpful referencing was in students' other subjects and how many students felt that this ability was focused on too much in the academic literacy course. Another discrepancy is in students' opinions of paraphrasing and summarising – while many students seemed to feel that these abilities were particularly useful in their content subjects, more than 50% of students indicated that these should be focused on less in the academic literacy course. Any changes made to how these abilities are addressed should therefore be carefully considered.

When asked what other comments students wanted to make regarding the academic literacy course, several interesting observations were made. General comments about the course were mostly positive, with comments including that it was a “fair and necessary introduction to the academic requirements of a higher learning institution”, and that “this module was extremely important and it really helps the first year students to understand how to write academically at university level”. One student indicated that “after the module my views have drastically changed towards the subject, it is extremely important”. Another student said that “it is a very enlightening course that empowers first years especially I know now to get integrated into the university (sic)”. One student even said that “this subject should be made compulsory (both AGLE111 and AGLE121) for all first year students”.

Though in the minority, there were also some negative comments regarding the course. One student stated that “I am a third year who did not take AGLE in my first year or second year and passed all my modules I did it without the aid of AGLE”. Another student indicated that “most of the things taught are already known. Most of it is a waste of time and time consuming”. It is to be expected that some students would be more prepared for university studies than others, yet the fact that the majority of students find the module valuable is an indication that it addresses a need that exists due to the gap between high school education and what is expected at university level.

One theme that came through strongly in the responses to this question was that of transfer to other subjects. Several students indicated that they use the abilities taught in the academic literacy module in their other subjects. Comments included that “It is helpful because I also apply it to other modules”, “It helps a lot in other subjects and you learn how to use it in other subjects”, “ALGE111 has helped me a lot in writing my academic assignments” and “It does help regarding marks when applying to other subjects (sic)”. However, several students indicated that transfer was not optimal, with one student stating that the module was “not as important as what I thought it was as it has little application in my field of study (sic)”. Another student suggested that “seeing that the subject is compulsory for certain fields, a class can be formed discussing relevant topics”. It would thus seem as though transfer to other subjects might be improved even more by making the modules more discipline-specific – a trend already followed by many similar academic literacy interventions.

6.5.5 Instrument 6: Correlating academic literacy achievements with other variables

A final instrument that was used for the current study was correlating students' academic literacy achievements with other variables. Students' AGLE111 and AGLE121 examination and final marks, their TALL marks, as well as the marks they had received in their content-subjects were correlated with each other. The same sample of 292 students that was used in Section 6.5.1 was used for these correlations.

6.5.5.1 Discussion

Spearman's rho was used to draw non-parametric correlations (see Table 6.7)⁴⁶. Significant and moderate to strong correlations were found between the TALL and a) the AGLE111 examination ($r_s = 0.48^{47}$; $p = < .001$) and b) the AGLE121 examination ($r_s = 0.50$; $p = < .001$), as well as between the AGLE111 and AGLE121 examinations ($r_s = 0.49$; $p = < .001$), supporting the assumption made in Section 6.5.1 that these three tests/examinations can be considered equivalent. Moderate to strong correlations were also found between the AGLE111 examination and the AGLE111 final mark ($r_s = 0.63$; $p = < .001$) as well as the AGLE121 examination and the AGLE121 final mark ($r_s = 0.40$; $p = < .001$), indicating that the continuous assessment and examinations in both modules are fairly well aligned, though more so in the case of the first semester than the second semester. It might be worthwhile, however, to consider how the second semester's (i.e. AGLE121) continuous assessment could be aligned even better with its examination.

As had been speculated in Section 5.5.4.1, a stronger correlation was found between students' marks in their other subjects, and the final marks they received in AGLE121 ($r_s = 0.49$; $p = < .001$) than was the case for the final marks they received in AGLE111 ($r_s = 0.31$; $p = < .001$). This is in line with the findings from previous studies (e.g. Van Dyk, 2015:180; Van Rooy & Coetzee-Van Rooy, 2015:10-11; Mhlongo, 2014:80-82) which found strong correlations between the results of academic literacy modules and students'

⁴⁶ Spearman rank order correlations were performed to indicate whether there is a relationship between students' various academic literacy test results as well as their results in their other modules. The purpose was not to make a prediction of any variable and therefore no regression analyses were performed.

⁴⁷ Once again, there is a difference between the figures reported on in this section and those reported on in Chapter 5 due to different samples being used.

results in their other subjects. The implication of this moderate to strong correlation is that it is likely that students' marks in their other subjects are likely to increase if their academic literacy marks increase. This correlation would then also suggest that the AGLE course, in its entirety, would seem to impact on students' other marks, indicating that transfer of academic literacy abilities to students' other subjects has occurred.

Table 6.7: Non-parametric correlations using Spearman's rho (2nd semester)

		TALL	AGLE111 exam	AGLE111 final	AGLE121 exam	AGLE121 final	Average other subjects
TALL	Correlation Coefficient (r_s)	1.00	.48	.27*	.50	.27	.01
	Sig. (2-tailed)	.	<.001	<.001	<.001	<.001	.945
	N	234	234	234	234	234	234
AGLE111 exam	Correlation Coefficient (r_s)	.48	1.00	.63	.49	.35	.07
	Sig. (2-tailed)	<.001	.	<.001	<.001	<.001	.285
	N	234	234	234	234	234	234
AGLE111 final	Correlation Coefficient (r_s)	.27	.63	1.00	.30	.50	.31
	Sig. (2-tailed)	<.001	<.001	.	<.001	<.001	<.001
	N	234	234	234	234	234	234
AGLE121 final mark	Correlation Coefficient (r_s)	.27	.35	.50	.40	1.00	.49
	Sig. (2-tailed)	<.001	<.001	<.001	<.001	.	<.001
	N	234	234	234	234	234	234
AGLE121 exam mark	Correlation Coefficient (r_s)	.50	.49	.30	1.00	.40	.00
	Sig. (2-tailed)	<.001	<.001	<.001	.	<.001	.992
	N	234	234	234	234	234	234
Average other subjects	Correlation Coefficient (r_s)	.01	.07	.31	.00	.49	1.00
	Sig. (2-tailed)	.945	.285	<.001	.992	<.001	.
	N	234	234	234	234	234	234

6.6 Conclusion

This study has used a mixed method approach. A strength of this type of paradigm is that the variety of evidence obtained from a mixed design can serve as a valuable validity check (Lynch, 2003:29), especially in studies such as the current one where it was impossible to use a control group. The various instruments that have been used have indeed provided rich evidence, and have served as an effective validity check.

This study has several limitations. Firstly, a quasi-experimental research design was used, as no control group was available. This is a limitation that would apply to most studies of this nature, as control groups become nearly impossible due to ethical considerations. This limitation was addressed by using a variety of instruments which acted as cross references for each other. Although such triangulation cannot rule out that an improvement in students' academic literacy levels was (at least in part) due to factors other than the academic literacy intervention, it does allow fairly strong assumptions regarding the causality of the intervention. Secondly, the study focused on summative data, and did not include descriptive data collected throughout the year which might have explained the successes and/or failures of the course (cf. Lynch, 1996:169). Thirdly, due to logistical reasons, only a relatively small sample size could be used in Sections 6.5.2 and 6.5.3. Finally, only the results of students who completed both modules were considered for the first three instruments. Future research should attempt to minimise these limitations as far as possible.

The findings of this study seem to indicate that the AGLE course comprehensively and effectively addresses a wide range of academic literacy abilities. This impact, however, is much more pronounced after a year-long academic literacy course than it is after only a semester-long academic literacy course – this is similar to findings by Van Dyk *et al.* (2011a:497), indicating that the duration of academic literacy courses is likely to influence their impact – at least in the case of students who had been identified as at-risk. Almost all abilities that were measured improved significantly with moderate or large effect sizes between the beginning and the end of the course. Student feedback regarding which abilities were necessary in their content subjects, and to which extent these abilities were addressed in the academic literacy course, support the conclusion that the course effectively addresses the range of academic literacy abilities listed in the questionnaire. Areas where statistically significant improvements could not be shown were students' ability to avoid plagiarism and the range of their academic vocabulary. It might be worthwhile to reconsider how these aspects are addressed in the course in future.

Another possibility, as raised by comments in Section 6.5.4.1, would be to consider making this generic course more discipline specific. Although it would seem as though much transfer already takes place from the academic literacy course to students' content

subjects, it would seem as though students feel that the course might have been more relevant to their studies, and that certain abilities (for example students' ability to use a range of academic vocabulary) are not currently being transferred optimally.

A final comment needs to be made regarding the assumptions of causality made in this chapter. Without a traditional experimental design, it is not possible to conclusively state that the academic literacy course was responsible for the improvement in students' academic literacy levels (cf. Bak, 2004:22). However, considering findings in the literature that it is highly unlikely that students' academic literacy abilities can improve significantly without some type of intervention (see Section 2.4.3), and students' own perceptions that the academic literacy course assisted them in their content subjects (see Section 2.5.4.1), it would seem likely that the academic literacy course had a meaningful influence in students' improved academic literacy abilities. Though conclusive proof cannot be provided in the context of this academic literacy intervention, as would be the case with most academic literacy interventions in South Africa, that should not deter researchers from collecting as much data as possible from which strong deductions can be drawn, and based upon which informed decisions can be taken.

While Chapters 5 and 6 have aimed to verify the evaluation design proposed in Chapter 4, Chapter 7 endeavours to validate this design by firstly reflecting on the efficacy of its implementation in the previous two chapters, and secondly by obtaining feedback from academic literacy intervention coordinators across the country on whether the evaluation design would be suitable for their respective contexts. Thereafter, a revised evaluation design is proposed.

CHAPTER 7

IMPACT MEASUREMENT: REFINING AN EVALUATION DESIGN FOR ACADEMIC LITERACY INTERVENTIONS THROUGH SELF- REFLECTION AND FEEDBACK (ARTICLE 5)

7.1 Prelude to Chapter 7

Chapter 4 proposed a preliminary evaluation design (based on relevant studies in the field, discussed in Chapter 3) which could be used to assess a variety of academic literacy interventions. An attempt to verify this design was made by applying it to an academic literacy course at the North-West University's Potchefstroom Campus. This process is reported on in Chapters 5 and 6. During this process, several problems were found with the implementation of various of the proposed instruments. The current chapter begins by examining these challenges.

Further, the evaluation design was validated by sending it to academic literacy intervention coordinators across South Africa. Coordinators were asked which of the instruments would be relevant for their particular interventions, how instruments could be improved, and which instruments could be added to the design.

Based on the results from this verification and validation process, a revised evaluation design for academic literacy interventions is proposed in this chapter. Where appropriate, revised instruments, or revised guidelines for using these instruments, are included.

7.2 Introduction

The goal of research, say Cole *et al.* (2005:2), should be to make a dual contribution to both practice and theory. Further, they state that “research should assist in solving practical problems of practitioners, problems that are either current or anticipated” (Cole *et al.*, 2005:2). In the sphere of “design research” (see Section 2.2.5), which centres around a viable artefact which has been created in the form of a model, method or construct, it is important that the design artefact's utility, quality and efficacy be demonstrated by rigorously evaluating it (Cole *et al.*, 2005:3-4; Hevner *et al.*, 2004:77).

Cole *et al.* (2005:4) contrast design research to action research, through which complex social processes are studied by means of introducing some type of change to these processes, and subsequently evaluating their effect.

Reflecting on a research process (a stage which Cole *et al.* [2005:16] argue is not formalised enough in design research, and a stage which design research could incorporate more formally from action research) is a pivotal part of responsible scholarship. It is crucial to understand what worked, what did not work, and what the reasons therefore might be. Dewey's (1933) definition of critical reflection is still drawn on today. According to Dewey (1933:6), "[r]eflection is an active, persistent and careful consideration of any belief or supposed form of knowledge in light of the grounds supporting it and future conclusions to which it tends". Yost, Sentner and Forlenza-Bailey (2000:39) add that "[r]eflection implies that something is believed in or disbelieved because of some evidence, proof, or grounds for that belief". For example, when assessing the impact of a phenomenon, it is important to reflect upon the measures and methods used to assess such impact to determine whether these resulted in reliable and valid evidence, or whether they could be altered to obtain richer, more valid and more reliable information in future. Only by doing this can scholarship in the field be improved upon.

The current chapter indicates how a specific design artefact – in this case an evaluation design which aims at assessing the impact of a wide range of academic literacy interventions – can be evaluated and be improved upon by critically reflecting upon its utility, quality and efficacy (cf. Cole *et al.*, 2005:3-4; Hevner *et al.*, 2004:77). The review of this specific evaluation design should be considered a case study to indicate the importance of reviewing an artefact as part of a cyclical process of continuous improvement.

7.3 Background to the development of the evaluation design

Most South African universities currently have some form of academic literacy intervention in place, be it credit-bearing courses or other interventions such as writing centres (see, amongst others, Carstens, 2013b; Winberg *et al.*, 2013; Van Dyk *et al.*, 2011a; Ngwenya, 2010; Fouché, 2009; Van Dyk *et al.*, 2009; Archer, 2008; Pienaar,

2005; Van Wyk, 2002). In fact, after having surveyed the websites of all 25 South African universities (including traditional universities which offer mainly degree courses, universities of technology, as well as comprehensive institutions which offer both traditional university and university of technology qualifications [cf. Higher Education South Africa, 2005], but excluding private universities), only two South African universities (Sol Plaatje University and the University of Mpumalanga) could be found which did not offer any form of academic literacy intervention, both of which were small universities established in the past two years. As discussed in Chapter 3, despite the abundance of academic literacy interventions, very few studies have attempted to determine the impact of these interventions, not only in South Africa, but also internationally (Fouché, 2015:4; Mhlongo, 2014:47; Butler, 2013:80; Terraschke & Wahid, 2011:174).

Determining the impact of an intervention falls within the sphere of evaluation research (see Section 2.2.5). Babbie (2005:360) defines evaluation research as the “process of determining whether a social intervention has produced the intended result”. Weinbach (2005:2) provides a similar definition, stating that evaluation research is “the systematic use of research methods to make judgments about the effectiveness and the overall merit, worth, or value of some form of (...) practice”. Brown (2001:15) focuses on evaluation in educational contexts, and defines it as an ongoing process of gathering, analysing and synthesising data, so as to improve the various aspects of a curriculum as well as the curriculum as a whole. Hawkey (2006:12-13) also addresses the concept of intentional positive washback which is alluded to above, where feedback from evaluations can positively influence both curriculum and teaching style.

De Vos *et al.* (2011:449) argue that, due to an increased focus on accountability, stakeholders want to see proof that interventions work, of how they work, and of how they can be improved. Lynch (2003:1), whose research focuses on the evaluation of language programmes, states that programme evaluation draws on the data gathered from language assessments so as to reflect on the intervention in question, make decisions and take appropriate actions based on the feedback received. Bachman and Palmer (2010:21) agree that in the process of evaluation, language assessments are used primarily to inform the decisions and value judgements that are made about a specific programme, intervention, or construct.

In an attempt to address the gap in the literature on assessing the impact of academic literacy interventions, Chapter 4 suggested an evaluation design which could be used to assess the impact of a variety of academic literacy interventions in the South African context. This originally proposed evaluation design (see Figure 4.2, Section 4.5) contained a variety of instruments from which the researcher could select a combination which was relevant to any particular academic literacy evaluation context. This evaluation design has since been implemented (see Chapters 5 and 6) by assessing the impact of a first-year academic literacy course at a South African university. Furthermore, the internal validity of the evaluation design was tested in several ways. Firstly, its construct validity was determined in Chapters 5 and 6. Secondly, its content and face validity were determined by sending it to academic literacy specialists at all South African universities to determine whether the evaluation design was appropriate and adequate for their respective contexts, and how they would adapt it so as to be more appropriate. This is discussed further in Section 7.6.

The purpose of the current chapter is to indicate how reflection and external feedback, as part of a research process, can contribute to an improved product – in the context of this study, an improved evaluation design. This is done by firstly reporting on challenges encountered in the implementation of the design (thus its use in practice), and to make suggestions regarding implementation based thereon (Section 7.4). Secondly, the chapter reports on the feedback received from South African academic literacy specialists, who provided feedback on the evaluation design’s theoretical construct as well as possible envisaged problems regarding its implementation (Section 7.5). Based on the feedback from the above, an adapted evaluation design is then proposed. Before the research pertaining to this chapter is reported upon, the originally proposed evaluation design (see Section 4.5) is briefly summarised.

7.4 A summary of the originally proposed evaluation design

Lynch (1996:156) states that a design can be seen as a methodological strategy for evaluating programmes (see Section 2.2.6). Bamberger *et al.* (2012:613) explain that an evaluation design “is the plan for conducting an evaluation, including methods and sources of data collection, deliverables, methods of analysis, and a timeline”. Judd and Keith (2012:36) argue that outcome data alone (e.g. data which show an improvement in

students' academic literacy levels) may not be sufficient to draw conclusions that the intervention caused the outcome, "but if it is collected in the context of a framework of other knowledge that supports causal inferences, then it may be justified".

The definition of impact in the context of academic literacy interventions put forward in Section 2.2.1 (also see Fouché, 2015:3) is that it is "i) the observable improvement in academic literacy abilities between the onset and the completion of an academic literacy intervention, and ii) the extent to which these abilities are necessary and applied in students' content subjects". Based on the definition above, Chapter 4 proposes that, in order to comprehensively assess the impact of an academic literacy intervention, two aspects need to be examined by means of an evaluation design. The first is to determine whether there was an improvement in students' academic literacy levels, and what the extent of this improvement might be. The second is to determine to which extent these abilities are needed in, and transferred to, students' content subjects.

Judd and Keith (2012:37) point out that the environments of institutions and departments differ; further, these environments can impact students as a result of different input characteristics, "resulting in outcomes that have multiple and compounded influences". Mhlongo (2014:4) echoes this sentiment and warns that the context of academic literacy interventions must be kept in mind when these are assessed, due to the unique challenges faced by various higher education institutions. The proposed evaluation design was therefore flexible in that evaluators could choose a combination of tools from a variety of evaluation instruments to suit their specific contexts. Using a combination of instruments would ensure that data are triangulated by both method and source, and would also contribute to the validity of findings (Judd & Keith, 2012:40; Lynch, 2003:152; 1996:59-60; Jick, 1979:602). Keeping this in mind, it was suggested in Chapter 4 that triangulation be used to validate findings. It was proposed that, where it was possible to use control groups, at least two instruments be used, and in cases where control groups were not possible, at least three instruments be used (see Section 4.5). Regardless of which combination of instruments was decided upon, it was recommended that at least one measures whether students' academic literacy abilities had improved, and at least one measures whether these abilities are required in, and/or transferred to, students' other subjects. By using several instruments that measure impact from different perspectives, triangulation by both method and source becomes possible.

Different combinations of these instruments were proposed for four broad types of courses, namely generic academic literacy interventions, subject-specific academic literacy interventions, collaborative academic literacy interventions and limited purpose academic literacy interventions (e.g. writing centres or reading programmes) (cf. Van de Poel & Van Dyk, 2015:169-173). The proposed instruments, which are expanded on in Chapter 4, are as follows:

Instruments which assess an improvement in students' academic literacy abilities, and the extent of this improvement:

- Generic academic literacy test
- Subject-specific academic literacy test⁴⁸
- Generic extended writing assignment (assessed by means of a rubric)
- Subject-specific extended writing assignment (assessed by means of a rubric)
- Extended writing assignment assessed by means of quantitative measures

Instruments which assess whether academic literacy abilities are needed in, and transferred to, students' other subjects:

- Student questionnaire
- Lecturer questionnaire
- Qualitative feedback from primary stakeholders
- Content analysis of study material
- Correlating academic literacy achievements with other variables

The following section reports on challenges experienced in the implementation of this evaluation design in an academic literacy course for first-year students at a South African university. In the discussion below, the reader should keep in mind that only five of the suggested instruments were found to be appropriate for this specific study (see Chapters 5 and 6). Further research using different instruments would likely encounter additional, or different challenges.

⁴⁸ Note that subject-specific academic literacy tests and subject-specific writing assignments could also provide information regarding the transfer of academic literacy abilities.

7.5 Reflecting on the challenges experienced in implementing the design

Before the challenges experienced are expanded upon, the context of the implementation of the evaluation design, as discussed in detail in Chapters 5 and 6, is summarised here. The generic academic literacy course which is the focus of the case study is aimed at first-year students from a wide variety of faculties. It runs over two semesters. Students who are identified as at-risk at the beginning of the year (by means of the Test of Academic Literacy Levels [TALL]) are required to complete both semesters of this course (624 students in 2015), while students who are not identified as at-risk are only required to complete the second semester (352 students in 2015; thus, a total of 976 students registered for the second semester of this course). Students have two 50-minute classes per week for 14 weeks per semester. Only students who had originally been identified as being at-risk were included in the sample used in the study in which the evaluation design was implemented. The following instruments were originally selected for assessing the impact of this course:

- Generic academic literacy test
- Generic extended writing assignment (assessed by means of a rubric)
- Generic extended writing assignment assessed by means of quantitative measures
- Student questionnaire
- Correlating academic literacy achievements with other variables

As the course is generic in nature (due to the wide variety of students it services), subject-specific instruments were not viable. For the same reason, content analysis of study material as well as lecturer questionnaires were seen to be impractical – the vast number of stakeholders and the large number of content-subjects (97 in total) made these instruments very difficult to implement effectively.

“Through praxis, critical consciousness develops, leading to further action” (Baum, Macdougall & Smith, 2006:856). The suggested evaluation design was accordingly based on theory and best praxis grounded in the literature (see Chapters 2, 3 and 4). Yet, as pointed out by Baum *et al.* (2006:856), theory must be put to the test in practice. The “critical consciousness” that develops during this process is crucial for further improvement (in the case of the current study, of an evaluation design as artefact). By

grounding research in both theory and practice, research consequently contributes to both practice and theory (cf. Cole *et al.*, 2005:2).

According to Yost *et al.* (2000:43-44), reflecting on one's own experience leads to improved praxis; "reflective individuals continually seek new information for the purpose of ongoing growth and change". That then is the purpose of reflecting on the challenges encountered in implementing the evaluation design (see Chapters 5 and 6). Four main challenges were identified in the implementation of the evaluation design in assessing the impact of a specific academic literacy course for first-year students at a South African university. These were 1) the dilemma of evaluator as insider or outsider, 2) obstacles related to using an academic literacy test as pre- and post-test, 3) obstacles related to using an extended writing assignment, and 4) obstacles related to using questionnaires. These challenges should be kept in mind by future researchers wishing to use the same instruments.

7.5.1 The insider/outsider dilemma

A preliminary point that should be considered before undertaking an evaluation of an intervention is that of insider and outsider roles with regard to evaluation studies (see Hawkey, 2006:5-6). In evaluating the impact of the course discussed in this section, the author of this study was what would be defined as an outsider. She does not work at this specific university, nor does she have any stake in the outcomes of the evaluation. What is more, the researcher was not on site during the evaluation period, and relied on the assistance of course coordinators to gather data on her behalf. Having an outsider conduct an evaluation has the advantage of objectivity being more likely. However, many of the challenges described in Sections 7.5.2 to 7.5.4 also stemmed from being an outsider. An insider (and specifically a course coordinator) would have had more control over modifying assessment practices to suit a planned evaluation design. Furthermore, being in closer contact with the students would have provided increased control over data gathering. Although coordinators from the course in question provided much cooperation, after having completed the evaluation, I would recommend that future studies assessing the impact of academic literacy interventions either have insiders conducting the evaluation, or have outsiders who can be on site during the evaluation period (which, of course, involves additional resources, in particular financial resources).

7.5.2 The use of an academic literacy test as pre- and post-test

A further challenge encountered involved the use of a generic academic literacy test as a pre- and a post-test. The TALL was used as a pre-test. This test has been thoroughly validated and shown to be reliable (Van Rooy & Coetzee-Van Rooy, 2015:8, 13; Van Dyk, 2010:17, 155, 199-284 ; Van Der Walt & Steyn, 2007:146-156; Van Dyk, 2005:42-44), and was thus a judicious choice as a pre-test. The academic literacy course's two examinations were used as post-tests; one was written in May, the other in November.

Due to practical constraints, in particular time constraints, it was not possible to have students write both the TALL and the subject's examinations at the end of each semester. As new examinations are set each year, it is not possible to subject them to the same process of validation as is the case with the TALL. Although both the TALL and the subject's examinations are based on the same construct and include the same sub-categories, the internal reliability rates of some of the examinations' sub-sections were not high enough to make it possible to compare the sub-sections of the respective tests so as to determine in which sections students had improved the most. However, the tests in their entirety could be considered equivalent, and the improvement between pre- and post-tests considered statistically significant (see Section 6.5.1). Thus, though it could be safely concluded that students' academic literacy levels as a whole improved between the TALL and the two post-tests, it was not possible to determine which academic literacy abilities improved in particular, and which did not. This makes it difficult to determine how the curriculum could best be adapted in future so as to optimally improve students' academic literacy abilities. A recommendation for future research would therefore be to ensure that the same test be used as both pre- and post-test. Where this is not possible, tests should ideally be piloted to ensure that they (and their various sub-sections) can be considered theoretically and practically equivalent before any valid conclusions can be drawn based on results from these tests.

7.5.3 The use of an extended writing assignment

Another set of challenges was encountered in assessing generic extended writing assignments written by students before, in the middle of, and at the end of the intervention. After students had completed the pre-assignment, a random sample was

selected and assessed. However, with the mid-year post-assignment, due to logistical problems very few students submitted assignments electronically as had been requested; this was only realised after hardcopy assignments had already been returned to students. As the researcher and the subject lecturers used different rubrics (and thus assessment criteria), lecturers' marks could not be used – the researcher needed the physical copies of the assignments. Also, for the third assessment instrument, namely assessing students' writing by means of quantitative measures, electronic copies of assignments were essential as word frequency counts could only be done effectively if assignments were in electronic format. Thus, the random sample had to be changed to a sample of convenience using the assignments of students who did submit electronically, which meant that most marked pre-assignments had to be discarded and a new sample had to be marked.

A further challenge that was experienced was with the second post-assignment (thus, the writing assignment submitted at the end of the second semester). It was not possible to include an individual essay into the assessment plan of the second semester for various reasons. Firstly, the student numbers were too high, and resources too few, to make it possible for lecturers to assess individually written assignments. It was also not possible for the researcher to mark more than the chosen sample of assignments. Further, the subject had already utilised all of the credits (and thus time allocated for the module) available; as a result, it was not possible to ask students to write an additional assignment over and above the assessments that were already included in the assessment plan. An alternative option was decided upon, namely to ask students to submit any individual extended written assignment (of a page or more) that was to be submitted to any of their content-subjects during the second semester. However, it became apparent that very few students are required to write individual assignments at first-year level. As a result, only a small sample of writing assignments was available at the end of the second semester. Still, asking students to submit subject assignments which are assigned by other subjects is a useful instrument for assessing whether abilities were transferred, and will thus be added to the revised evaluation design proposed in the current chapter.

Even with the advantage of hindsight, these challenges would have been difficult to overcome. The challenge of most students not submitting mid-year assignments electronically might have been addressed by more careful planning and monitoring. It is

thus vital for future researchers to ensure that communication mechanisms to lecturers and students are in place and are followed up upon. The second challenge, namely of very few individual assignments being written in students' first year of study (even in the academic literacy module), probably due to ever-increasing student numbers (cf. Calderon, 2012; Teichler, 1998), is a limitation to this type of research where a large number of students is concerned.

Another challenge with regard to the written assignments concerned assessing these assignments by means of quantitative measures. Specifically assessing accuracy scores is extremely work-intensive, and it is not practically possible to do this effectively unless significant resources (time as well as money for several well-qualified markers) are available. In this study it was only possible to use this instrument with a sample of 50 scripts, and even that took several weeks to complete. It should thus be kept in mind by future researchers that this instrument cannot be used effectively with large samples without accompanying resources.

7.5.4 The use of questionnaires

A final challenge that was encountered was with regard to the student questionnaire. Originally, the researcher had planned to send out student questionnaires at the end of the first semester, at the end of the second semester, and also to second-year students who had completed the course the year before. Questionnaires were sent electronically at the end of the first semester. The questionnaires were sent to all 624 students who had completed the first semester module. After two e-mails requesting students to complete the questionnaire, only two students had done so. The same was the case with sending questionnaires to students who had completed the course the year before. This is consistent with research that has indicated that electronic surveys tend to have a very low response rate (Singer & Ye, 2013:112; Newcomer & Triplett, 2010:269; Michaelidou & Dibb, 2006:290). Due to the problems experienced in the first semester, questionnaires were handed out in class at the end of the second semester. This yielded better results, with 84 of 976 students completing the questionnaires. This number was still fairly low, possibly due to factors such as the 2015 South African #FeesMustFall protests as well as relatively few students attending classes in the last week of the semester, yet the sample was large enough to draw deductions from. It is thus recommended that if researchers

wish to send out questionnaires electronically, an incentive be linked to completing the questionnaire (see Cobanoglu & Cobanoglu, 2003:485 for examples of how and which incentives could be provided). It should be kept in mind though that linking incentives to data collection comes with its own potential ethical problems (Cobanoglu & Cobanoglu, 2003:486). Otherwise, questionnaires should be handed out in class – late enough in the semester for the curriculum to have been completed, but at a point where most students still attend class. A further problem here might be that the questionnaire was quite long. This chapter thus proposes that the need-press questionnaire, that determines which abilities students believe they need, and which they believe were addressed in the intervention (see Section 4.4.2.1), be shortened where possible. A revised student questionnaire can be found under Appendix F.

Many of the challenges highlight the need for meticulous planning before a large-scale evaluation of the impact of an intervention is undertaken. Even if this is done (as was the case in this study), unforeseen obstacles are likely to arise which will necessitate the implementation of alternative plans. By continuously reflecting on new challenges that arise, the accompanying critical consciousness that will develop (cf. Baum *et al.*, 2006:856) will enable the researcher to keep refining theory, which can in turn be positively implemented in practice. In addition to critical self-reflection, it was considered valuable to obtain feedback on the evaluation design from other specialists in the field. The following section reports on this feedback received from academic literacy specialists from across South Africa on the usefulness and relevance of the proposed evaluation design. This “reflection of others” supplements the self-reflection as discussed in the current section.

7.6 Feedback from academic literacy specialists

A questionnaire (Appendix G) was designed to determine whether the proposed evaluation design met the needs of academic literacy specialists across a broad spectrum of universities and interventions. The coordinator questionnaire was first piloted by sending it to colleagues at two different universities and was adapted based on their feedback. Thereafter, it was sent to a selection of academic literacy specialists at all South African traditional universities, universities of technology, and comprehensive universities (cf. Higher Education South Africa, 2005). Of the 23 universities contacted,

academic literacy specialists from 14 universities replied. Of these, eight are from traditional universities, three are from comprehensive universities, and three are from universities of technology. In several cases, two or three academic literacy specialists from a university completed the questionnaire (for example, in cases where a university had an academic literacy course [or several such courses] as well as a writing centre intervention). In total, 23 responses were received. These responses represent the different types of universities, as well as the geographical locations of the various universities. Furthermore, the responses cover a broad range of academic literacy interventions. Seven specialists who responded are responsible for generic undergraduate academic literacy courses, fourteen are responsible for subject-specific academic literacy courses, two are responsible for generic postgraduate academic literacy courses (here I include one respondent who is responsible for facilitating writing retreats for postgraduate students, most of whom are PhD students, although this response might also have been appropriately classified under subject-specific postgraduate interventions), three are responsible for subject-specific postgraduate academic literacy courses, four are responsible for collaborative academic literacy interventions, eight are responsible for writing centres, and one is responsible for a reading laboratory. The number of interventions indicated above (39 in total) is greater than the total number of responses (23) as several specialists are responsible for more than one type of intervention.

7.6.1 Quantitative feedback from academic literacy specialists

The academic literacy specialists were first asked whether the various proposed instruments were “very applicable”, “applicable to a limited extent”, or “not applicable” to their respective contexts. The results obtained from the 23 respondents are summarised in Table 7.1. As can be seen from this table, the majority of respondents felt that each of the proposed instruments were either “very applicable” or “applicable to a limited extent”. The relatively large number of respondents who felt that generic academic literacy tests and generic extended writing assignments were “not applicable” makes sense if one considers the large number of undergraduate and postgraduate subject-specific courses (n=21) represented in this survey as opposed to the nine generic undergraduate and postgraduate courses.

Table 7.1: Academic literacy specialists’ perceptions of usefulness of instruments

Instruments	Very applicable	Applicable to a limited extent	Not applicable
Generic academic literacy test	8	8	7
Subject-specific academic literacy test	14	6	3
Generic extended writing assignment (rubric)	4	12	7
Subject-specific extended writing assignment (rubric)	17	4	2
Quantitatively assessing writing assignment	9	9	5
Student questionnaire	14	8	1
Lecturer questionnaire	9	12	2
Content analysis of study material	13	8	2
Correlating academic literacy results with other variables	11	7	5
Qualitative feedback from primary stakeholders	21	2	0

An attempt was made to determine whether certain instruments were more applicable to certain types of academic literacy interventions. However, a limitation of the questionnaire was that it did not allow academic literacy specialists who are responsible for more than one intervention to indicate how applicable the various instruments are to the various interventions they are responsible for (thus, one set of responses was given for all the relevant interventions, instead of a set of responses for each intervention the academic literacy specialist is responsible for). To account for this limitation, only the responses of academic literacy specialists who are responsible for only one intervention (n=16) are used in Table 7.2. Only four types of interventions are represented by this reduced sample, namely generic undergraduate academic literacy courses (abbreviated Gen UG AL), subject-specific undergraduate academic literacy courses (abbreviated SS UG AL), collaborative academic literacy interventions (abbreviated Collab AL), and writing centres (abbreviated WC).

Although only approximately half of the original responses could be used for the more detailed information presented in Table 7.2, the same trends as in Table 7.1 are evident. Although all instruments are either very applicable or applicable to a limited extent to some academic literacy interventions (meaning that none can be discarded in the revised evaluation design), there does seem to be a strong preference for subject-specific instruments (even among generic interventions – it is possible that even these interventions are integrating subject-specific elements into their curricula). Furthermore, there is an overwhelming preference for qualitative feedback from primary stakeholders,

Table 7.2: Academic literacy specialists' perceptions of usefulness of instruments, categorised into types of interventions

Instruments	Type of academic literacy intervention	Total number of responses	Number of academic literacy specialists who found the instrument...		
			Very applicable	Applicable to a limited extent	Not applicable
Generic academic literacy test	Gen UG AL	4	3	0	1
	SS UG AL	8	2	5	1
	Collab AL	2	0	0	2
	WC	2	0	1	1
Subject-specific academic literacy test	Gen UG AL	4	1	2	1
	SS UG AL	8	8	0	0
	Collab AL	2	0	1	1
	WC	2	0	1	1
Generic extended writing assignment (rubric)	Gen UG AL	4	2	2	0
	SS UG AL	8	0	7	1
	Collab AL	2	0	0	2
	WC	2	0	1	1
Subject-specific extended writing assignment (rubric)	Gen UG AL	4	2	2	0
	SS UG AL	8	8	0	0
	Collab AL	2	1	1	0
	WC	2	1	0	1
Quantitatively assessing writing assignment	Gen UG AL	4	2	2	0
	SS UG AL	8	3	4	1
	Collab AL	2	0	1	1
	WC	2	1	0	1
Student questionnaire	Gen UG AL	4	2	2	0
	SS UG AL	8	6	2	0
	Collab AL	2	0	2	0
	WC	2	2	0	0
Lecturer questionnaire	Gen UG AL	4	1	3	0
	SS UG AL	8	6	1	1
	Collab AL	2	0	2	0
	WC	2	0	2	0
Content analysis of study material	Gen UG AL	4	4	0	0
	SS UG AL	8	6	2	0
	Collab AL	2	0	1	1
	WC	2	0	2	0
Correlating academic literacy results with other variables	Gen UG AL	4	3	1	0
	SS UG AL	8	4	3	1
	Collab AL	2	0	1	1
	WC	2	0	0	2
Qualitative feedback from primary stakeholders	Gen UG AL	4	4	0	0
	SS UG AL	8	8	0	0
	Collab AL	2	1	1	0
	WC	2	2	0	0

indicating a need for thorough, descriptive, and explanatory instruments which can be used to understand the impact of academic literacy interventions. Such instruments might need to be separately developed for each specific evaluation context. Future research could however determine whether credible, transferable and dependable templates aimed at various primary stakeholders could be developed.

7.6.2 Qualitative feedback from academic literacy specialists

Though the quantitative feedback indicates that all instruments in the proposed evaluation design should be retained, qualitative feedback was essential in understanding how some instruments could be refined, and whether there was a need for additional instruments to be added to this design. This was done by asking academic literacy specialists three questions which are discussed in the remainder of this section. Lecturer responses were analysed by means of ATLAS.ti. Some themes were identified which are discussed below.

Improving instruments

The first question posed to academic literacy specialists was “Are there any ways in which you believe any of the proposed instruments could be improved?”. Half (n = 10) of the participants who answered this question indicated that the instruments were “suitable” and “addressed the intended outcomes”. A further two responses indicated that they could not comment on the question – in the first case, because the academic literacy specialist felt s/he needed to first implement the instruments before commenting on how they could be improved, and in the second case, because the academic literacy specialist felt many of the instruments were not applicable to his/her specific context, namely facilitating writing retreats for postgraduate students.

Several of the participants did, however, have suggestions as to how some of the instruments could be improved. The first comment was that the student questionnaire was too long, and that “this might serve as a deterrent to student response”. This is in line with my own deduction that a possible reason for so few electronic student questionnaires being returned after the first semester of the course might have been the length of the questionnaire (see Section 7.5.4). The original 48 questions were condensed

to 33 questions in a revised questionnaire (Appendix F). To ensure that the student and lecturer questionnaires still correspond so as to facilitate triangulation, the latter (see Appendix H) was similarly revised. Another participant highlighted the fact that the lecturer questionnaire would “have to be explained to the lecturer, and close co-operation will be necessary” – this corresponds to Jacobs’ (2005:480) findings that academic literacy specialists can assist subject specialists with bringing what they “already [know] tacitly, into the realm of overt and explicit teaching”. As is the case with the student questionnaires, this also indicates the importance of personal contact (i.e. facilitating the completion of the questionnaire). Yet, such personal contact is not always feasible due to the resources required. Trade-offs will thus have to be made, depending on the resources available and the evaluation context.

One participant commented on the appropriateness of the writing rubric for postgraduate students. S/he indicated that more categories, for example genre and audience, might be needed. The writing rubric (Appendix I) has been adapted to reflect these suggested categories. As was done in Chapters 5 and 6, any categories that are not deemed appropriate could simply be left out during assessment for different contexts. This participant also indicated that the rubric was “very limited for the types of genres that students have to produce”. This is a valid point – the rubric is aimed at relatively standard essay-type assignments, and not, for example, at genres such as laboratory reports or legal writing. However, developing appropriate rubrics for every type of academic writing genre lies outside the scope of the current study. The instrument provided is a guideline for what might be considered the most common type of academic writing, and it was proven useful in assessing subjects such as history, social anthropology, geography, physiology, and urban morphology in Section 6.5.2 (cf. Carstens, 2009:161-162). However, to assess specialised and niche genres of writing, alternative rubrics are likely to be needed. Any such subject-specific rubric would have to be, as the participant points out, “negotiated first with all the collaborating stakeholders”. Future research should investigate whether rubric templates could be developed for specific writing genres, specifically within discipline-specific contexts.

Two participants proposed that the rubric be adapted to also enable formative assessment in addition to summative assessment. One of these participants indicated that the “draft method” could be used in impact measurement – thus, that the impact of an intervention

on students' writing be assessed through several drafts. One challenge with regard to this suggestion is that extensive resources (for example time and qualified markers) would be necessary for marking several drafts, especially in the South African context where classes are frequently large, and lecturers tend to have a heavy teaching load. Yet, there might be contexts where the required resources would be available, or where small class sizes allow for additional marking. In such contexts, this would seem to be a valuable addition to the evaluation design, and the writing rubric (Appendix I) was adapted by adding more details so as to facilitate formative assessment.

Another suggestion with regard to using writing assignments is that group assignments be used. This seems to be a sensible suggestion, considering how little individual writing seems to be done by students in their first-year, especially where large student numbers are a reality. Even the second semester of the academic literacy intervention which was examined in Chapter 6 required group assignments rather than individual assignments for the end-of-year extended assignment. This would have to be thoroughly planned and managed though (for an example of best practices in this regard, see Michaelsen & Sweet, 2011), so that both pre- and post-assignments are similar, and so that all students contribute equally to the end product. Furthermore, the group members for both pre- and post-assignments would ideally have to stay consistent so that various groups' assignments could be compared with each other. Without proper planning and management, the reliability of this instrument would suffer; appropriate checks and balances throughout the group work process would be essential.

Finally, one participant suggested that an assessment be included which assesses postgraduate students' "ability to identify arguments in the texts they read for research projects". Chapter 4 suggests the TALL as an appropriate instrument for assessing undergraduate students' academic literacy levels. However, as the proposed evaluation design should be adaptable for both undergraduate and postgraduate students, it would be sensible to propose an academic literacy test aimed at postgraduate students, which includes abilities such as the one that this participant raised. The Test of Academic Literacy for Postgraduate Students (TALPS) (Rambiritch, 2013, 2012; Butler, 2009) is a widely-used, validated test aimed specifically at postgraduate students, and is suitable for assessing the academic literacy abilities which are required after students' undergraduate studies.

Additional instruments

The second question asked was “Are there any instruments that you would add to this evaluation design?”. Although the majority of participants felt that the range of tools was sufficient to assess the impact of their respective interventions, some valuable suggestions were provided for possible additional tools.

One participant indicated that s/he would have included listening and speaking abilities in the generic or discipline-specific tests. The participant specifically focused on vocabulary acquisition: whether the student would be able to understand a word when hearing it in context, or whether the student would be able to use new words correctly when speaking, for example in a class conversation or a debate where the student does not have much time to think about the usage of the word. Though it might be possible to integrate such abilities in discipline-specific tests (specifically if these are developed for a specific course which has outcomes focusing on listening and speaking), it is not feasible to do this in some of the more widely-used validated and reliable generic academic literacy tests, due to the resources (such as test time, marking time, experienced assessors and concomitant financial resources) this would require. Still, as there seem to be interventions which do focus on these outcomes, it would seem useful to add tools to the evaluation design which could be used to assess these abilities.

A possible rubric for assessing oral presentations is attached as Appendix J. This rubric is taken from a course called Language and Study Skills which is presented at the University of Pretoria (Fouché & Immelman, 2015), and is aimed at assessing formal oral presentations for first-year students. It might thus not be applicable to all contexts. Further research would be necessary to determine whether it could be used across a wider range of contexts. For academic literacy interventions that have listening abilities as a major outcome, it would be valuable to consider the research done by Marais (2009) and Marais and Van Dyk (2010). Marais (2009:119-127) proposes a listening test (the Academic Listening Test [ALT]) aimed at first-year students which was shown to have construct, content, and face validity, in addition to being shown to be reliable. The revised evaluation design includes the possibility for assessing students’ speaking and listening abilities.

A further suggestion was that qualitative student data be used. Although the proposed student questionnaire (Appendix D) does contain some open-ended questions⁴⁹, several other possibilities exist to collect rich and detailed information “so as to obtain a more holistic picture of development as well as the nature of students' challenges with the literacy practices of their discipline”. Options which were suggested by participants include retrospective self-observation blogs as well as student interviews. To this, one could add other options such as focus group interviews. Although the originally proposed evaluation design did include “qualitative data from primary stakeholders”, the discussion surrounding this in Chapter 4 was focused on academic literacy lecturers and content-subject lecturers. In the revised evaluation design, a distinction is made between qualitative feedback from (academic literacy and content-subject) lecturers and from students.

A final instrument that was suggested was benchmarking with other institutions. This can indeed be useful in determining whether students at various institutions are at similar levels with regard to academic literacy, whether institutions focus on similar outcomes in their academic literacy interventions, and whether academic interventions at some institutions seem to be more effective than those at others. At the very least, such benchmarking could assist in identifying gaps and weaknesses (or strengths) in an institution's own academic literacy intervention(s). At best, this might lead to collaboration and shared expertise, which would in turn benefit the interventions involved.

Additional aspects to keep in mind

The third question posed to academic literacy specialists was “Are there any other aspects that you believe the proposed evaluation design should take into consideration, given the context of the academic literacy intervention you are responsible for?”. Several constructive responses were received for this question.

Three participants commented on the duration of the academic literacy intervention or the duration of the assessment of its impact (cf. Bamberger *et al.*, 2012:613). As one

⁴⁹ The revised need-press questionnaire on academic literacy abilities (Appendix F) contains an additional open-ended question so as to facilitate richer data-collection. This question reads: “How could the academic literacy course be adapted to be of more value to you in your other subjects”.

participant noted, instruments “would have to be longitudinal in nature, tracking the progress and development of academic literacy skills over a period of time”. Another participant pointed out that such an impact assessment should not be “once off” – it is necessary to regularly determine the impact of academic literacy interventions so as to ensure that they still fulfil their outcomes optimally. One participant commented on the fact that the “length of the intervention has an impact on the extent to which students can be said to have successfully learnt the literacies of their learning contexts. As has been confirmed by Van Dyk *et al.* (2011a:497) and in Section 6.6 of the current study, the impact of academic literacy interventions is often only noticeable after a longer intervention – in their case, little impact could be seen after a semester, but after a year the impact of the intervention was much clearer. The participant echoes this when stating that the “courses we teach are offered for a semester only due to timetabling issues, [and] resources in the form of teaching staff who are available to offer more extended interventions. Student feedback has indicated that they would have benefitted from a whole year rather than a semester”. The duration of the intervention must thus be kept in mind when the impact thereof is reported upon. One participant touched on a similar aspect, namely the structure of the intervention, for example whether it is a foundational or a mainstream course. Factors such as time allocated on the timetable and class size would also influence the structure of the course. As the participant points out, it is necessary to determine how the structure of the course “impacts on student experiences as they develop as academic writers”. One way of doing this would be to include more qualitative student data in an impact assessment, as was suggested as an additional instrument earlier in this section. In addition to that though, any report on the impact of an academic literacy intervention would have to be accompanied by a thorough description of the structure of the course to put the findings into perspective.

Other aspects that were raised by some participants were the background as well as the academic strength of the students. With regard to background, it would be important to consider students’ home languages and their writing and reading proficiency in these languages. These factors might potentially influence students’ performance in an academic literacy intervention (see, for example, Yamashita, 2002; Cummins, 2001)⁵⁰. For instance, English second language students might not have acquired appropriate

⁵⁰ Also consider Coetzee-van Rooy's (2010) critique of applying Cummins' model to the South African context.

reading and writing proficiency in their first languages, and might thus be unable to transfer certain abilities from one language to the other; this could of course also be true of students who only speak English. The needs of such students might differ from those of students who are fully proficient in another language. Another participant stated that a “strong student can easily pick up a skill, whether the intervention is good or bad. It is the poorer student, however, whose improvement is proof of a good intervention”. This claim could be assessed by categorising students into various quartiles based on pre-intervention data (for example Grade 12 or pre-test marks). These data could also indicate whether the intervention has the same impact on students at various levels, and whether it might be necessary to adapt interventions for students of different academic levels.

Two participants stressed the need for evaluating discipline-specific courses. One participant notes that an evaluation design for academic literacy interventions “should consider the nuanced nature of different disciplines within a university, and how the AL [academic literacy] intervention needs to be constantly re-modelled in order to accommodate these differences”. Another states that “we work with students from their disciplines, so [an evaluation design] needs to evaluate what has been provided in that particular discipline”. This study attempts to address these concerns by including options for discipline-specific and collaborative courses in the evaluation design. Developing specific academic literacy tests and questionnaires for each specific discipline falls outside of the scope of the current study, but this is certainly something that should be researched in future. Such research would have to consider how many resources would be required to undertake this task, whether such discipline-specific academic literacy tests and questionnaires could be used across a variety of universities or whether each would have to be contextualised in the setting of a specific university, and whether such an undertaking would ultimately be feasible and sustainable. Until such research has been conducted, the evaluator would either have to adapt the instruments proposed in this study to be appropriate for discipline-specific interventions, or use alternative instruments that are designed for such specific disciplines. A further comment that relates to the type of academic literacy intervention being assessed was made by one participant who suggested that a definition of academic literacy “could include mathematical and graphic literacy for academic purposes”. Though this is a valid point, and warrants further research, it does fall outside the scope of the current study which accepts Van

Dyk and Van de Poel's (2013:56) view of academic literacy as "being able to use, manipulate, and control language and cognitive abilities for specific purposes and in specific contexts". Based on this definition, the student and lecturer questionnaires were adapted to remove the questions they had originally contained on computer literacy.

Another aspect that should be kept in mind is the planning stage of the evaluation. Lynch (1996:3) argues that when evaluating a programme, the first step would be identifying and consulting with relevant stakeholders. One participant echoes this notion: "I would first determine from the stakeholders what they will need, e.g. what does the ethics committee require? What does the disciplinary supervisor require? What about the funders?". Answering these questions would be pivotal in selecting suitable evaluation instruments. A related comment made by a participant was that the academic literacy specialist should work with subject lecturers, and that content-subject lecturers must be involved in the evaluation process; this would be specifically relevant in the case of discipline-specific academic literacy interventions. Yet another comment suggested that students "be given the platform to tell the lecturers (...) how they would like to be assessed". Both of these suggestions could be fruitful, as content-subject lecturers and students are likely to be able to give valuable input as to which instruments might measure the transfer of academic literacy abilities the best.

Three final comments were made that are particularly valuable. Firstly, it was pointed out that "students' responses to questionnaires etc. would obviously be subjective and likely to be influenced by current academic performance so the interviews / questionnaires would need to be considered alongside their writing". This highlights the need for triangulation in an evaluation design, and the necessity of including a variety of instruments. A second comment highlighted the need for the evaluation design to be flexible: "It can never be a one size fits all. My use of the design and approach of academic literacy is that you have to be sensitive to context, to social practices and these are always dynamic, changing and negotiated. In other words, evaluation has to also be flexible and depends on space, time, and purpose". A final suggestion was that the evaluation process itself should ultimately be evaluated. Section 7.5 attempted to do just that, and it was indeed a fruitful exercise, as I would approach future impact assessments slightly differently, based on the challenges experienced in this evaluation process.

The following section proposes a revised evaluation design based on personal experience of implementing some of the instruments proposed in the initial design (see Section 7.5), as well as feedback from academic literacy specialists across South Africa (Section 7.6).

7.7 Revised evaluation design

The revised evaluation design (Figure 7.1) includes some added instruments, some revised instruments, as well as general guidelines for the implementation of the design.

Instruments that were added to the initially proposed evaluation design (Figure 4.2, Section 4.5) are as follows:

- assessing students' academic speaking abilities (Appendix J);
- assessing students' listening abilities;
- using qualitative student data; and
- benchmarking with other institutions.

Instruments that were revised are:

- the student need-press questionnaire on academic literacy abilities (Appendix F);
- the lecturer questionnaire on academic literacy abilities (Appendix H);
- the writing rubric (Appendix I);
- the individual extended written assignment (to also allow for group assignments or assignments from students' content subjects).

The initially proposed evaluation design was further adapted in that the categories of “recommended instruments” and “additional optional instruments” were removed. As the quantitative responses in Section 7.6.1 indicate, those instruments that had originally been classified as “recommended” are not necessarily seen as more applicable to academic literacy specialists responsible for specific categories of interventions than any of the other instruments. Note that each of the four categories of interventions contains the same instruments. However, instruments are ordered, from top to bottom for each respective category (indicated by different colours⁵¹), by how relevant each instrument

⁵¹ Instruments that aim to measure an improvement in students' academic literacy levels are indicated in blue. Instruments that attempt to determine whether these academic literacy abilities are needed in, and transferred to, students' content subjects are indicated in green. Instruments which do not clearly fall into either of these categories, but might still shed some light onto the impact of the intervention in question, are indicated in orange.

might be based on feedback from academic literacy specialists (see Section 7.6.1). Note that due to the relatively small number of participants, this order is merely a rough estimation of usefulness, and the order of relevance of instruments will without doubt be different for each individual intervention within these four categories. It is also important to note that the colour-coded categories can overlap in some cases. For instance, a subject-specific academic literacy test, or a writing assignment completed for another subject would indicate whether there was an improvement in students' academic literacy levels (and would thus be categorised under the blue-coded research instruments), but could also provide information regarding the transfer of such abilities (thus, it might also be applicable under the green-coded categories) (see Figure 7.1).

The recommendation that at least three instruments be used in cases where control groups are not available, and at least two instruments be used where control groups are available, remains unchanged (of course, the more instruments that are used, the more complete the picture that will emerge). So does the recommendation of using at least one instrument that indicates whether there was an improvement in students' academic literacy levels (indicated in blue), and one instrument that indicates transfer of these abilities or the need for them in students' content subjects (indicated in green). Instruments that do not clearly fall into either of these categories, but that might still be useful, are indicated in orange. By using a number of instruments from various categories, the researcher strengthens triangulation by both source and method (see Section 4.5). To further strengthen triangulation, it is suggested that a combination of qualitative and quantitative instruments be used wherever possible.

One of the key characteristics of this evaluation design for academic literacy interventions is its flexibility. Evaluators must be able to choose instruments that are applicable to their respective contexts, and as can be seen from the data in Section 7.6.1, even interventions in the same category (for example generic academic literacy courses) might need to use very different instruments than other interventions in the same category. The greater the variety of instruments used (for example qualitative and quantitative instruments, and instruments that measure an improvement in academic literacy abilities as well as those that determine the necessity of abilities in students' content subjects and the possible transfer of these abilities), and the larger the number of

GENERIC ACADEMIC LITERACY INTERVENTIONS	SUBJECT-SPECIFIC ACADEMIC LITERACY INTERVENTIONS	COLLABORATIVE ACADEMIC LITERACY INTERVENTIONS	LIMITED PURPOSE INTERVENTIONS <small>(e.g. writing centres or reading programmes)</small>
Generic extended writing assignment (assessed by means of rubric)	Subject-specific extended writing assignment (assessed by means of rubric)	Subject-specific extended writing assignment (assessed by means of rubric)	Subject-specific extended writing assignment (assessed by means of rubric)
Generic / subject-specific extended writing assignment (assessed quantitatively)	Subject-specific academic literacy test	Subject-specific academic literacy test	Generic / subject-specific extended writing assignment (assessed quantitatively)
Generic academic literacy test	Generic / subject-specific extended writing assignment (assessed quantitatively)	Generic / subject-specific extended writing assignment (assessed quantitatively)	Subject-specific academic literacy test
Subject-specific extended writing assignment (assessed by means of rubric)	Generic academic literacy test	Generic academic literacy test	Generic academic literacy test
Subject-specific academic literacy test	Generic extended writing assignment (assessed by means of rubric)	Generic extended writing assignment (assessed by means of rubric)	Generic extended writing assignment (assessed by means of rubric)
Listening test	Listening test	Listening test	Listening test
Assessing students' oral abilities			
Qualitative feedback from students	Qualitative feedback from academic literacy and content-subject lecturers	Qualitative feedback from academic literacy and content-subject lecturers	Qualitative feedback from academic literacy and content-subject lecturers
Qualitative feedback from academic literacy and content-subject lecturers	Qualitative feedback from students	Qualitative feedback from students	Qualitative feedback from students
Content analysis of study material	Student questionnaire	Student questionnaire	Student questionnaire
Student questionnaire	Content analysis of study material	Content analysis of study material	Content analysis of study material
Lecturer questionnaire	Lecturer questionnaire	Lecturer questionnaire	Lecturer questionnaire
Correlating academic literacy results with other variables			
Benchmarking with other institutions			

(Cont.) Guidelines for using the evaluation design for academic literacy interventions

- At least one instrument should indicate whether there was an improvement in students' academic literacy levels (blue-coloured instruments) and at least one instrument should indicate whether the abilities acquired are needed, or were transferred to, students' other subjects (green-coloured instruments). Instruments that do not clearly fall into either of these categories, but that might still be useful, are indicated in orange.
- Instruments at the top of each colour category are generally considered more relevant for each category of academic literacy interventions, while instruments at the bottom of each colour category are generally considered less relevant. This will, however, depend on each specific evaluation context.
- Where possible, a combination of qualitative and quantitative instruments should be used.
- Interventions where control groups are available should make use of at least two of the instruments above, and interventions where control groups are not available should make use of at least three of the instruments above.
- Extended subject-specific assignments could be used in addition to, or instead of, writing assignments given by the academic literacy intervention.
- Group work assignments could be used in addition to, or instead of, individual assignments, provided that variables such as group members stay consistent between pre- and post-assessments.
- Questionnaires should preferably be distributed in person, or an incentive should be linked to completing online questionnaires, so as to ensure an acceptable return rate.
- Should a generic academic literacy test be decided upon, it should be aimed at the appropriate level (e.g. undergraduate or postgraduate students).
- The impact assessment should contain a thorough description of the structure of the course (for example duration, contact time and class sizes) so as to contextualise findings.
- The impact assessment could consider students' background (for example Grade 12 marks) as well as their academic literacy levels before the start of the interventions. By analysing data in various quartiles based on pre-evaluation data, valuable information might be obtained.
- Students as well as content-subject lecturers could be consulted with to determine which evaluation instruments would best measure the impact of the academic literacy intervention in question.

Figure 7.1: Evaluation design for academic literacy interventions

instruments used, the more valid the deductions regarding impact and its causality that can be made based on the triangulated data obtained from these instruments.

Even if a range of instruments is used and data are triangulated, this design still has limitations. In the course that was evaluated in Chapters 5 and 6, it seemed clear that students improved significantly in several academic literacy abilities, that students believe that the areas they had shown improvement in were necessary for success in their other subjects, and that these abilities were generally sufficiently addressed in their academic literacy course. It would also seem as though most of the academic literacy areas in which students had improved showed a significant improvement when content-subject assignments were assessed at the end of the year. However, there is still no definite empirical proof that the academic literacy course itself was responsible for such improvement, despite strong indications from the literature that academic literacy abilities are unlikely to improve without a specific intervention (see, for example, De Graaff & Housen, 2009:729; Holder *et al.*, 1999:27; Farnill & Hayes, 1996:264; Rosenthal, 1996:24; Thompson, 1990:101; Eskey, 1983:322). However, based on the rich and varied evidence provided, it would seem likely that the academic literacy course did have a meaningful impact.

Ultimately, it remains the researcher's responsibility to ensure that the most comprehensive combination of instruments that is feasible is used so as to strengthen inferences regarding the impact of the intervention as far as is possible, whilst acknowledging any limitations that might still remain in the research design. If data are responsibly triangulated by both source and method (cf. Lynch, 1996:59; Denzin, 1989:237; also see Section 2.5.3), much value will be added to the field of academic literacy by ensuring effective, tried and tested interventions that have the highest impact possible for their specific contexts.

7.8 Conclusion

According to Kiely (2009:99), “[p]rogramme evaluation is a form of enquiry which describes the achievements of a given programme, provides explanations for these, and sets out ways in which further development might be realized”. The instruments

proposed in this chapter could all contribute towards this aim, even more so if several are used in conjunction with each other and data are triangulated.

Kiely (2009:99) further calls for research that ensures that “the research-type knowledge-building enterprise and the ongoing quality management processes are mutually informing, and that programme evaluation becomes a socially-situated cycle of enquiry, dialogue, and action”. Cole *et al.* (2005:12) echo this sentiment, and point out that this process is characterised by a “generate-test cycle”. The instruments proposed for the evaluation design discussed in this study should therefore be seen as artefacts that can, and must be adapted over time. The self-reflection and feedback from specialists discussed in this chapter formed part of this process, but it is a process that must be continued in future research for various contexts.

Cole *et al.* (2005:12) point out though that this “generate-test cycle” is constrained by available resources as well as technology – this might be why so few studies have attempted to comprehensively measure impact, as can be seen from the review in Chapter 3. Where such constraints exist, it might be wise to use existing instruments, albeit imperfect, and to adapt these where possible. It is certainly preferable to follow that route rather than to avoid measuring the impact of an intervention altogether since no perfect instruments exist with which to do this.

The current chapter has aimed at indicating how research, in particular self-reflection as well as feedback from specialists in the field, can assist in solving either current or anticipated problems of practitioners (cf. Cole *et al.*, 2005:2). By continuously integrating such reflection into the research process, the artefact’s (in the case of this study, an evaluation design and accompanying suggested instruments) utility, quality and efficacy (cf. Cole *et al.*, 2005:3-4; Hevner *et al.*, 2004:77) will continuously be improved upon, and the field as a whole will benefit.

The following chapter concludes this study. It starts by summarising each chapter. Thereafter, conclusions are made based on the evidence provided in this study, and the contribution of the study is discussed. Finally, the limitations of the study are indicated, and recommendations for further research are made.

CHAPTER 8

CONCLUSION AND RECOMMENDATIONS

8.1 Introduction

Due to the dearth of studies assessing the impact of academic literacy interventions, there is consensus in the literature that this is a field that needs to be reported on more (Mhlongo, 2014:47; Terraschke & Wahid, 2011:174; Carstens & Fletcher, 2009b:319; Storch & Tapper, 2009:208; Holder *et al.*, 1999:20), so as to be accountable to stakeholders and to work towards improving academic literacy interventions (De Vos *et al.*, 2011:449; Hatry & Newcomer, 2010:678). As can be seen from the literature reviewed in Chapter 3, those studies that have attempted to assess the impact of academic literacy interventions have often either used inadequate instruments or an insufficient number and combination of instruments (i.e., the scope of these studies was very narrow). The aim of the current study was to address this gap in the research by suggesting a multi-faceted, verified and validated evaluation design that could be used to effectively and comprehensively assess a variety of academic literacy interventions.

The evaluation design is grounded in a postpositivist paradigm (see Section 2.5.1) in that, although it focuses strongly on empirical evidence for determining an improvement of students' academic literacy abilities, it also allows for rich qualitative data to be collected to help explain certain phenomena, and give insights that would not be possible with purely empirical data. As Lynch (2003:29) points out, the greatest strength of using a postpositivist paradigm is that using various types of evidence serves as a valuable validity check.

8.2 Summary of the study

Chapter 1 contextualised the study in the South African higher education setting. It indicated why there is a need for comprehensive impact studies of academic literacy interventions within this context. It then elaborated on the research design used for the current study. Thereafter, the research questions, objectives, instruments and parameters of the study were described. An overview of the study was provided, followed by a

discussion of how ethical considerations were addressed. It concluded by indicating what the contribution of this study would be.

An in-depth review of the literature was done in Chapter 2. The chapter elaborated on concepts which could not be expanded on in the various articles (Chapters 3 to 7) due to the nature of academic articles. This chapter started by defining various key concepts. It then examined students' underpreparedness for higher education studies. Finally, the concepts of "academic literacy" and "programme evaluation" were thoroughly examined.

Considering the increasingly resource-scarce higher education environment, it is becoming crucial for those who are responsible for academic literacy interventions to provide evidence of their impact on student success. Taking this into consideration, Chapter 3 aimed to provide a broad overview and critique of studies conducted thus far that attempted to assess the impact of various academic literacy (or related) interventions. This chapter proceeded by identifying instruments that are commonly used when assessing the impact of these interventions. From the literature surveyed, it would seem that there are two broad aspects that are considered when evaluating impact, namely students' improved academic literacy levels between the onset and the completion of the course, and the extent to which these acquired academic literacy abilities are transferred to students' other subjects.

Chapter 4 aimed to provide an overview of instruments that are commonly used to assess impact, and to discuss guidelines regarding the use of these instruments, their strengths and their weaknesses. The instruments were divided into two broad categories, namely those that measure the observable improvement in students' academic literacy abilities between the onset and the completion of an intervention, and those that measure the extent to which these abilities are necessary and applied in students' content subjects. A conceptual evaluation design was then proposed that could be used in evaluating the impact of a range of academic literacy interventions.

In order to verify the evaluation design for academic literacy interventions proposed in Chapter 4, Chapter 5 investigated whether there was an improvement in students' academic literacy levels between the onset and completion of the first semester of an academic literacy course at a South African university. This was done by using a

combination of instruments selected from this proposed evaluation design. A pre-test / post-test design was used, where, firstly, students' results in a validated and reliable generic academic literacy test were considered. Secondly, students' writing abilities were assessed by means of two instruments: a rubric and quantitative measures. Finally, students' academic literacy marks were correlated to other variables, and interpreted within the context of the study, to give additional insight into the impact of the academic literacy course. Findings indicate that students showed an improvement across a wide array of academic literacy abilities, in particular their ability to use source material in their writing assignments, and their usage of a wider range of academic vocabulary. However, students did not appear to be more proficient in terms of their visual literacy after having completed the course; further, they were slightly more likely to commit plagiarism (i.e. copying word for word, even though sources were acknowledged to a greater extent) at the end of the course than at the beginning.

In Chapter 6, the impact of the second module of the course that was the subject of Chapter 5, as well as the impact of the course in its entirety, were examined. Once again, a combination of instruments suggested in the evaluation design proposed in Chapter 4 was utilised. Firstly, a pre-test / post-test design was used with three instruments, namely a valid and reliable academic literacy test, a writing assignment which was assessed by means of a rubric, and a quantitative assessment of the same writing assessment. Secondly, student questionnaires were used to determine whether the abilities addressed in the course are those relevant to students' content subjects, and to which extent relevant abilities are addressed in the academic literacy course. Finally, students' academic literacy results were correlated with their marks in other subjects. Findings indicate that the academic literacy course in question effectively and comprehensively (both statistically significantly, and with moderate to large effect sizes) addressed a wide range of academic literacy abilities. Student feedback corroborated these findings. As was the case in Chapter 5, students' ability to avoid plagiarism would not seem to have improved over the course of the year. When considering students' range of academic vocabulary in a writing assignment from a content subject, it seems as though the increased range of academic vocabulary seen in the mid-year writing assignment was not transferred to students' content subjects. On the whole, however, this course seems to have effectively and comprehensively improved students' academic literacy levels. As for the applicability of the proposed evaluation design, this design was generally considered to

be appropriate for assessing the impact of this academic literacy intervention. Possible improvements to the design were suggested in Chapter 7.

Reflecting on the research process through self-reflection as well as feedback from stakeholders is a vital step in responsible scholarship so as to ensure that research contributes to both theory and practice (cf. Cole *et al.*, 2005:2, 16). Chapter 7 considered how such reflection could contribute to the improvement of a research artefact by considering an impact study (Chapters 5 and 6) in which one such artefact, namely an evaluation design for academic literacy interventions, was implemented. This chapter reported on the challenges experienced in the implementation of the originally proposed evaluation design in this impact study. In addition, the evaluation design was sent to academic literacy specialists at all South African universities to determine whether the design would be appropriate for their respective contexts (thus, to validate the design), and whether they would recommend improvements to make the design more applicable to a variety of contexts. After having considered 1) the researcher's reflection on the challenges posed in the implementation of the evaluation design, and 2) the academic literacy specialists' feedback regarding possible improvements to the evaluation design, a revised evaluation design was proposed in Section 7.7.

8.3 Conclusions and contribution of the study

The study set out to address four objectives (that correspond to the research questions posed in Chapter 1). The first objective, namely determining which methods have been used in the literature thus far to assess the impact of academic literacy interventions, was addressed in Chapter 3. The second objective of developing an evaluation design (based on trends and suggestions in the literature) to assess the impact of an academic literacy intervention on student success, was addressed in Chapter 4; this evaluation design was consequently refined in Chapter 7. The third objective was to verify the proposed evaluation design to assess the impact of a specific academic literacy intervention; the outcomes of this stage of the study were reported on in Chapters 5 and 6. The fourth objective, namely determining to which degree the evaluation design could be generalised to other settings, so as to assess the impact of a variety of academic literacy interventions in a variety of contexts, was addressed in Chapter 7.

The primary contribution of the study has been the development of a comprehensive and flexible evaluation design for assessing the impact of academic literacy interventions. The evaluation design, which was verified and validated in this study, should be able to assist evaluators in assessing the impact of a wide range of academic literacy interventions in the South African context. Once an evaluator has determined the extent of the intervention's impact, the intervention's approach and/or content could be adjusted accordingly, thus resulting in a more effective, applicable and useful intervention. An indirect consequence of more responsibly designed impact studies would be that course developers at various institutions would be able to learn from each other's failures and successes. This would ultimately strengthen the field of academic literacy studies as a whole.

While the evaluation design itself contains specific types of instruments that could be effectively used in assessing the impact of academic literacy interventions, it also proposes specific examples of instruments which are available in, and applicable to, the South African context. These specific instruments (many of which are included in the appendices of this study) can assist lecturers who are interested in evaluating the impact of the interventions they are responsible for, but who do not have time to develop alternative, valid and reliable instruments. Prospective evaluators are encouraged to use and adapt these instruments where necessary, so as to make them applicable to each intervention's specific context and needs (as has been done in Chapters 5 and 6 of this study).

Striving towards more meaningful interventions becomes an ethical obligation in a country where so few students are sufficiently prepared for higher education studies, yet where resources in the higher education sector are becoming increasingly scarce. I believe that the evaluation design put forward in this study can contribute towards enabling more academic literacy specialists to critically evaluate, and subsequently improve, their respective academic literacy interventions.

8.4 Limitations

It is appropriate to revisit the parameters set in Chapter 1. Firstly, the evaluation design was developed by keeping the South African context in mind, and was refined by

consulting a wide variety of South African academic literacy specialists. Although it is possible that it might be effective in other contexts as well, further research would be needed to establish that. Based on the research conducted for this study, however, no claims beyond the fact that the evaluation design is applicable to and effective for the South African context can be made. Secondly, the proposed evaluation design focuses on the short-term impact of academic literacy interventions, though several instruments (for example the academic literacy tests and student questionnaires) could be used for more longitudinal studies. A third limitation is that the evaluation design was only implemented by assessing an academic literacy course at one South African university. Implementing it in more contexts in future would doubtlessly provide valuable additional information that might be used to refine the evaluation design and the instruments proposed in this design. A fourth limitation is that the evaluation design only focused on academic literacy interventions as conceptualised in Sections 2.4.1 and 2.4.2. It cannot be used to evaluate other types of literacies that are sometimes included under the term “academic literacies”, for example mathematical or information literacy.

A limitation of the proposed evaluation design (and not necessarily of the study, as is the case with the previous four limitations), is that even when using a variety of instruments (as was the case with the implementation of the evaluation design in Chapters 5 and 6), if a traditional experimental design cannot be followed, it is not possible to conclusively claim that the academic literacy intervention was responsible for students’ improved academic literacy abilities. However, several studies have indicated that academic literacy abilities are not likely to develop without explicit instruction (see Section 2.4.3). Keeping this in mind, combined with triangulated empirical research from a variety of sources and using a quasi-experimental pre- and post-test design, strong inferences could be drawn regarding the likely impact that an academic literacy intervention might have had, and the nature of this impact. Such a quasi-experimental design might not be able to determine causality with the same level of certainty as would be the case with a traditional experimental design, but responsibly using multiple instruments can assist in making claims regarding the likelihood of causality – in the case of this study, that the academic literacy intervention positively impacted students’ academic literacy levels.

Beretta (1992:9) warns that various extraneous variables, for example students dropping out, would complicate an evaluation study that was conducted over a long period. This

was also true for the implementation of the proposed evaluation design over the period of a year (see Chapters 5 and 6). It was impossible to obtain post-tests from the random sample identified at the beginning of the year due to such extraneous variables, and the study constantly had to compensate for these complications. Had the evaluation study been conducted over an even longer period, these complications would have been much more pronounced and difficult to overcome. Although large enough samples were used to make statistical analyses feasible, larger samples would have benefited the study by making conclusions more valid and reliable. In the context of this study, that would, however, not have been possible. On the other hand, too large a sample might have caused even more complications (cf. Beretta, 1992:9). Future researchers should try to follow Beretta's (1992:9) guidelines by selecting enough participants to ensure valid and reliable statistical conclusions and saturation in qualitative research, but to avoid excessive sample sizes so as to keep the study contained.

8.5 Recommendations for further research

Possibly the most prominent recommendation for future research is that the evaluation design proposed in this study be implemented in a variety of South African contexts. Even though academic literacy specialists who work in a wide variety of contexts have evaluated the appropriateness of this design, practically implementing it across a wide variety of contexts would likely provide additional ways in which it can be refined. It is further recommended that future researchers continuously reflect on new challenges that arise (as was done in Section 7.5), and to report on these challenges. The critical consciousness that will develop from continuously reflecting on new challenges will help researchers to keep refining theory, which can ultimately be implemented in practice.

Another suggestion for further research is that this evaluation design be adapted for contexts outside of South Africa. This should be done by first consulting with academic literacy specialists in these contexts. Thereafter, the adapted design should be implemented. I believe that many of the instruments suggested in this evaluation design could be appropriate for settings outside of South Africa, but this would have to be verified and validated through further research.

From the feedback received from course coordinators in Chapter 7, two themes arose that warrant further research. Firstly, there would seem to be a need for evaluation instruments that are directed at various disciplines, as generic instruments are not always ideal for assessing discipline-specific interventions. Future researchers could determine whether it is possible to create such instruments that can be used in specific disciplines, but that would also be transferable to a variety of institutions. Such an arsenal of discipline-specific evaluation instruments would certainly enrich the current evaluation design.

Secondly, there seems to be a need for more qualitative evaluation instruments. Due to the nature of qualitative research, it might be that tailor-made qualitative instruments would need to be developed for each unique evaluation context. However, it would be worthwhile to investigate whether dependable, credible and transferable templates which are aimed at various primary stakeholders could be developed.

As far as the specific testing instruments proposed in this evaluation design are concerned, I would recommend that the same test be used as both pre- and post-test. This would enable the researcher to make stronger conclusions on the improvement of students' academic literacy levels, and would prevent situations such as those reported on in Chapters 5 and 6, where it becomes impossible to compare various sections within a test. Where it is not possible to use the same test as pre- and post-test, tests should be piloted to make sure that these tests and their sub-sections can be considered theoretically and practically equivalent. This will greatly assist the research in drawing valid conclusions based on the results of such tests.

It is also important that future researchers ensure that communication mechanisms to various stakeholders (such as lecturers and students) are put in place. This will assist in eliminating misunderstandings regarding various requirements that the research might have for the seamless implementation of the evaluation design.

Another recommendation to future researchers is that studies which assess the impact of academic literacy interventions be conducted either by outsiders who can be on site during the entire evaluation period, or by insiders (see Sections 2.5.2.1 and 7.5.1). Although outsiders might be more objective about the outcomes of an impact evaluation,

having outsiders on site for the duration of such an evaluation would involve additional resources, and might not guarantee objectivity. Being on site during the evaluation process would, however, help in eliminating many of the challenges experienced in the study (see Section 7.5).

8.6 Conclusion

Universities are not able to equally address all the barriers to student success (for example students' secondary education, their socio-economic circumstances or inherent factors). However, where they are able to address certain barriers by, for example, providing academic literacy interventions, they have the responsibility to ensure that such interventions are designed and constantly improved so as to have the greatest possible impact on student success. This can only be done once the impact of these interventions (including their strengths and weaknesses) has been determined.

I conclude with Van Dyk and Van de Poel's (2013) statement that "[i]n the end it will come down to our, and our institutions' maturity (readiness) to design and apply solutions to the benefit of all involved, or put differently, we need solutions that have impact". Even though it is not easy to make definitive conclusions on the impact of academic literacy courses, I agree with these authors that "this should not prevent us from consistently and continuously investigating, enquiring and critiquing our own ventures" (Van Dyk & Van de Poel, 2013:60).

BIBLIOGRAPHY

- Alderson, J.C. 1992. Guidelines for the evaluation of language education. In: Alderson, C.J. & Beretta, A. (eds). *Evaluating second language education*. Cambridge: Cambridge University Press.
- Alderson, J.C. & Beretta, A. 1992. *Evaluating second language education*. Cambridge: Cambridge University Press.
- Alderson, J.C. & Scott, M. 1992. Insiders, outsiders and participatory evaluation. In: Alderson, C.J. & Beretta, A. (eds). *Evaluating second language education*. Cambridge: Cambridge University Press.
- Amos, T. & Quinn, L. 1997. Management education and training: The role of integrated language development. *South African Journal of Higher Education*, 11(1):186-191.
- Archer, A. 2006. A multimodal approach to academic 'literacies': Problematizing the visual/verbal divide. *Language and Education*, 20(6):449-462.
- Archer, A. 2008. Investigating the effect of Writing Centre interventions on student writing. *South African Journal of Higher Education*, 22(2):248-264.
- Babbie, E. 2005. *The basics of social research, 3rd ed.* Belmont: Wadsworth.
- Bachman, L.F. & Palmer, A.S. 1996. *Language testing in practice: Designing and developing useful language tests*. Oxford: Oxford University Press.
- Bachman, L.F. & Palmer, A.S. 2010. *Language assessment in practice*. Oxford: Oxford University Press.
- Baer, J., Cook, A. & Baldi, S. 2006. *The literacy of America's college students*. Washington DC: American Institutes for Research.
- Baik, C. & Greig, J. 2009. Improving the academic outcomes of undergraduate ESL students: The case for discipline-based academic skills programs. *Higher Education Research and Development*, 28(4):401-416.
- Bak, N. 2004. *Completing your thesis: A practical guide*. Pretoria: Van Schaik.
- Bamberger, M., Rugh, J. & Mabry, L. 2012. *Real world evaluation: Working under budget, time, data and political constraints, 2nd ed.* Thousand Oaks: Sage.
- Bardovi-Harlig, K. & Bofman, T. 1989. Attainment of syntactic and morphological accuracy by advanced language learners. *Studies in Second Language Acquisition*, 11(1):17-34.
- Basson, D. 2006. *Die verwantskap tussen enkele veranderlikes en akademiese sukses op universiteit*. Ongepubliseerde magisterverhandeling. Stellenbosch: Universiteit van Stellenbosch.
- Baum, F., Macdougall, C. & Smith, D. 2006. Participatory action research. *Journal of Epidemiology and Community Health*, 60(10):854.
- Beretta, A. 1986. Program-fair language teaching evaluation. *TESOL Quarterly*, 20(3):431-444.
- Beretta, A. 1992. Evaluation of language education: An overview. In: Alderson, C.J. & Beretta, A. (eds). *Evaluating second language education*. Cambridge: Cambridge University Press.
- Berkenkotter, C., Huckin, T.N. & Ackerman, J. 1991. Social context and socially constructed texts: The initiation of a graduate student into a writing research community. In: Bazerman, C. & Paradis, J. (eds). *Textual dynamics of the professions: Historical and contemporary studies of writing in professional communities*. Madison: University of Wisconsin Press.

- Bharuthram, S. 2012. Making a case for the teaching of reading across the curriculum in higher education. *South African Journal of Education*, 32(2):205-214.
- Bharuthram, S. & Mckenna, S. 2006. A writer-respondent intervention as a means of developing academic literacy. *Teaching in Higher Education*, 11(4):495-507.
- Bojuwoye, O. 2002. Stressful experiences of first year students of selected universities in South Africa. *Counselling Psychology Quarterly*, 15(3):277-290.
- Botha, H. & Cilliers, C. 1999. Preparedness for university study: Designing a thinking skills test. *South African Journal of Higher Education*, 13(1):144-152.
- Boughey, C. 2000. *Contrasting constructions of students' literacy-related experiences at a historically black South African university*. Cape Town: University of the Western Cape.
- Boughey, C. 2013. What are we thinking of? A critical overview of approaches to developing academic literacy in South African higher education. *Journal for Language Teaching*, 47(2):25-41.
- Braine, G. 1995. Writing in the natural sciences and engineering. In: Belcher, D. & Braine, G. (eds). *Academic writing in a second language: Essays on research and pedagogy*. Norwood: Ablex Publishing Corporation.
- Brindley, G. 1998. Outcomes-based assessment and reporting in language learning programmes: A review of the issues. *Language Testing*, 15(1):45-85.
- Brousselle, A. & Champagne, F. 2011. Program theory evaluation: Logic analysis. *Evaluation and Program Planning*, 34(1):69-78.
- Brown, J.D. 2001. *Using surveys in language programs*. Cambridge: Cambridge University Press.
- Butler, G. 2009. The design of a postgraduate test of academic literacy: Accommodating student and supervisor perceptions. *Southern African Linguistics and Applied Language Studies*, 27(3):291-300.
- Butler, G. 2013. Discipline-specific versus generic academic literacy intervention for university education: An issue of impact? *Journal for Language Teaching*, 47(2):71-87.
- Butler, G. & Van Dyk, T. 2004. An academic English language intervention for first year engineering students. *Southern African Linguistics and Applied Language Studies*, 22(1-2):1-8.
- Calderon, A. 2012. *Massification continues to transform higher education*. [Online]. Available: <http://www.universityworldnews.com/article.php?story> [Accessed 20 October 2013].
- Carstens, A. 2008. Preferred genres and rhetorical modes in the humanities and social sciences. *Language Matters*, 39(1):49-65.
- Carstens, A. 2009. *The effectiveness of genre-based approaches in teaching academic writing: Subject-specific versus cross-disciplinary emphases*. Unpublished PhD dissertation. Pretoria: University of Pretoria.
- Carstens, A. 2011. Generic versus discipline-specific writing interventions: Report on a quasi-experiment. *Southern African Linguistics and Applied Language Studies*, 29(2):149-165.
- Carstens, A. 2013a. Collaboration: The key to integration of language and content in academic literacy interventions. *Journal for Language Teaching*, 47(2):109-125.
- Carstens, A. 2013b. Using literacy narratives to scaffold academic literacy in the Bachelor of Education: A pedagogical framework. *Journal for Language Teaching*, 46(2):9-25.
- Carstens, A. & Fletcher, L. 2009a. Evaluating the effectiveness of a cross-disciplinary genre-focused writing intervention. *Journal for Language Teaching*, 43(1):54-65.

- Carstens, A. & Fletcher, L. 2009b. Quantitative evaluation of a subject-specific essay-writing intervention. *Southern African Linguistics and Applied Language Studies*, 27(3):319-332.
- Carstens, W. & Van de Poel, K. 2012. *Teksredaksie*. Stellenbosch: African Sun Media.
- Cheetham, J., Fuller, R., Mcivor, G. & Petch, A. 1992. *Evaluating social work effectiveness*. Buckingham: McGraw-Hill International.
- Cilliers, F.J. & Herman, N. 2010. Impact of an educational development programme on teaching practice of academics at a research-intensive university. *International Journal for Academic Development*, 15(3):253-267.
- Clarence-Fincham, J. 2000. Cummins' contribution to the understanding of language learning and achievement at school. In: Inglis, M., Thomson, C. & Macdonald, A. (eds). *Language in learning and teaching (LILT)*. Pietermaritzburg: University of Natal Press.
- Clark, S. 1998. Tutor development: finding a language for teaching. In: Angelil-Carter, S. (ed.). *Access to success: Literacy in academic contexts*. Cape Town: University of Cape Town Press.
- Cliff, A. 2014. Entry-level students' reading abilities and what these abilities might mean for academic readiness. *Language Matters*, 45(3):313-324.
- Cliff, A. & Hanslo, M. 2009. The design and use of 'alternate' assessments of academic literacy as selection mechanisms in higher education. *Southern African Linguistics and Applied Language Studies*, 27(3):265-276.
- Cliff, A. & Yeld, N. 2006. Academic literacy. In: Griesel, H. (ed.). *Access and entry level benchmarks: The national benchmark tests project*. Pretoria: Higher Education South Africa.
- Cobanoglu, C. & Cobanoglu, N. 2003. The effect of incentives in web surveys: application and ethical considerations. *International Journal of Market Research*, 45(4):475-488.
- Cobb, T. 2015. *Web Vocabprofile*. [Online]. Available: <http://www.lex tutor.ca/vp/eng/> [Accessed 5 March 2015].
- Coetzee-Van Rooy, S. 2010. Evaluation of the Cummins theoretical framework for higher education in South Africa. *Southern African Linguistics and Applied Language Studies*, 28(1):25-38.
- Cole, R., Purao, S., Rossi, M. & Sein, M. 2005. Being proactive: Where action research meets design research. *ICIS 2005 Proceedings*:27.
- Connor, U. & Kramer, M. 1995. Writing from sources: Case studies of graduate students in business management. In: Belcher, D. & Braine, G. (eds). *Academic writing in a second language: Essays on research and pedagogy*. Norwood: Ablex Publishing Cooperation.
- Council on Higher Education. 2009. *The state of higher education in South Africa*. [Online]. Available: http://www.che.ac.za/media_and_publications/higher-education-monitor/higher-education-monitor-8-state-higher-education [Accessed 17 March 2013].
- Coxhead, A. 2000. A new academic wordlist. *TESOL Quarterly*, 34(2):213-238.
- Cummins, J. 2001. Linguistic interdependence and the educational development of bilingual children. In: Baker, C. & Hornberger, N.H. (eds). *Linguistic interdependence and the educational development of bilingual children*. Clevedon: Multilingual Matters.
- Cummins, J. & Swain, M. 1986. *Bilingualism in education: Aspects of theory, research and practice*. London: Longman.

- Davies, A. 2001. The logic of testing Languages for Specific Purposes. *Language Testing*, 18(2):133-147.
- Davies, A. 2010. Assessing the academic literacy of additional-language students. *Southern African Linguistics and Applied Language Studies*, 27(3):xi-xii.
- Davies, A., Brown, A., Elder, C., Hill, K., Lumley, T. & Mcnamara, T. 1999. *Dictionary of language testing*. Cambridge: Cambridge University Press.
- De Graaff, R. & Housen, A. 2009. Investigating the effects and effectiveness of L2 instruction. In: Long, M.H. & Doughty, C.J. (eds). *The handbook of language teaching*. Chichester: Wiley-Blackwell.
- De Klerk, E., Van Deventer, I. & Van Schalkwyk, S. 2006. Small victories over time: The impact of an academic development intervention at Stellenbosch University. *Education as Change*, 10(2):149-169.
- De Vos, A.S., Fouché, C.B., Strydom, H. & Delport, C.S.L. 2011. *Research at grassroots for the social sciences and human service professions*. Pretoria: Van Schaik.
- Defazio, J., Jones, J., Tennant, F. & Hooke, S.A. 2010. Academic literacy: The importance and impact of writing across the curriculum. *Journal of the Scholarship of Teaching and Learning*, 10(2):34-47.
- Denzin, N.K. 1989. *The research act: A theoretical introduction to sociological methods*, 3rd ed. Englewood Cliffs: Prentice Hall.
- Department of Higher Education and Training. 2014. *White paper for post-school education and training: Building on expanded, effective and integrated post-school education*. [Online]. Available: www.gov.za/documents/download.php?f=207591 [Accessed 31 January 2014].
- Dewey, J. 1933. Analysis of reflective thinking. In: Hickman, L.A. & Alexander, T.M. (eds). *Reprinted in 1998 in The Essential Dewey*. Bloomington: Indiana University Press.
- Douglas, D. 2000. *Assessing language for specific purposes*. Cambridge: Cambridge University Press.
- Douglas, D. 2001. Language for Specific Purposes assessment criteria: Where do they come from? *Language Testing*, 18(2):171-185.
- Dudley-Evans, T. 1995. Common-core and specific approaches to the teaching of academic writing. In: Belcher, D. & Braine, G. (eds). *Academic writing in a second language: Essays on research and pedagogy*. Norwood: Ablex Publishing Corporation.
- Dudley-Evans, T. & St John, M.J. 1998. *Developments in English for Specific Purposes: A multi-disciplinary approach*. Cambridge: Cambridge University Press.
- Elder, C. 2001. Assessing the language proficiency of teachers: Are there any border controls? *Language Testing*, 18(2):149-170.
- Engstrom, C. & Tinto, V. 2008. Access without support is not opportunity. *Change – The Magazine of Higher Learning*, 40(1):46-50.
- Eskey, D.E. 1983. Meanwhile, back in the real world... Accuracy and fluency in second language teaching. *TESOL Quarterly*, 17(2):315-323.
- Farnill, D. & Hayes, S. 1996. Do NESB university students with poor English skills make rapid linguistic gains in mainstream studies? *Higher Education Research and Development*, 15(2):261-268.
- Feast, V. 2002. The impact of IELTS scores on performance at university. *International Education Journal*, 3(4):70-85.
- Flowerdew, J. 1993. An educational, or process, approach to the teaching of professional genres. *ELT Journal*, 47(4):305-316.

- Flowerdew, J. & Peacock, M. 2001. Issues in EAP: A preliminary perspective. In: Flowerdew, J. & Peacock, M. (eds). *Research perspectives on English for academic purposes*. Cambridge: Cambridge University Press.
- Foddy, W. 1993. *Constructing questions for interviews and questionnaires*. Cambridge: Cambridge University Press.
- Fouché, I. 2009. *Improving the academic literacy levels of first-year Natural Sciences students by means of an academic literacy intervention*. Unpublished MA dissertation. Pretoria: University of Pretoria.
- Fouché, I. 2015. Towards impact measurement: An overview of approaches for assessing the impact of academic literacy abilities. *Stellenbosch Papers in Linguistics*, 44:19-35.
- Fouché, I. & Immelman, S. 2015. Rubric for assessing oral presentations: LST 133 and LST 143. Pretoria: University of Pretoria.
- Fulcher, G. 1999. Assessment in English for academic purposes: Putting content validity in its place. *Applied Linguistics*, 20:221-236.
- Gardner, D. & Davies, M. 2013. A new academic vocabulary list. *Applied Linguistics*, 1:1-24.
- Gee, J. 1991. Socio-cultural approaches to literacy (literacies). *Annual Review of Applied Linguistics*, 12:31-48.
- Gee, J. 2008. *Social linguistics and literacies, 3rd ed.* Abingdon: Routledge.
- Gibbs, G. & Coffey, M. 2004. The impact of training of university teachers on their teaching skills, their approach to teaching and the approach to learning of their students. *Active Learning in Higher Education*, 5(1):87-100.
- Gliem, J.A. & Gliem, R.R. 2003. *Calculating, interpreting, and reporting Cronbach's alpha reliability coefficient for Likert-type scales*. Columbus: Midwest Research-to-Practice Conference in Adult, Continuing, and Community Education.
- Gnutzman, C. 2009. Language for specific purposes vs. general language. In: Knapp, K., Seidlhofer, B. & Widdowson, H. (eds). *Handbook of foreign language communication and learning*. New York: Mouton de Gruyter.
- Goodier, C. & Parkinson, J. 2005. Discipline-based academic literacy in two contexts. *Journal for Language Teaching*, 39(1):66-79.
- Grabe, W. 2009. Teaching and testing reading. In: Long, M.H. & Doughty, C.J. (eds). *The handbook of language teaching*. Chichester: Wiley-Blackwell.
- Grabe, W. & Stoller, F.L. 1997. Content-based instruction: Research foundations. In: Snow, M.A. & Brinton, D. (eds). *The Content-based Classroom: Perspectives on Integrating Language and Content*. White Plains: Addison Wesley Longman.
- Grabowski, J. 1996. Writing and speaking: Common grounds and differences toward a regulation theory of written language production. In: Levy, C.M. & Randsdell, S. (eds). *The science of writing: Theories, methods, individual differences, and applications*. Mahwah: Laurence Erlbaum Associates.
- Green, A. 2005. EAP study recommendations and score gains on the IELTS academic writing test. *Assessing Writing*, 10(1):44-60.
- Halliday, M.A. 1993. Towards a language-based theory of learning. *Linguistics and education*, 5(2):93-116.
- Hamp-Lyons, L. & Lumley, T. 2001. Assessing Language for Specific Purposes. *Language Testing*, 18(2):127-132.
- Harker, M. & Koutsantoni, D. 2005. Can it be as effective? Distance versus blended learning in a web-based EAP programme. *ReCALL*, 17(2):197-216.

- Hatry, H.P. & Newcomer, K.E. 2010. Pitfalls in evaluations. In: Wholey, J.S., Hatry, H.P. & Newcomer, K.E. (eds). *Handbook of practical program evaluation*. 3rd ed. San Francisco: Jossey-Bass.
- Hawkey, R. 2006. *Studies in language testing: Impact theory and practice*. Cambridge: Cambridge University Press.
- Heatley, A., Nation, I.S.P. & Coxhead, A. 2002. *RANGE and FREQUENCY programs*. [Online]. Available: <http://www.victoria.ac.nz/lals/staff/paul-nation.aspx> [Accessed 20 October 2014].
- Heaton, J.B. 1988. *Writing English language tests*. New York: Longman.
- Henning, E., Van Rensburg, W. & Smit, B. 2004. *Finding your way in qualitative research*. Pretoria: Van Schaik.
- Henning, G. 1987. *A guide to language testing: Development, evaluation, research*. Cambridge: Newbury House.
- Hevner, A.R., March, S.T., Park, J. & Ram, S. 2004. Design science in information systems research. *MIS quarterly*, 28(1):75-105.
- Higher Education South Africa. 2005. *HESA structures*. [Online]. Available: <http://web.archive.org/web/20050301015907/http://www.sauvca.org.za/highered/> [Accessed 7 October 2015].
- Higher Education South Africa. 2008. *Address to the Parliamentary Portfolio Committee on Education*. [Online]. Available: http://www.pmg.org.za/files/docs/080624_hesa.pdf [Accessed 31 March 2013].
- Holder, G.M., Jones, J., Robinson, R.A. & Krass, I. 1999. Academic literacy skills and progression rates amongst pharmacy students. *Higher Education Research & Development*, 18(1):19-30.
- Holliday, A. 2004. The value of reconstruction in revealing hidden or counter cultures. *Journal of Applied Linguistics*, 1(3):275-294.
- Holtzman, J.M., Elliot, N., Biber, C.L. & Sanders, R.M. 2005. Computerized assessment of dental student writing skills. *Journal of Dental Education*, 69(2):285-295.
- Hornsby, D.J. & Osman, R. 2014. Massification in higher education: Large classes and student learning. *Higher Education*, 67(6):711-719.
- Howes, A. 2003. Teaching reforms and the impact of paid adult support on participation and learning in mainstream schools. *Support for Learning*, 18(4):147-153.
- Hughes, A. 2003. *Testing for language teachers*, 2nd ed. Cambridge: Cambridge University Press.
- Hutchinson, T. & Waters, A. 1987. *English for Specific Purposes*. Cambridge: Cambridge University Press.
- Hyland, K. 2004. *Disciplinary discourses: Social interactions in academic writing*. Michigan: University of Michigan Press.
- Hyland, K. 2009. 12 specific purpose programs. In: Long, M.H. & Doughty, C.J. (eds). *The Handbook of Language Teaching*. Chichester: Wiley-Blackwell.
- Hyland, K. & Hamp-Lyons, L. 2002. EAP: Issues and directions. *Journal of English for Academic Purposes*, 1(1):1-12.
- Hyon, S. 2001. Long-term effects of genre-based instruction – a follow-up study of an EAP reading course. *English for Specific Purposes*, 20:417-438.
- Jacobs, C. 2005. On being an insider on the outside: New spaces for integrating academic literacies. *Teaching in Higher Education*, 10(4):475-487.
- Jacobs, C. 2013. Academic literacies and the question of knowledge. *Journal for Language Teaching*, 47(2):127-139.
- Jacoby, S., Leech, D. & Holten, C. 1995. A genre-based developmental writing course for undergraduate ESL science majors. In: Belcher, D. & Braine, G. (eds).

- Academic writing in a second language: Essays on research and pedagogy.* Norwood: Ablex Publishing Cooperation.
- Jama, M., Mapesela, M. & Beyliefeld, A. 2008. Theoretical perspectives on factors affecting the academic performance of students. *South African Journal of Higher Education*, 22(5):992-1005.
- Jansen, J.V. 2004. *Taalvaardigheid as moderator in die voorspelling van akademiese prestasie*. Ongepubliseerde MA dissertasie. Bloemfontein: Universiteit van die Vrystaat.
- Jeffery, J. 2009. Constructs of writing proficiency in US state and national writing assessments: Exploring variability. *Assessing Writing*, 14:3-24.
- Jick, T.D. 1979. Mixing qualitative and quantitative methods: Triangulation in action. *Administrative Science Quarterly*, 24(4):602-611.
- Johnes, J. & Taylor, J. 1989. Undergraduate non-completion rates: Differences between UK universities. *Higher Education*, 18(2):209-225.
- Johns, A. 1995. Teaching classroom and authentic genres: Initiating students into academic cultures and discourses. In: Belcher, D. & Braine, G. (eds). *Academic writing in a second language: Essays on research and pedagogy*. Norwood: Ablex Publishing Corporation.
- Johnson, G.M. & Buck, G.H. 1995. Students' personal and academic attributions of university withdrawal. *Canadian Journal of Higher Education*, 25(2):53-77.
- Joint Committee on Standards for Educational Evaluation. 2014. *Program evaluation standards statements*. [Online]. Available: <http://www.jcsee.org/program-evaluation-standards-statements> [Accessed 5 December 2013].
- Judd, T. & Keith, B. 2012. Student learning outcomes assessment at the program and institutional levels. In: Secolsky, C. & Denison, D.B. (eds). *Handbook on measurement, assessment, and evaluation in higher education*. New York: Routledge.
- Kaburise, P. 2012. Foundation English at the University of Venda: How responsive is it? *South African Journal of Higher Education*, 26(1):77-90.
- Kasper, L.F. 1994. Improved reading performance for ESL students through academic course pairing. *Journal of Reading*, 37(5):376-384.
- Kasper, L.F. 1997. The impact of content-based instructional programs on the academic progress of ESL students. *English for Specific Purposes*, 16(4):309-320.
- Keeve, A., Naudé, L. & Esterhuyse, K. 2012. Enkele voorspellers van die akademiese prestasie van eerste-jaarstudente in drie- en vierjaar-kurrikulums. *Acta Academica*, 44(1):121-158.
- Kennedy, C. & Bolitho, R. 1984. *English for Specific Purposes*. London: Macmillan.
- Kern, R. 2000. *Literacy and language teaching*. Oxford: Oxford University Press.
- Kiely, R. 2009. Small answers to the big question: Learning from language programme evaluation. *Language Teaching Research*, 13(1):99-130.
- Krekeler, C. 2006. Language for special academic purposes (LSAP) testing: The effect of background knowledge revisited. *Language Testing*, 23(1):99-130.
- Kreuger, L.W. & Neuman, W.L. 2006. *Social work research methods: Qualitative and quantitative applications*. Boston: Pearson & Allyn Bacon.
- Kwiek, M., Lebeau, Y. & Brown, R. 2014. *Who shall pay for the public good? Comparative trends in the funding crisis of public higher education*. [Online]. Available: https://repozytorium.amu.edu.pl/jspui/handle/10593/9857?mode=full&submit_simple>Show+full+item+record [Accessed 19 February 2015].
- Lea, M.R. 2004. Academic literacies: A pedagogy for course design. *Studies in Higher Education*, 29(6):739-756.

- Lea, M.R. & Street, B.V. 1998. Student writing in higher education: An academic literacies approach. *Studies in Higher Education*, 23(2):157-172.
- Leibowitz, B. 2001. *Students' prior learning and their acquisition of academic literacy at a multilingual South African university*. Sheffield: University of Sheffield.
- Leibowitz, B. 2010. *The significance of academic literacy in higher education: Students' prior learning and their acquisition of academic literacy at a multilingual South African university*. Saarbrücken: Lap Lambert.
- Lillis, T. 2003. Student writing as 'academic literacies': Drawing on Bakhtin to move from critique to design. *Language and Education*, 17(3):192-207.
- Lincoln, Y., Lynham, S. & Guba, E. 2011. Paradigmatic controversies, contradictions, and emerging confluences, revisited. In: Denzin, N.K. & Lincoln, Y.S. (eds). *The SAGE handbook of qualitative research*. Thousand Oaks: Sage.
- LLC. 2015. *Turnitin*. [Online]. Available: <http://www.turnitin.com> [Accessed 2 February 2015].
- Luke, A. & Freebody, P. 1999. Further notes on the four resources model. *Reading online*, 3. [Online]. Available: <http://www.readingonline.org/research/lukefreebody.html> [Accessed 5 March 2014].
- Lynch, B.K. 1990. A context-adaptive model for program evaluation. *TESOL Quarterly*, 24(1):23-42.
- Lynch, B.K. 1992. Evaluating a program inside and out. In: Alderson, C.J. & Beretta, A. (eds). *Evaluating second language education*.
- Lynch, B.K. 1996. *Language program evaluation: Theory and practice*. Cambridge: Cambridge University Press.
- Lynch, B.K. 2001. Rethinking assessment from a critical perspective. *Language Testing*, 18(4):351-372.
- Lynch, B.K. 2003. *Language assessment and programme evaluation*. Edinburgh: Edinburgh University Press.
- Marais, F. & Van Dyk, T. 2010. Put listening to the test: An aid to decision making in language placement. *Per Linguam*, 26(2):1-19.
- Marais, F.C. 2009. *An investigation into the significance of listening proficiency in the assessment of academic literacy levels at Stellenbosch University*. Unpublished MA dissertation. Stellenbosch: University of Stellenbosch.
- Martin, J.R. 2012. Grammar meets genre: reflections on the 'Sydney School'. *Arts: The Journal of the Sydney University Arts Association*, 22:47-95.
- McGhie, V., Van Der Walt, C., Van Schalkwyk, S., Walt, C.V.D. & Schalkwyk, S.V. 2012. Successful first-year learning: A social cognitive view of academic literacy. *Journal for Language Teaching*, 46(2):27-41.
- Mckenzie, K. & Schweitzer, R. 2001. Who succeeds at university? Factors predicting academic performance in first year Australian university students. *Higher Education Research and Development*, 20(1):21-33.
- Mhlongo, G.J. 2014. *The impact of an academic literacy intervention on the academic literacy levels of first year students: The NWU (Vaal Triangle Campus) experience*. Unpublished MA dissertation. Potchefstroom: North West University.
- Michaelidou, N. & Dibb, S. 2006. Using email questionnaires for research: Good practice in tackling non-response. *Journal of Targeting, Measurement and Analysis for Marketing*, 14(4):289-296.
- Michaelsen, L.K. & Sweet, M. 2011. Team-based learning. *New Directions for Teaching and Learning*, 2011(128):41-51.
- Misselhorn, A. 1997. *An investigation into the role of informal study groups in the acquisition of academic literacy among undergraduate black students at the*

- University of Natal in Durban*. Unpublished MA dissertation. Durban: University of Natal.
- Murie, R. & Thomson, R. 2001. When ESL is developmental: A model program for the freshman year. In: Higbee, J.L. (ed.). Warrensburg: National Association of Developmental Education.
- Murphy, S. & Yancey, K.B. 2008. Construct and consequence: Validity in writing assessment. In: Bazerman, C. (ed.). *Handbook of research on writing: History, society, school, individual, text*. Madison: University of Wisconsin Press.
- Nagy, W. & Townsend, D. 2012. Words as tools: Learning academic vocabulary as language acquisition. *Reading Research Quarterly*, 47(1):91-108.
- Nair, P. 2004. A theoretical framework for an access programme encompassing further education training: remedy for educational wastage? *South African Journal of Higher Education*, 16(2):94-103.
- Newcomer, K.E., Hatry, H.P. & Wholey, J.S. 2010. Planning and designing useful evaluations. In: Wholey, J.S., Hatry, H.P. & Newcomer, K.E. (eds). *Handbook of Practical Program Evaluation*. 3rd ed. San Francisco: Jossey-Bass.
- Newcomer, K.E. & Triplett, T. 2010. Using surveys. In: Wholey, J.S., Hatry, H.P. & Newcomer, K.E. (eds). 3rd ed. San Francisco: Jossey-Bass.
- Ngoepe, L.J. 2007. *The University of Limpopo mathematics and science foundation year course in English and study skills: An evaluation*. Unpublished Phd dissertation. Potchefstroom: North-West University.
- Ngwenya, T. 2010. Correlating first-year law students' profile with the language demands of their content subjects. *Per Linguam*, 26(1):74-99.
- Nitko, A.J. & Brookhart, S.M. 2011. *Educational assessment of students*, 6th ed. Boston: Pearson.
- O'Sullivan, B. 2012. Assessment issues in languages for specific purposes. *Modern Language Journal*, 96:71-88.
- Paltridge, B. 2001. Linguistic research and EAP pedagogy. In: Flowerdew, J. & Peacock, M. (eds). *Research perspectives on English for academic purposes*. Cambridge: Cambridge University Press.
- Paltridge, B. 2004. Academic writing. *Language teaching*, 37(2):87-105.
- Parkinson, J., Jackson, L., Kirkwood, T. & Padayachee, V. 2007. A scaffolded reading and writing course for foundation level science students. *English for Specific Purposes*, 26(4):443-461.
- Parkinson, J., Jackson, L., Kirkwood, T. & Padayachee, V. 2008. Evaluating the effectiveness of an academic literacy course: Do students benefit? *Per Linguam*, 24(1):11-29.
- Patton, M.Q. 2002. *Qualitative research and evaluation methods*. Thousand Oaks: Sage.
- Penrose, A.M. 2002. Academic literacy perceptions and performance. *Research in the Teaching of English*, 36(4):437-461.
- Petersen-Waughtal, M. & Van Dyk, T. 2011. Towards informed decision making: The importance of baseline academic literacy assessment in promoting responsible university access and support. *Journal for Language Teaching*, 45(1):99-114.
- Pienaar, C. 2005. Shared assessment: Empowering student writers. *Language Matters*, 36(2):193-204.
- Potter, C. & Van der Merwe, E. 1993. Academic performance in engineering. *South African Journal of Higher Education*, 7(1):33-40.
- Pretorius, E.J. 1995. Reading as an interactive process: Implications for studying through the medium of a second language. *Communicatio*, 21(2):33-43.

- Prinsloo, P. & Subotzky, G. 2011. *Towards a framework and strategy for enhancing student success, retention, graduation and satisfaction at UNISA: Final Report Task Team*. Pretoria: UNISA.
- Radloff, A. & De La Harpe, B. 2000. Helping students develop their writing skills. A resource for lecturers. *Actas del ASET Higher Education Research and Development in South Asia Conference*. Queensland.
- Rambiritch, A. 2012. *Accessibility, transparency and accountability as regulative conditions for a post-graduate test of academic literacy*. Unpublished PhD dissertation. Bloemfontein: University of the Free State.
- Rambiritch, A. 2013. Validating the Test of Academic Literacy for Postgraduate Students (TALPS). *Journal for Language Teaching*, 47(1):175-193.
- Rao, T.V. & Saxena, S. 2014. *Measuring and managing institutional environment of institutes of professional education*. [Online]. Available: <http://iimahd.ernet.in/assets/snippets/workingpaperpdf/14346925982014-03-05.pdf> [Accessed 27 April 2014].
- Renger, R. & Titcomb, A. 2002. A three-step approach to teaching logic models. *American Journal of Evaluation*, 23(4):493-503.
- Renkema, J. 1998. De C3-analyse. *Tekst[blad]*, 4(4):29-31.
- Renkema, J. 2001. Undercover research into text quality as a tool for communication management. In: Janssen, D. & Neutelings, R. (eds). *Reading and writing public documents: Problems, solutions and characteristics*. Amsterdam: John Benjamins.
- Robinson, P.C. 1980. *ESP (English for Specific Purposes): The present position*. Oxford: Pergamon Press.
- Rosenthal, J.W. 1996. *Teaching science language minority students: Theory and practice*. Clevedon: Multilingual Matters.
- Ross, S.J. 2009. Program evaluation. In: Long, M.H. & Doughty, C.J. (eds). *The handbook of language teaching*. Chichester: Wiley-Blackwell.
- Scott, I., Ndebele, N., Badsha, N., Figaji, B., Gevers, W. & Pityana, B. 2013. A proposal for undergraduate curriculum reform in South Africa: The case for a flexible curriculum structure. Pretoria: Council on Higher Education.
- Scott, I., Yeld, N. & Hendry, J. 2007. *Higher education monitor: A case for improving teaching and learning in South African higher education*. Pretoria: Council on Higher Education.
- Scouller, K. 2012. The influence of assessment method on students' learning approaches: Multiple choice question examination versus assignment essay. *Higher Education*, 35:453-472.
- Seabi, J. 2012. Research designs and data collection techniques. In: Maree, J.G. (ed.). *Complete your thesis or dissertation successfully: Practical guidelines*. Claremont: Juta & Company.
- Sebolai, K. 2014. Evaluating academic literacy teaching at a South African university : A case study of an academic literacy programme. *Journal for Language Teaching*, 48(1):51-69.
- Shaw, P. & Liu, E.T. 1998. What develops in the development of second-language writing? *Applied Linguistics*, 19(2):225-254.
- Shaw, S.D. & Weir, C.J. 2007. *Examining writing: Research and practice in assessing second language writing*. Cambridge: Cambridge University Press.
- Shay, S. 2015. Curriculum reform in higher education: A contested space. *Teaching in Higher Education*, 20(4):431-441.

- Sinclair, K.E. 1995. *Identifying future leaders: A study of career progression and development*. Melbourne: Business/Higher Education Round Table.
- Singer, E. & Ye, C. 2013. The use and effects of incentives in surveys. *The ANNALS of the American Academy of Political and Social Science*, 645(1):112-141.
- Smit, R. 2012. Towards a clearer understanding of student disadvantage in higher education: Problematising deficit thinking. *Higher Education Research & Development*, 31(3):369-380.
- Snow, C.E. & Kim, Y. 2007. Large problem spaces: The challenge of vocabulary for English language learners. In: Wagner, R.K., Muse, A.E. & Tannenbaum, K.R. (eds). *Vocabulary acquisition: Implications for reading comprehension* New York: Guilford.
- Snow, M.A. & Brinton, D.M. 1988. Content-based language instruction: Investigating the effectiveness of the adjunct model. *TESOL Quarterly*, 22(4):553-574.
- Song, B. 2006. Content-based ESL instruction: Long-term effects and outcomes. *English for Specific Purposes*, 25(4):420-437.
- Stephen, D.F., Welman, J.C. & Jordaan, W.J. 2004. English language proficiency as an indicator of academic performance at a tertiary institution. *South African Journal of Human Resource Management*, 2(3):42-53.
- Steyn, S. 2012. *Investigating issues of equivalence in the design of parallel English and Afrikaans tests of academic literacy for grade 11 learners*. . University of the Free State.
- Steyn, S. Forthcoming. *Investigating issues of equivalence in the design of parallel English and Afrikaans tests of advanced language ability for learners at FET level*. North-West University.
- Stoller, F.L. 2004. Content-based instruction: Perspectives on curriculum planning. *Annual Review of Applied Linguistics*, 24: 261-283.
- Storch, N. & Tapper, J. 2009. The impact of an EAP course on postgraduate writing. *Journal of English for Academic Purposes*, 8(3):207-223.
- Street, B. 1984. *Literacy in theory and practice*. Cambridge: Cambridge University Press.
- Street, B. 1993. Literacy and nationalism. *History of European Ideas*, 16(1-3):225-228.
- Street, B. 2004. Academic literacies and the "new orders": Implications for research and practice in student writing in higher education. *LATISS: Learning and Teaching in the Social Sciences*, 1(1):9-20.
- Street, B. 2008. New literacies, new times: Developments in literacy studies. *Encyclopedia of Language and Education*, 2:3-14.
- Swales, J.M. 1990. *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.
- Taylor, G., Ballard, B., Beasley, V., Bock, H., Clanchy, J. & Nightingale, P. 1988. *Literacy by degrees*. Buckingham: Open University Press.
- Teichler, U. 1998. Massification: A challenge for institutions of higher education. *Tertiary Education & Management*, 4(1):17-27.
- Terraschke, A. & Wahid, R. 2011. The impact of EAP study on the academic experiences of international postgraduate students in Australia. *Journal of English for Academic Purposes*, 10(3):173-182.
- The Concise Oxford English Dictionary. 2002. 10th ed. Oxford: Oxford University Press.
- Thompson, J.L. 2011. *An evaluation of a university level English for tourism program*. Chiang Mai: Payap University.
- Thompson, R.M. 1990. Writing-proficiency tests and remediation: Some cultural differences. *TESOL Quarterly*, 24(1):99-102.

- Tinto, V. & Pusser, B. 2006. Moving from theory to action: Building a model of institutional action for student success. *National Postsecondary Education Cooperative*, June:1-51.
- Van de Poel, K. & Van Dyk, T. 2015. Discipline-specific academic literacy and academic integration. In: Wilkinson, R. & Walsh, M.L. (eds). *Integrating Content and Language in Higher Education*. Frankfurt am Main: Peter Lang.
- Van der Merwe, C.A. & Van der Merwe, S. 2009. Student success: Data mining measures what matters. *Journal for New Generation Sciences*, 7(2):284-302.
- Van Der Slik, F. & Weideman, A. 2008. Measures of improvement in academic literacy. *Southern African Linguistics and Applied Language Studies*, 26(3):363-378.
- Van Der Walt, J.L. & Steyn, H. 2007. Pragmatic validation of a test of academic literacy at tertiary level. *Ensovoort*, 11(2):138-153.
- Van Dyk, T. 2005. Towards providing effective academic literacy intervention. *Per Linguam*, 21(2):38-51.
- Van Dyk, T. 2010. *Konstitutiewe voorwaardes vir die ontwerp van 'n toets van akademiese geletterdheid*. Ongepubliseerde PhD dissertasie. Bloemfontein: Universiteit van die Vrystaat.
- Van Dyk, T. 2014. *Vraelys oor akademiese geletterdheid / Questionnaire on academic literacy*. *Sentrum vir Akademiese en Professionele Taalpraktyk*. Potchefstroom: Noordwes-Universiteit.
- Van Dyk, T. 2015. Tried and tested. *Tijdschrift voor Taalbeheersing*, 37(2):159-186.
- Van Dyk, T., Cillié, K., Coetzee, M., Ross, S. & Zybrands, H. 2011a. Ondersoek na die impak van 'n akademiese geletterdhedsofintervensie op eerstejaarstudente se akademiese taalvermoë. *LitNet Akademies*, 8(3):487-506.
- Van Dyk, T. & Van de Poel, K. 2013. Towards a responsible agenda for academic literacy development: Considerations that will benefit students and society. *Journal for Language Teaching*, 47(2):43-69.
- Van Dyk, T., Van de Poel, K. & Van der Slik, F. 2013. Reading ability and academic acculturation: The case of South African students entering higher education. *Stellenbosch Papers in Linguistics Plus*, 42:353-369.
- Van Dyk, T., Van Rensburg, A. & Marais, F. 2011b. Levelling the playing field: An investigation into the translation of academic literacy tests. *Journal for Language Teaching*, 45(1):153-169.
- Van Dyk, T. & Weideman, A. 2004. Switching constructs: On the selection of an appropriate blueprint for academic literacy assessment. *Journal for Language Teaching*, 38(1):1-13.
- Van Dyk, T., Zybrands, H., Cillié, K. & Coetzee, M. 2009. On being reflective practitioners: The evaluation of a writing module for first-year students in the Health Sciences. *Southern African Linguistics and Applied Language Studies*, 27(3):333-344.
- Van Rooy, B. & Coetzee-Van Rooy, S. 2015. The language issue and academic performance at a South African university. *Southern African Linguistics and Applied Language Studies*, 33(1):31-46.
- Van Wyk, A. 2002. A university bridging course focusing on academic reading and writing skills. *Journal for Language Teaching*, 36(3&4):220-232.
- Van Wyk, A.L. & Greyling, W.J. 2008. Developing reading in a first-year academic literacy course. *Stellenbosch Papers in Linguistics*, 38:205-219.
- Wait, T.H. 2007. *The relationship between the use of language learning strategies and performance on a standardised English proficiency test*. Unpublished MA dissertation. Stellenbosch: Stellenbosch University.

- Wallace, M.C., Shorten, A. & Crookes, P.A. 2000. Teaching information literacy skills: an evaluation. *Nurse Education Today*, 20(6):485-489.
- Weideman, A. 2003a. *Academic literacy: Prepare to learn*. Pretoria: Van Schaik.
- Weideman, A. 2003b. Justifying course and task design in language teaching. *Acta Academica*, 35(3):26-48.
- Weideman, A. 2009. Constitutive and regulative conditions for the assessment of academic literacy. *Southern African Linguistics and Applied Language Studies*, 27(3):235-251.
- Weideman, A. 2013. Academic literacy interventions: What are we not yet doing, or not yet doing right? *Journal for Language Teaching*, 47(2):11-23.
- Weigle, S. 2002. *Assessing writing*. Cambridge: Cambridge University Press.
- Weinbach, R.W. 2005. *Evaluating social work services and programs*. Boston: Pearson/Allyn and Bacon.
- Wells, C.S. & Wollack, J.A. 2003. *An instructor's guide to understanding test reliability*. Wisconsin: University of Wisconsin.
- Wentland, E. 2011. Survey use in academic contexts: considerations and guidelines. In: Secolsky, C. & Denison, D.B. (eds). *Handbook on measurement, assessment, and evaluation in higher education*. New York: Routledge.
- Winberg, C., Wright, J., Birch, B.W. & Jacobs, C. 2013. Conceptualising linguistic access to knowledge as interdisciplinary collaboration. *Journal for Language Teaching*, 47(2):89-107.
- Wingate, U. & Tribble, C. 2012. The best of both worlds? Towards an English for Academic Purposes/Academic Literacies writing pedagogy. *Studies in Higher Education*, 37(4):481-495.
- Yamashita, J. 2002. Mutual compensation between L1 reading ability and L2 language proficiency in L2 reading comprehension. *Journal of Research in Reading*, 25(1):81-95.
- Yeh, S.S. 1998. Empowering education: Teaching argumentative writing to cultural minority middle-school students. *Research in the Teaching of English*, 33(1):49-83.
- Yeld, N. 2001. *Equity, assessment and language of learning: Key issues for higher education selection and access in South Africa*. Cape Town: University of Cape Town.
- Yeld, N. 2010. Some challenges and responses: Higher education in South Africa. *Discourse*, 38(1):24-36.
- Yost, D.S., Sentner, S.M. & Forlenza-Bailey, A. 2000. An examination of the construct of critical reflection: Implications for teacher education programming in the 21st century. *Journal of Teacher Education*, 51(1):39.
- Zambo, R. & Cleland, J. 2005. Integrating mathematics and the language arts. *Academic Exchange Quarterly*, 9(3):156-161.
- Zhu, W. 2009. Performing argumentative writing in English: Difficulties, processes, and strategies. *TESL Canada Journal*, 19(1):34-50.
- Zuengler, J. & Miller, E.R. 2006. Cognitive and sociocultural perspectives: Two parallel SLA worlds? *TESOL Quarterly*, 40(1):35-58.

Appendix A: Rubric for assessing writing assignments

		1 (Poor)	2 (Below average)	3 (Average)	4 (Good)	5 (Excellent)	N/A
Structure and development							
1.	Introduction						
2.	Thesis statement						
3.	Development of main argument						
4.	Paragraph development (topic sentences, main ideas, supporting information)						
5.	Relevance of content to topic						
6.	Linking devices (structuring thought with discourse markers, pronouns etc)						
7.	Conclusion						
Academic writing style							
8.	Syntax: phrase and clause structure						
9.	Academic vocabulary						
10.	Technical vocabulary						
11.	Style (formality; rhetorical mode)						
12.	Integration of visual data						
Editing							
13.	Spelling, capitalisation and punctuation						
14.	Concord and tense						
15.	Layout and format						
16.	Referencing technique						
Use of source material							
17.	Relevance of source data						
18.	Appropriately citing quotations						
19.	Paraphrasing information from source texts						
20.	Integration of source data with text (synthesising)						

Appendix B: The C3-model for evaluation of text quality

Renkema (2001:40-44; 1998:29-31) proposes the following model for evaluating a text's quality. Fifteen evaluation points are included, and these are organised from top to bottom and left to right. Thus, says Renkema (2001:43), the more to the top or to the left the evaluation point is, the more important it is.

	Analysis criteria		
Text levels	Correspondence	Consistence	Correctness
Text type	1. Appropriateness	2. Purity of genre	3. Application of genre rules
Content	4 Sufficient information	5. Agreement between facts	6. Correctness of facts
Structure	7. Sufficient coherence	8. Consistent structure	9. Correct linking words
Wording	10. Appropriate working	11. Unity of style	12. Correct syntax and choice of words
Presentation	13. Appropriate lay-out	14. Lay-out adapted to text	15. Correct spelling and punctuation

Appendix C: Storch and Tapper quantitative analysis

Storch and Tapper (2009) use the following measures to quantitatively assess student writing⁵².

A. Fluency

Measured in terms of:

- Number of words
- Words per T-unit

B. Accuracy

Scripts are coded for the following 18 categories of errors (categorised into six broad themes, namely “syntactical”, “morphological: nominal”, “morphological: verbal”, “grammatical”, “lexical” and “mechanics” (taken from Storch and Tapper 2009:219-220)

Syntactical

1. Word order
2. Absence of major constituent, such as subject, verb, object
3. Absence of minor constituent (e.g. ‘Enterprises may not be professional [enough] to master the coordination of ...’)
4. Errors in linking ideas (missing, redundant, or incorrect)

Morphological:nominal

5. Plural
6. Agreement (noun or pronoun with verb)
7. Possessive

Morphological:verbal

8. Tense and verb form. Errors of tense, aspect, mood and form for the same verb were counted as one error.
9. Agreement of verb with subject. Agreement errors involving both subject and verb in the same phrase were counted as one error.
e.g. ‘Every details [detail] need [needs] to be considered’.
10. Passive form (missing or incorrect)
11. Derivational (word form) e.g. ‘very technologic [technological] parameters’

Grammatical

12. Determiners (e.g. this, that, it, those). Missing, redundant or incorrect. e.g. ‘when building cantilever bridges. Those [These] bridges ...’
13. e.g. ‘Looking at its [this] background and current situation, .’
14. Articles. Errors of article and noun plurals were counted as one error.
e.g. When the context shows that ‘ the problem’ should be ‘problems’, one error was counted.
15. Prepositions (missing or redundant)

⁵² Text structure and rhetorical quality are also investigated in the Storch and Tapper study using a more traditional rubric. However, the current study is primarily interested in the quantitative measures employed in the Storch and Tapper study. As rubrics are dealt with separately in the current study, the section on text structure and rhetorical quality is not included in this summary.

Lexical

16. Word choice. (Register errors such as ‘lots of’ were not included).
e.g. ‘Many countries still out of [lack] responsibility’.
e.g. ‘especially in developing countries, such as my hometown [home] – China’.
Prepositions were coded as word choice if the choice was incorrect.
e.g. ‘The glaciers in [at] the two poles of the earth’.
17. Collocation. Erroneous expressions and phrasal verbs were counted as one error.
e.g. the key of the [to] success e.g. I am interested to conduct [in conducting]. If meaning was so obscure that reformulation was impossible, a phrase or clause was counted as one collocation error.
e.g. ‘The definition should “with which” or “follow with” conclude the rights, the duties .’ was one error.

Mechanics (Spelling omitted)

18. Capitalisation
19. Punctuation (if meaning was affected)
A repeated error was counted each time it occurred. Errors were counted according to the minimal number of corrections required to make a phrase or clause error-free, while maintaining the apparent meaning indicated by the context. For example, when taking context into account, a minimum reformulation of the following sentence yields 5 errors.

The following accuracy scores were calculated:

- a ratio of error free T-units per total T-units (EFT/T),
- a ratio of error free clauses per total clauses (EFC/C),
- and the total number of errors per total number of words (E/W).

C. Use of academic vocabulary

Measured by means of Academic Word List (AWL) developed by Coxhead (2000).

Counts performed:

- Numbers of AWL types
- Numbers of AWL tokens
- Number of tokens as a percentage of words written

Appendix D: Need-press questionnaire on academic literacy abilities

This questionnaire consists of two sections. Firstly, you should indicate how important you believe the academic literacy ability is for you to succeed at your 1st or 2nd year studies. Secondly, you should indicate to what extent the academic literacy ability was addressed in your academic literacy course.

	How important is this ability for you to succeed in your studies?				To which extent was this ability addressed in your academic literacy intervention?				
	This ability is...				This ability is was addressed...				
	Not important at all	Not very important	Important	Very important	Far too little	Too little	About right	Too much	Far too much
1. Listen effectively in class									
2. Take effective notes during class									
3. Take notes from reading material (such as annotating, linear outlines, mind maps)									
4. Paraphrase and summarise information									
5. Have an appropriate reading speed									
6. Use appropriate reading strategies for different goals									
7. Understand assigned reading									
8. Understand academic vocabulary									
9. Use academic vocabulary									
10. Understand subject terminology									
11. Use subject terminology									
12. Use the conventions of academic language (formality, vocabulary, exact language, objective language etc.)									
13. Use subject-specific conventions									
14. Participate in academic discussions (during and outside of class, with students and lectures, in spoken or written form)									
15. Do oral presentations in class									
16. Analyse and comprehend (know what to do) assignment and exam questions									
17. Plan a strategy for writing tasks (for exams, tests or assignments)									

	How important is this ability for you to succeed in your studies?				To which extent was this ability addressed in your academic literacy intervention?				
	This ability is...				This ability is was addressed...				
	Not important at all	Not very important	Important	Very important	Far too little	Too little	About right	Too much	Far too much
18. Structure writing (for exams, tests or assignments)									
19. Produce writing (for exams, tests or assignments)									
20. Apply relevant processes involved in academic argumentation (fact/opinion, ir/relevant information)									
21. Develop a main argument or thesis									
22. Write short coherent pieces of text									
23. Write long coherent pieces of text									
24. Interpret visual data									
25. Create visual data									
26. Integrate visual data with written work									
27. Understand underlying concepts of empirical research (including methodologies)									
28. Apply underlying concepts of empirical research (including methodologies)									
29. Use different sources for research (databases, books, scientific journals, the Internet, etc.)									
30. Process and interpret gathered data									
31. Report on gathered data									
32. Use appropriate search strategies for research purposes									
33. Reference a variety of sources (in-text [direct and indirect quoting] and bibliography)									
34. Use evidence from texts to support ideas									
35. Use evidence from texts to challenge ideas									
36. Identify relevant information									
37. Identify reliable information									
38. Synthesise (integrate) information from various sources									
39. Refer to different points of view appropriately									

	How important is this ability for you to succeed in your studies?				To which extent was this ability addressed in your academic literacy intervention?				
	This ability is...				This ability is was addressed...				
	Not important at all	Not very important	Important	Very important	Far too little	Too little	About right	Too much	Far too much
40. Use appropriate time-management strategies									
41. Use appropriate learning strategies (such as using various learning styles)									
42. Use a computer									
43. Understand the functions a computer offers (e.g. using MS Word, MS Excel, MS PowerPoint etc; creating graphs, inserting pictures etc.)									
44. Using appropriate format and layout when typing assignments									

45. Which abilities or strategies taught in the academic literacy intervention did you find most helpful in your other subjects?
46. Which abilities or strategies do you think need to be focused on more in this academic literacy intervention?
47. Which abilities or strategies do you think need to be focused on less in the academic literacy intervention?
48. Which other comments would you like to make regarding the academic literacy intervention?

This questionnaire is adapted from Van Dyk (2014).

Appendix E: Lecturer questionnaire on academic literacy abilities

In your opinion, how important is it for an academic literacy course to address the following abilities so as to ultimately assist students in succeeding in the 1st and 2nd year courses that you teach?

For students to successfully complete my course, they need support in the following academic literacy abilities:	The ability needs to be addressed			
	Not at all	Occasionally	Frequently	Extensively
1. Listen effectively in class				
2. Take effective notes during class				
3. Take notes from reading material (such as annotating, linear outlines, mind maps)				
4. Paraphrase and summarise information				
5. Have an appropriate reading speed				
6. Use appropriate reading strategies for different goals				
7. Understand assigned reading				
8. Understand academic vocabulary				
9. Use academic vocabulary				
10. Understand subject terminology				
11. Use subject terminology				
12. Use the conventions of academic language (formality, vocabulary, exact language, objective language etc.)				
13. Use subject-specific conventions				
14. Participate in academic discussions (during and outside of class, with students and lectures, in spoken or written form)				
15. Do oral presentations in class				
16. Analyse and comprehend (know what to do) assignment and exam questions				
17. Plan a strategy for writing tasks (for exams, tests or assignments)				
18. Structure writing (for exams, tests or assignments)				
19. Produce writing (for exams, tests or assignments)				
20. Apply relevant processes involved in academic argumentation (fact/opinion, ir/relevant information)				
21. Develop a main argument or thesis				
22. Write short coherent pieces of text				
23. Write long coherent pieces of text				
24. Interpret visual data				

For students to successfully complete my course, they need support in the following academic literacy abilities:	The ability needs to be addressed			
	Extensively	Frequently	Occasionally	Not at all
25. Create visual data				
26. Integrate visual data with written work				
27. Understand underlying concepts of empirical research (including methodologies)				
28. Apply underlying concepts of empirical research (including methodologies)				
29. Use different sources for research (databases, books, scientific journals, the Internet, etc.)				
30. Process and interpret gathered data				
31. Report on gathered data				
32. Use appropriate search strategies for research purposes				
33. Reference a variety of sources (in-text [direct and indirect quoting] and bibliography)				
34. Use evidence from texts to support ideas				
35. Use evidence from texts to challenge ideas				
36. Identify relevant information				
37. Identify reliable information				
38. Synthesise (integrate) information from various sources				
39. Refer to different points of view appropriately				
40. Use appropriate time-management strategies				
41. Use appropriate learning strategies (such as using various learning styles)				
42. Use a computer				
43. Understand the functions a computer offers (e.g. using MS Word, MS Excel, MS PowerPoint etc; creating graphs, inserting pictures etc.)				
44. Use appropriate format and layout when typing assignments				

45. What do you believe to be your students' academic literacy strengths and weaknesses?

46. Are there any other abilities that you believe an academic literacy course that is presented at first-year level should address?

This questionnaire is adapted from Van Dyk (2014).

Appendix F: Revised need-press questionnaire on academic literacy abilities

This questionnaire consists of two sections. Firstly, you should indicate how important you believe the academic literacy ability is for you to succeed at your 1st or 2nd year studies. Secondly, you should indicate to what extent the academic literacy ability was addressed in your academic literacy course.

	How important is this ability for you to succeed in your studies?				To which extent was this ability addressed in your academic literacy intervention?				
	This ability is...				This ability is was addressed...				
	Not important at all	Not very important	Important	Very important	Far too little	Too little	About right	Too much	Far too much
1. Listen effectively in class									
2. Take effective notes during class and from reading material									
3. Paraphrase and summarise information									
4. Have an appropriate reading speed									
5. Use appropriate reading strategies for different goals									
6. Understand assigned reading									
7. Understand and use academic vocabulary									
8. Understand and use subject terminology									
9. Use the conventions of academic language (formality, vocabulary, exact language, objective language etc.)									
10. Use subject-specific conventions									
11. Do oral presentations in class and participate in academic discussions									
12. Analyse and understand assignment and exam questions									
13. Structure writing (for exams, tests or assignments)									
14. Write short coherent pieces of text									
15. Write long coherent pieces of text									
16. Apply relevant processes involved in academic argumentation (fact/opinion, ir/relevant information)									
17. Develop a main argument or thesis									
18. Interpret visual data									
19. Create visual data and integrate visual data with written work									
20. Understand and apply the underlying concepts of empirical research (including methodologies)									

	How important is this ability for you to succeed in your studies?				To which extent was this ability addressed in your academic literacy intervention?				
	This ability is...				This ability is was addressed...				
	Not important at all	Not very important	Important	Very important	Far too little	Too little	About right	Too much	Far too much
21. Use different sources for research (databases, books, scientific journals, the Internet, etc.) and use appropriate research strategies									
22. Process, interpret and report on gathered data									
23. Reference a variety of sources (in-text [direct and indirect quoting] and bibliography)									
24. Use evidence from texts to support and challenge ideas									
25. Identify relevant and reliable information									
26. Synthesise (integrate) information from various sources									
27. Refer to different points of view appropriately									
28. Use appropriate time-management strategies									
29. Use appropriate learning strategies (such as using various learning styles)									

30. How could the academic literacy course be adapted to be of more value to you in your other subjects?

31. Which abilities or strategies taught in the academic literacy intervention did you find most helpful in your other subjects?

32. Which abilities or strategies do you think need to be focused on more in this academic literacy intervention?

33. Which abilities or strategies do you think need to be focused on less in the academic literacy intervention?

34. Which other comments would you like to make regarding the academic literacy intervention?

Appendix G: Coordinator questionnaire⁵³

1. What type of academic literacy intervention are you responsible for? (Generic academic literacy course, subject-specific academic literacy course, writing centre, or other [please specify])
-

2. How applicable would each of the following instruments be when assessing the impact of the academic literacy intervention you are responsible for? The section which explains each instrument in detail is indicated in brackets.	Not applicable	Applicable to a limited extent	Very applicable
47. Generic academic literacy test			
48. Subject-specific (LSP) academic literacy test			
49. Generic extended writing assignment			
50. Subject-specific extended writing assignment			
51. Quantitatively assessing a writing assignment			
52. Student need-pess questionnaire			
53. Lecturer questionnaire			
54. Content analysis of study material			
55. Correlating academic literacy results with other variables			
56. Qualitative feedback from primary stakeholders			

3. Are there any ways in which you believe any of the proposed instruments could be improved?
4. Are there any instruments that you would add to this evaluation design?
5. Are there any other aspects that you believe the proposed evaluation design should take into consideration, given the context of the academic literacy intervention you are responsible for?

⁵³ Additional information on all of these instruments were provided to course-coordinators in the form of addenda.

Appendix H: Revised lecturer questionnaire on academic literacy abilities

In your opinion, how important is it for an academic literacy course to address the following abilities so as to ultimately assist students in succeeding in the 1st and 2nd year courses that you teach?

For students to successfully complete my course, they need support in the following academic literacy abilities:	The ability needs to be addressed			
	Not at all	Occasionally	Frequently	Extensively
1. Listen effectively in class				
2. Take effective notes during class and from reading material				
3. Paraphrase and summarise information				
4. Have an appropriate reading speed				
5. Use appropriate reading strategies for different goals				
6. Understand assigned reading				
7. Understand and use academic vocabulary				
8. Understand and use subject terminology				
9. Use the conventions of academic language (formality, vocabulary, exact language, objective language etc.)				
10. Use subject-specific conventions				
11. Do oral presentations in class and participate in academic discussions				
12. Analyse and understand assignment and exam questions				
13. Structure writing (for exams, tests or assignments)				
14. Write short coherent pieces of text				
15. Write long coherent pieces of text				
16. Apply relevant processes involved in academic argumentation (fact/opinion, ir/relevant information)				
17. Develop a main argument or thesis				
18. Interpret visual data				
19. Create visual data and integrate visual data with written work				
20. Understand and apply the underlying concepts of empirical research (including methodologies)				
21. Use different sources for research (databases, books, scientific journals, the Internet, etc.) and use appropriate research strategies				
22. Process, interpret and report on gathered data				
23. Reference a variety of sources (in-text [direct and indirect quoting] and bibliography)				
24. Use evidence from texts to support and challenge ideas				
25. Identify relevant and reliable information				
26. Synthesise (integrate) information from various sources				
27. Refer to different points of view appropriately				
28. Use appropriate time-management strategies				
29. Use appropriate learning strategies (such as using various learning styles)				

30. What do you believe to be your students' academic literacy strengths and weaknesses?

31. Are there any other abilities that you believe an academic literacy course that is presented at first-year level should address?

Appendix I: Revised writing rubric

		1 (Poor)	2 (Below average)	3 (Average)	4 (Good)	5 (Excellent)	N/A
Structure and development							
1.	Introduction	No attempt at an introduction is evident.		An attempt at an introduction is evident.		The introduction effectively introduces the topic, grabs the reader's interest, and indicates what will be discussed in the text.	
2.	Thesis statement	No thesis statement is evident.		A thesis statement is present, but it is not as clear as it could be, and/or not sufficiently proven by means of argument.		A clear and logical thesis statement is provided that is proven by means of argument.	
3.	Development of main argument	No argument is made in this text.		The text contains an argument, though this is not always clear and logical.		A clear and logical argument is built throughout the text.	
4.	Paragraph development (topic sentences, main ideas, supporting information)	Paragraphs do not have topic sentences. Paragraphs do not centre around one main idea. The information in the paragraph does not support the topic sentence.		Many paragraphs have clear topic sentences, centre around a single main idea, and contain supporting information that relates to the topic sentence.		All paragraphs have clear topic sentences, centre around a single main idea, and all supporting information relates to the topic sentence.	
5.	Relevance of content to topic	Content is not relevant to the topic.		Content is mostly relevant to the topic.		Content is highly relevant to the topic.	
6.	Linking devices (structuring thought with discourse markers, pronouns etc.)	No attempt is made at using linking devices effectively.		An attempt at using linking devices effectively is evident throughout the text.		Linking devices are used logically and appropriately throughout the text.	
7.	Conclusion	No attempt at a conclusion is evident.		An attempt at a conclusion which summarises the main points of the text, ties ideas together, and finishes off the text is evident.		The conclusion succinctly and effectively summarises the main points of the text, ties ideas together, and finishes off the text.	
Academic writing style							
8.	Syntax: phrase and clause structure	Numerous incomplete sentences and incorrect clauses are evident.		Complete sentences with correctly formulated clauses are generally evident.		Complete sentences with correctly formulated clauses are used throughout.	
9.	Academic vocabulary	No attempt is made to correctly make use of academic vocabulary.		Some attempt is made to correctly make use of a range of academic vocabulary.		A wide range of academic vocabulary is used correctly throughout the text.	

10.	Technical vocabulary	No attempt is made to correctly make use of technical vocabulary.		Some attempt is made to correctly make use of a range of technical vocabulary.		A wide range of technical vocabulary is used correctly throughout the text.	
11.	Style (formality; rhetorical mode)	The writing style does not reflect the level of formality of the text.		The writing style generally reflects the level of formality of the text.		The writing style excellently reflects the level of formality of the text.	
12.	Genre	The writing style is not appropriate for the intended genre.		The writing style is partially appropriate for the intended genre.		The writing style suits the intended genre perfectly.	
13.	Audience	The writing style is not appropriate for the intended audience.		The writing style is partially appropriate for the intended audience.		The writing style suits the intended audience perfectly.	
14.	Integration of visual data	Visual data are not integrated effectively at all.		An attempt to integrate visual data into the text is clear.		Visual data are fully and effectively integrated with the text, using correct captions with clear references to the visual data throughout the text.	
Editing							
15.	Spelling, capitalisation and punctuation	Spelling, capitalisation and punctuation errors obstruct the meaning of the text.		Spelling, capitalisation and punctuation are generally accurate.		Spelling, capitalisation and punctuation are accurate throughout the text.	
16.	Concord and tense	Numerous concord and tense errors are evident in the text.		Some concord and tense errors are evident in the text.		No concord and tense errors are evident in the text.	
17.	Layout and format	The format requirements have not been met.		Format requirements are generally adhered to.		Format requirements are strictly adhered to. The format is polished and professional.	
Use of source material							
18.	Referencing technique (list of references)	The referencing is absent or unsystematic, and the prescribed referencing system is not used.		The referencing is generally accurate using the prescribed referencing system.		The referencing is consistently accurate using the prescribed referencing system.	
19.	Relevance of source data	Source data are not relevant to the topic.		Source data are generally relevant to the topic, and indicate sufficient research.		Source data are very relevant to the text, and indicate extensive research.	
20.	Appropriately citing direct and indirect quotations	Quotations are not cited in the text.		Quotations are mostly cited in the text using the prescribed referencing system.		Quotations are consistently and accurately referenced in the text using the prescribed referencing system.	
21.	Paraphrasing information from source texts	Information from source texts is not paraphrased.		A clear attempt was made to paraphrase most of the information from source texts.		Information from source texts are paraphrased correctly throughout the text.	
22.	Integration of source data with text (synthesising)	No attempt to synthesise source data with the author's argument is evident.		A clear attempt to synthesise source data with the author's argument is evident.		Source data are excellently synthesised with the author's argument throughout the text.	

Appendix J: Rubric for oral presentations

ASSESSMENT CRITERIA	PERFORMANCE INDICATOR LEVELS		
	0	1	2
1. Introduction	No introduction	An attempted introduction	Stimulating introduction, with clear links to content
2. Body of presentation / cohesion	Muddled; unclear (not coherent)	Some attention paid to the coherence of the presentation	Coherent, clear and to the point
3. Conclusion	No conclusion	An attempted conclusion	Conclusion effectively highlighting key points
4. Preparation	Not prepared	Partially prepared	Fully prepared and rehearsed
5. Support material	None	Some support material used	Comprehensive and fully integrated with content
6. Use of voice and pace	Voice too soft / pace much too slow or fast	Generally acceptable voice projection and pace	Excellent voice projection; pace is just right
7. Enthusiasm/ effort	Little effort evident and lack of enthusiasm	Some effort and enthusiasm evident	Excellent effort and enthusiasm
8. Contents	Inadequate and/or often inappropriate	Sufficiently relevant content based on research	Comprehensive, relevant, evidence of extensive research
9. Responses to questions	No answers to questions from audience members provided	A sufficient attempt at answering questions posed from audience members	Can answer interpretation and application questions, and respond appropriately if questions cannot be answered
10. Overall impression	Little effort was made with this presentation	Some effort was made with this presentation	A lot of effort was made with this presentation

(Fouché & Immelman, 2015)