

Impact measurement: towards creating a flexible evaluation design for academic literacy interventions

Ilse Fouché

University of Pretoria
E-mail: ilse.fouche@up.ac.za

Tobie van Dyk

Centre for Academic and Professional Language Practice, North-West University
E-mail: tobie.vandyk@nwu.ac.za

Gustav Butler

School of Languages, North-West University
E-mail: gustav.butler@nwu.ac.za

Abstract

Considering the vast array of academic literacy interventions that are presented both nationally and internationally, and the resources required to present these interventions, it is becoming increasingly important for those who are responsible for these interventions to provide evidence of their impact. The aim of this paper is 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 are 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 is then proposed that could be used in evaluating the impact of a range of academic literacy interventions. Avenues to explore in future include testing the design in the South African context.

Keywords: academic literacy, programme evaluation, impact measurement, effect, interventions

1. 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 (Davies 2010: xi; Cliff 2014: 322; Sebolai 2014: 52). What is still largely lacking, though, is evidence of the effectiveness (or lack thereof) of such interventions, and the

extent of the impact they have (Holder, Jones, Robinson and Krass 1999: 20; Carstens and Fletcher 2009a: 319; Storch and Tapper 2009: 218; Terraschke and Wahid 2011: 174; Butler 2013: 80; Sebolai 2014: 52). For the purposes of the current study, impact (or effect) is defined as “1) the observable improvement in academic literacy abilities between the onset and the completion of an academic literacy intervention and 2) the extent to which these abilities are necessary and applied in students’ content subjects” (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 Fouché (2015), even though some attempts to assess the impact of academic literacy interventions have been made, the type and variety of research instruments used 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. Reasons for this include that, typically, too few instruments have been used in most studies conducted thus far to make valid conclusions regarding the impact of modules, that instruments used thus far have been varying to such a degree that it might be difficult for future researchers to decide which to adopt, and that they are often designed for local contexts and are not necessarily as robust as one would like in order to make valid and reliable conclusions, which can be compared to those of similar interventions elsewhere.

However, effectively assessing the impact of an academic literacy intervention poses challenges. A considerable challenge is the wide variety of academic literacy interventions offered not only in South Africa, but also elsewhere. These interventions include, but are not limited to, generic interventions (e.g. Van Wyk 2002; Weideman 2003), subject-specific interventions (e.g. Goodier and Parkinson 2005; Ngwenya 2010; Carstens 2013b), 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)¹, writing centres (e.g. Archer 2008), and reading interventions (e.g. Van Wyk and 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 address such a variety of interventions effectively, as each has 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 all) 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 have an impact on students’ academic literacy levels. Yet, increasingly, South African higher education students are expected to participate in 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 3) is generally seen as a desirable method of determining the validity of findings (Lynch 1996: 59–61, 2003: 152), in cases where control groups are not possible, triangulation becomes vital in reliably and validly assessing the impact of academic literacy interventions.

¹ See Van de Poel and Van Dyk (2015: 169–173) for a more in-depth discussion of collaborative approaches.

The current article forms part of a larger study which aims to develop and test an evaluation design that could be used to assess the impact of a wide variety of academic literacy interventions. A previous article (Fouché 2015) provides an overview of studies that have aimed to evaluate the impact of a range of academic literacy interventions. That article discusses each study by considering its research methodology, results, strengths and weaknesses. Based on the findings from the literature, broad suggestions of possible useful instruments for impact evaluation are made. The aim of the current article is to build on the literature surveyed in Fouché (2015), and to propose a conceptual evaluation design that could be used to assess the effectiveness of various academic literacy interventions based on that literature. The findings from Fouché (2015) will not be repeated in the current article – rather, for a thorough overview of instruments that have been used to assess the impact of academic literacy interventions, these two articles should be considered in tandem. Future articles will empirically verify this conceptual evaluation design, and the results from these articles will further inform the final evaluation design.

Section 2 briefly describes the variety of instruments that have been used in previous studies to assess the impact of academic literacy interventions; whereas Fouché (2015) focuses on the studies that have been conducted in the field of impact assessment, section 2 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 3, an evaluation design is proposed through which a combination of these instruments could be used to assess various academic literacy interventions.

2. Instruments for determining the impact of academic literacy interventions

This study argues that impact of interventions 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 article, 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 paper is based on strategies that have been followed in previous studies (discussed in Fouché 2015). However, it will attempt to overcome the weaknesses that were 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) and by data sources (for example, obtaining data from students, lecturers and the literature) (Denzin 1989: 237; Lynch 1996: 59). 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, Alderson (1992: 285) argues that:

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.

Lynch (1996: 60) agrees 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” (cf. Judd and Keith 2012: 40). Therefore, by examining a variety of factors that may shed light on the agency of an academic literacy intervention in the ultimate improvement of students’ academic literacy abilities, and the extent to which these were transferred to other subjects, a more valid inference can be made regarding the causal relationship between such improvement and the academic literacy intervention.

2.1 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 article) 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 ones seemed to be unchanged after the completion of the intervention. Thus, this section addresses the first aspect of the definition of impact as given in section 1.

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 (cf. e.g., Hughes 2003: 1–2) as the course is refined. This is a process that 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, Brown, Elder, Hill, Lumley and Mcnamara 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 3. 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 competence is to be measured, a portfolio might be necessary (Hughes 2003: 87), as this would adequately show the student's full range of competence, 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 (Shaw and Liu 1998: 228–230; Carstens and Fletcher 2009a: 322; Storch and Tapper 2009: 210). Furthermore, the assessment must be relevant to its context to assess a set of outcomes appropriately. 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 an 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 instruments proposed in this article.

This paper 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 (cf. Fouché 2015). 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.

2.1.1 Instrument 1: Reliable, validated generic academic literacy tests

Several reliable, validated academic literacy tests (cf. Davies et al. 1999: 168, 221) are widely used to assess students' academic literacy abilities, both before and after interventions. These include the Test of Academic Literacy Levels (TALL) (Van Dyk 2005), the literacy section of the National Benchmark Test (NBT) (Yeld 2010: 29), the Test of Academic Literacy for Postgraduate Students (TALPS) (Butler 2009), the Placement Test in English for Educational Purposes (PTEEP) (Cliff and Hanslo 2009: 268–274) and the English Literacy Skills Assessment for Higher Education and Training (ESLA-Plus) (Van Dyk and Weideman 2004: 3). They 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 and 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 that is to be expected if the assumption is made that the academic literacy intervention would have a significant impact on student success (Van Dyk 2010: 272; Van Dyk 2015: 180; Van Rooy and Coetzee-Van Rooy 2015: 10).

Such tests have several advantages. These include that they make it possible to mark large quantities of tests easily and make replicating research easier than would be the case with assessments that have not been validated and verified. In addition, such tests generally have fixed and uniform procedures, materials, scoring instruments and administration that ensure that assessment will occur in the same manner at different places and times. Furthermore, these tests have undergone rigorous developmental cycles (Beretta 1986: 434; Heaton 1988: 167; Nitko and Brookhart 2011: 346; Seabi 2012: 93). 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)². 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 (cf. section 2.1.2 for a further discussion of this type of test).

As was mentioned earlier, several reliable and valid academic literacy tests could be used for the purpose of assessing students' academic literacy proficiency. This study uses characteristics of the TALL to illustrate the typical strengths and weaknesses of this type of instrument. The TALL is considered an adequate and appropriate test in the South African context for various reasons. Firstly, it has a theoretically sound test construct that takes the most important aspects of academic literacy into consideration, and 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 Bachman and Palmer's (1996: 75–76) suggestion into consideration 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 reliable 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: 155). Thirdly, the TALL is viewed as a valid instrument (Van Der Walt and Steyn 2007: 146–156; Van Dyk 2010: 285–286), which is standardised in terms of its scoring and administration (Petersen-Waughtal and 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 and Coetzee-Van Rooy 2015: 8, 13). Finally, it is an efficient test choice from a logistical stance. The TALL is a 60-minute 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, as well as 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.

² In reality, of course, the concept of “universal questions” is not always quite so straightforward.

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 further 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.

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

The limitations mentioned in this section can be addressed by triangulating data through multiple sources of evidence. LSAP tests are one such an additional source of evidence that could be used to make more valid inferences with regard to the possible transfer of abilities to students' content subjects.

2.1.2 Instrument 2: LSAP tests

Despite the advantages of generic academic literacy tests, evaluators might feel the need to use tests that reflect how academic literacy abilities are contextually applied in students' disciplines. LSAP tests deal with relevant academic subjects (Krekeler 2006:104). These tests might use texts from various subjects, such as business studies, engineering and medicine (Fulcher 1999; Krekeler 2006). 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 be a more reliable indicator of 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 2.1.1, as by 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 tests. Furthermore, these tests are difficult to create, since they demand more resources and skills than generic tests (Hamp-Lyons and Lumley 2001: 129). Examples of these resources and skills are more manpower and time (with concomitant financial implications) as several discipline-specific tests need to be created

instead of one generic test, and staff members with English for specific purposes (ESP) experience and knowledge are required.

It should be noted that some researchers (Davies 2001: 144; Elder 2001: 164)³ have questioned the use of ESP tests 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 English for academic purposes (EAP) tests in assessing students' command of academic language competence (cf. e.g., 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 (cf. section 2.1.1). This could be addressed if the researcher can justify, empirically or theoretically, that the pre- and post-tests are equivalent.

If the evaluator decides 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:

- ensuring that experts in the content-subject field are involved in setting the test (Gnutzman 2009: 530);
- ensuring 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 and Palmer 1996: 19–20);
- having an appropriate test construct that does not leave out important areas addressed in the intervention, or add aspects that were not the focus of the intervention (Bachman and Palmer 1996: 21; Fulcher 1999: 225–226; Hamp-Lyons and Lumley 2001: 131);
- using 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 (Fulcher 1999: 222; Douglas 2001: 176; Hamp-Lyons and Lumley 2001: 131);
- taking the purpose of the test into consideration (Hamp-Lyons and Lumley 2001: 130);
- ensuring 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 and Palmer 1996: 23–24]);
- using 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) (Douglas 2000: 16–18; Hamp-Lyons and Lumley 2001: 131);
- ensuring 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

³ See Weideman (2013) for similar arguments regarding ESP courses.

- students not only display knowledge of the language, but that they also show they can use the language in context (Bachman and Palmer 1996: 25);
- considering 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 and Palmer 1996: 29);
 - keeping practicality in mind – thus, consider how many resources are required by the test, and how many are available in reality (Bachman and Palmer 1996: 35);
 - not assuming previous knowledge in the specific field and taking steps to ensure 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
 - piloting the test (Fulcher 1999: 226; Douglas 2000: 5–6; Krekeler 2006: 106).

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. Data should also preferably be used as part of a larger triangulated study in which additional sources of evidence are used to verify and validate results.

Many academic literacy tests have limitations though. Firstly, these tests usually do not expect students to, for example, produce language, build an argument, or communicate research – often due to practical constraints in the administration of the test (Butler 2009: 294). Butler (2013: 83) cautions that an improvement in writing might not necessarily be reflected in such tests. Secondly, academic literacy 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 put forward is that multiple choice questions are not able to test students' normal language processing abilities; in other words, the processes that students engage in when processing information and producing language, for example, in written form (Murphy and Yancey 2008: 366–368; Scouller 2012: 469). 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 cross-validate findings. A combination of assessment tools can also assist the researcher to view impact from various angles, which could lead to a more holistic impact measurement. Therefore, it is worthwhile to consider additional methods of assessment if the impact of an academic literacy intervention is to be measured comprehensively.

2.1.3 Instrument 3: Assessing writing assignments by means of a rubric

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” (cf. 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 inquiry in the field.

Researchers recommend that writing assessments adhere to various principles (summarised in figure 1). Firstly, students should be given a clearly defined topic that will motivate them to write and guide them in their writing endeavour (Heaton 1988: 137, 144; Weigle 2002: 53, 93). 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 the content area 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 and 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 that 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 (Heaton, 1988: 137; Weigle 2002: 52). 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 (Heaton 1988: 138; Weigle 2002: 52) – 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 (Heaton 1988: 148; Weigle 2002: 108–139). Rubrics attempt to separate various features of written text for scoring purposes (Heaton 1988: 148; Weigle 2002: 122). 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.

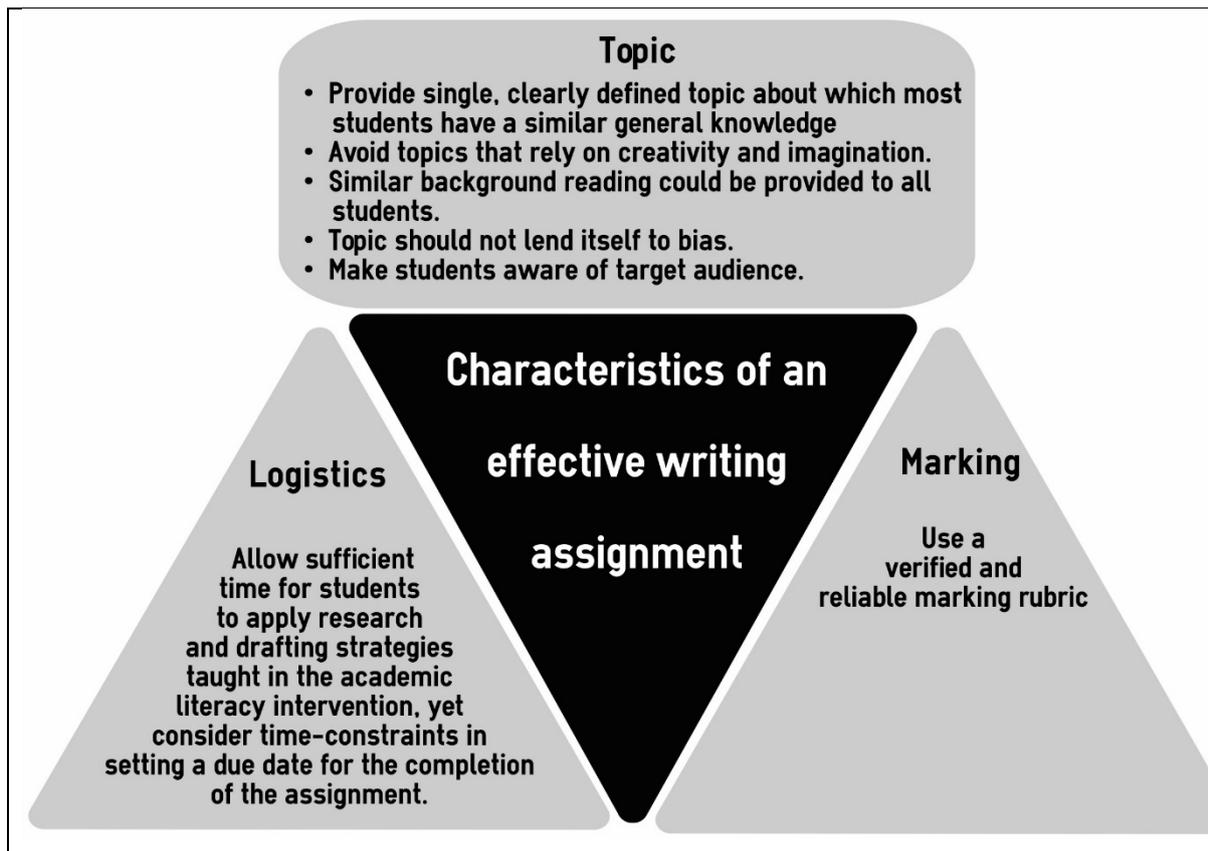


Figure 1: Characteristics of an effective writing assessment

Keeping in mind the principles discussed in this section, this study proposes that an extended⁴ 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, an essay pre- and post-assignment could be used. In such a case, a 300-word text of argumentative writing (cf. e.g., Laufer and Nation 1995: 314) on the same or on a similar topic could be written at the beginning of the intervention and at the end of the intervention. The conditions of the pre- and post-assignments would be similar, so as to exclude as many external variables as possible. Students could be provided with a topic and sources they could use to support their argument. As mentioned earlier though, a framework for assessing the impact of an academic literacy intervention should be flexible. Although argumentative writing is a commonly used genre in academic literacy courses and university writing in general (Yeh 1998: 49–51; Hyland 2004: 18, 122; Paltridge 2004: 87; Zhu 2009: 34), 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

⁴ “Extended” here refers to an assignment that is longer than a typical paragraph answer. For the purposes of this article, a 300-word assignment is considered “extended” at first-year level. However, longer assignments might be required for more advanced students.

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 A. 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 encompasses a wide variety of general academic literacy outcomes under the headings “structure and development”, “academic writing style”, “editing” and “use of source material”. A strength of this rubric is that the various sections are divided into subsections, so that it is clear in which specific areas students improved. However, it can also be argued that these subsections represent distinct skills, and that academic literacy should not be compartmentalised in this way. A further strength is that the combined essay marks, and 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 (1998: 29–31; 2001: 40–44) C3 model (refer to Appendix B). 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 is presented in Carstens and Van de Poel (2012: 58–79). A possible strength of this rubric is that it does not compartmentalise academic literacy into specific skills to the same extent as in the previous rubric. A potential weakness, on the other hand, is that it might be more difficult to determine how a curriculum should be improved based on the feedback from this rubric, as several categories are quite broad.

2.1.4 Instrument 4: Assessing writing assignments by means of quantitative measures

An additional option when assessing students’ writing assignments 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 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 Word List).

Accuracy counts can be done by coding various categories of mistakes and indicating these by means of comments in Microsoft Word. Frequency counts can then be done by means of the “find” function. Academic vocabulary counts using Coxhead’s (2000) Academic Word List (AWL) can be done by using a programme called Vocabprofile developed by Heatley, Nation and Coxhead (2002) and adapted by Cobb (2015). Note that some criticism has arisen against the AWL, for instance that it is not ideal for use in various disciplines, and it is biased in favour of disciplines such as business and law at the expense of fields such as the natural sciences (Read 2007: 109). Therefore, where discipline-specific word lists are available, these might be more appropriate. However, the Coxhead (2000) list is still an excellent resource where discipline-specific lists are not available. Appendix C provides more information on the methods proposed by Storch and Tapper (2009: 210–212).

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

The second aspect of the definition of impact given in section 1, 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 Fouché (2015), 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, Fouché, Strydom and Delpont 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)⁵. 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 lecturers⁶.

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 matter that permits quantification and comparison with other programs or with other features of the same 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 2.2.5.

In addition to using questionnaires (discussed in more detail in sections 2.2.1 and 2.2.2), some valuable information regarding which academic literacy abilities are necessary and applied in students' content-subjects could be obtained by analysing students' content-subject study guides, tests and examinations (cf. section 2.2.3), correlating academic literacy results to other variables (cf. section 2.2.4) and obtaining qualitative feedback from primary stakeholders (see section 2.2.5). 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.

⁵ However, such information should be combined with empirical evidence regarding an improvement in academic literacy abilities, as discussed in section 2.1. Only when considering multiple sources of evidence can valid inferences be drawn.

⁶ Using additional sources of evidence, such as an analysis of study guides or diagnostic test data, would be valuable in corroborating such results.

2.2.1 Instrument 5: 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. They argue that the success of such interventions 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 skills are transferred. It is particularly important to determine whether students consciously transfer abilities acquired in the academic literacy classroom to their content subjects⁷. One way of doing that is by asking students about their behaviour. 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 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 and open-ended descriptive questions in a student questionnaire, which allows the collection of both quantitative and qualitative data.

Some studies use in-house official feedback forms to assess students' perceptions regarding academic literacy interventions (cf. e.g. Van Dyk, Cillié, Coetzee, Ross and Zybrands 2011: 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 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 of transfer is a “need-press” questionnaire (Lynch 2003: 68; Rao and Saxena 2014: 22–25). 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

⁷ However, it is 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.

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 adapted 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 telling finding of this study, namely that perceived preparedness does not reflect actual student preparedness (Van Dyk et al. 2013: 363), should serve as a warning that students' perceptions are not necessarily always valid, and thus must be used as part of a triangulated study. 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.

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 such cases, 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.

2.2.2 Instrument 6: 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 additional sources.

One way of doing so, 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 the intervention could potentially have.

This study proposes that the adapted 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 adapted questionnaire can be found in Appendix E. This questionnaire also contains Likert scale and open-ended questions, so as to ensure that not only empirical, but also descriptive data⁸ are gathered (cf. Lynch 1996 discussed in section 2.2.1). In fact, this questionnaire is very similar to the student questionnaire discussed in 2.2.1, and purposefully so – the researchers believe it to be very important that these two questionnaires correspond with one another, so that findings can easily and validly be compared in a triangulated study.

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 to take place. However, conclusions regarding transfer could be strengthened by using some of the instruments discussed below.

2.2.3 Instrument 7: Content analysis of study material

Another option of determining which academic literacy abilities students need to use in their content subjects, in addition to the ones above, would be to do a content analysis of students' study material, specifically their assignments, tests and examinations, as in Fouché (2009: 83–101). In this study, the assignments for Unisa's Science Foundation Programme were analysed. First, appropriate subjects (i.e. those that required some writing as part of their assignments) were identified. Assignments were then categorised based on Bloom's Taxonomy (Fouché 2009: 95–100), which differentiates between the following cognitive levels: knowledge, comprehension, application, analysis, synthesis, and evaluation. 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. That was indeed the case in the Fouché (2009) study, where it became evident that first-year students of Unisa's science faculty are rarely required to work at cognitive levels beyond the comprehension level.

However, 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 analysis. Therefore, this instrument might 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

⁸ Section 2.2.5 discusses further methods of gathering qualitative data from stakeholders.

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.

Due to these limitations, 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 might be refined 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.

2.2.4 Instrument 8: 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: 10–11), among others, correlated the results of an academic literacy course with students' general academic achievement (Van Dyk 2010: 262–274; Mhlongo 2014: 80–82; Van Dyk 2015: 178–180). 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. However, 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 article). 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, while 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 could be drawn between students' performance in an academic literacy intervention and their class attendance (cf. 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 2.1.

2.2.5 Instrument 9: 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, Wright, Birch and Jacobs (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 and 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, Zybrands, Cillié and Coetzee (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 students' writing.

3. 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.

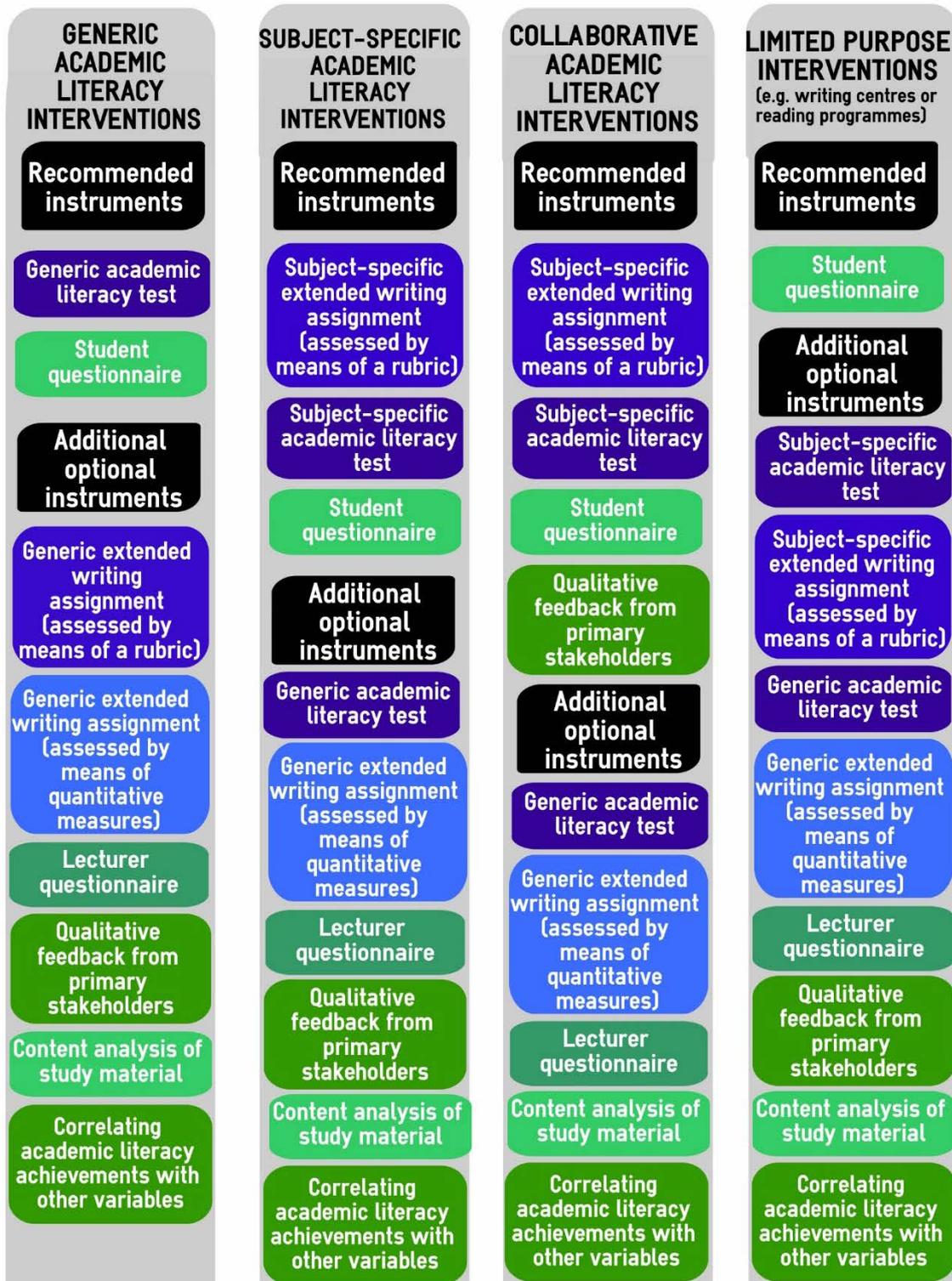
As argued by Judd and Keith (2012: 36), using a valid framework (or evaluation design) is possibly the best way of ensuring that change (for example in students’ academic literacy levels over a period of time, and the transference thereof) is indeed due to a particular agent (for example an academic 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 (Jick 1979: 602; Lynch 1996: 59–60). 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 (Lynch 1996: 59–61; Judd and Keith 2012: 40).

Based on the approaches discussed in the previous section and the definition of impact that was put forward in the introduction of this article, the current study proposes that 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”. As mentioned before, since academic literacy interventions 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 core components that would address integral aspects that should be part of each academic literacy intervention. However, the researcher would have to be able to adapt some research tools so as to assess the impact of each individual academic literacy intervention most effectively, as not all academic literacy interventions have the same foci or objectives.

Based on the variety of potential evaluation instruments discussed above, the evaluation design depicted in figure 2 is proposed as a tool 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 and Van Dyk 2015: 169–173). The current study proposes using at least three 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 four 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.

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 and Van Dyk, 2015: 169–173). Ultimately, it is the responsibility of the researcher to ensure that the most comprehensive combination of instruments is used so as to strengthen inferences regarding the impact of the intervention as far as is possible, while acknowledging any limitations that might still remain in the research design.



- 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 2: Proposed evaluation design for academic literacy interventions

4. 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 article has aimed to contribute to the discussion regarding the impact and accountability of academic literacy interventions by considering eight instruments that 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 that determine which academic literacy abilities were needed in students’ content subjects and whether these abilities were transferred to students’ content subjects. Finally, a flexible evaluation design was proposed that could be used to validly assess the impact of a range of academic literacy interventions.

In future articles, 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, 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 to be effectively applied to their specific contexts. A final evaluation design will then be proposed. 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).

References

- Alderson, J.C. 1992. Guidelines for the evaluation of language education. In C.J. Alderson and A. Beretta (Eds.) *Evaluating second language education*. Cambridge: Cambridge University Press. pp. 274–304.
- Archer, A. 2008. Investigating the effect of Writing Centre interventions on student writing. *South African Journal of Higher Education* 22(2): 248–264.
- Bachman, L.F. and A.S. Palmer. 1996. *Language testing in practice: Designing and developing useful language tests*. Oxford: Oxford University Press.
- Beretta, A. 1986. Program-fair language teaching evaluation. *TESOL Quarterly* 20(3): 431–444.
- 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.
- 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. Available online: <http://repository.up.ac.za/handle/2263/24689> (Accessed 7 July 2013).

- 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. and L. Fletcher. 2009a. Evaluating the effectiveness of a cross-disciplinary genre-focused writing intervention. *Journal for Language Teaching* 43(1): 54–65.
- Carstens, A. and L. Fletcher. 2009b. Quantitative evaluation of a subject-specific essay-writing intervention. *Southern African Linguistics and Applied Language Studies* 27(3): 319–332.
- Carstens, W. and K. Van de Poel. 2012. *Teksredaksie*. Stellenbosch: African Sun Media.
- 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. and M. Hanslo. 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.
- Cobb, T. 2015. *Web Vocabprofile*. Available online: <http://www.lex Tutor.ca/vp/eng/> (Accessed 5 March 2015).
- Coxhead, A. 2000. A new academic wordlist. *TESOL Quarterly* 34(2): 213–238.
- 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., A. Brown, C. Elder, K. Hill, T. Lumley and T. Mcnamara. 1999. *Dictionary of language testing*. Cambridge: Cambridge University Press.
- De Vos, A.S., C.B. Fouché, H. Strydom and C.S.L. Delport. 2011. *Research at grassroots for the social sciences and human service professions*. Pretoria: Van Schaik.
- Denzin, N.K. 1989. *The research act: A theoretical introduction to sociological methods, 3rd ed.* Englewood Cliffs: Prentice Hall.
- 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 D. Belcher and G. Braine (Eds.) *Academic writing in a second language: Essays on research and pedagogy*. Norwood: Ablex Publishing Corporation. pp. 298–312.

Elder, C. 2001. Assessing the language proficiency of teachers: are there any border controls? *Language Testing* 18(2): 149–170.

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. Available online: <http://repository.up.ac.za/handle/2263/26500> (Accessed 7 July 2013).

Fouché, I. 2015. Towards impact measurement: An overview of approaches for assessing the impact of academic literacy abilities. *Stellenbosch Papers in Linguistics* 44: 1–17.

Fulcher, G. 1999. Assessment in English for academic purposes: Putting content validity in its place. *Applied Linguistics* 20: 221–236.

Gnutzman, C. 2009. Language for specific purposes vs. general language. In K. Knapp, B. Seidlhofer and H. Widdowson (Eds.) *Handbook of foreign language communication and learning*. New York: Mouton de Gruyter. pp. 517–544.

Goodier, C. and J. Parkinson. 2005. Discipline-based academic literacy in two contexts. *Journal for Language Teaching* 39(1): 66–79.

Grabe, W. 2009. Teaching and testing reading. In M.H. Long and C.J. Doughty (Eds.) *The handbook of language teaching*. Chichester: Wiley-Blackwell. pp. 441–462.

Green, A. 2005. EAP study recommendations and score gains on the IELTS academic writing test. *Assessing Writing* 10(1): 44–60.

Hamp-Lyons, L. and T. Lumley. 2001. Assessing language for specific purposes. *Language Testing* 18(2): 127–132.

Hatry, H.P. and K.E. Newcomer. 2010. Pitfalls in evaluations. In J.S. Wholey, H.P. Hatry and K.E. Newcomer (Eds.) *Handbook of practical program evaluation*. 3rd ed. San Francisco: Jossey-Bass. pp. 557–580.

Heatley, A., I.S.P. Nation and A. Coxhead. 2002. *RANGE and FREQUENCY programs*. Available online: <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, G. 1987. *A guide to language testing: Development, evaluation, research*. Cambridge: Newbury House.

- Holder, G.M., J. Jones, R.A. Robinson and I. Krass. 1999. Academic literacy skills and progression rates amongst pharmacy students. *Higher Education Research and Development* 18(1): 19–30.
- Hughes, A. 2003. *Testing for language teachers, 2nd ed.* Cambridge: Cambridge University Press.
- Hyland, K. 2004. *Disciplinary discourses: Social interactions in academic writing.* Michigan: University of Michigan Press.
- 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.
- Judd, T. and B. Keith. 2012. Student learning outcomes assessment at the program and institutional levels. In C. Secolsky and D.B. Denison (Eds.) *Handbook on measurement, assessment, and evaluation in higher education.* New York: Routledge. pp. 31–46.
- 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.
- Laufer, B. and P. Nation. 1995. Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics* 16(3): 307–322.
- Lynch, B.K. 1996. *Language program evaluation: Theory and practice.* Cambridge: Cambridge University Press.
- Lynch, B.K. 2003. *Language assessment and programme evaluation.* Edinburgh: Edinburgh University Press.
- 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. Available online: <http://dspace.nwu.ac.za/handle/10394/13174> (Accessed 7 July 2015).
- Murphy, S. and K.B. Yancey. 2008. Construct and consequence: Validity in writing assessment. In C. Bazerman (Ed.) *Handbook of research on writing: History, society, school, individual, text.* Madison: University of Wisconsin Press. pp. 365–385.
- 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. and S.M. Brookhart. 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. 2004. Academic writing. *Language teaching* 37(2): 87–105.
- Petersen-Waughtal, M. and T. Van Dyk. 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.
- Rao, T.V. and S. Saxena. 2014. *Measuring and managing institutional environment of institutes of professional education*. Available online: <http://iimahd.ernet.in/assets/snippets/workingpaperpdf/14346925982014-03-05.pdf> (Accessed 27 April 2014).
- Read, J. 2007. Second language vocabulary assessment: Current practices and new directions. *International Journal of English Studies* 7(2): 105–126.
- 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 D. Janssen and R. Neutelings (Eds.) *Reading and writing public documents: problems, solutions and characteristics*. Amsterdam: John Benjamins. pp. 37–57.
- 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 J.G. Maree (Ed.) *Complete your thesis or dissertation successfully: Practical guidelines*. Claremont: Juta & Company. pp. 81–95.
- 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. and E.T. Liu. 1998. What develops in the development of second-language writing? *Applied Linguistics* 19(2): 225–254.
- Shaw, S.D. and C.J. Weir. 2007. *Examining writing: research and practice in assessing second language writing*. Cambridge: Cambridge University Press.
- Storch, N. and J. Tapper. 2009. The impact of an EAP course on postgraduate writing. *Journal of English for Academic Purposes* 8(3): 207–223.
- Terraschke, A. and R. Wahid. 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.
- Van de Poel, K. and T. Van Dyk. 2015. Discipline-specific academic literacy and academic integration. In R. Wilkinson and M.L. Walsh (Eds.) *Integrating Content and Language in Higher Education*. Frankfurt am Main: Peter Lang. pp. 161–180.

- Van Der Walt, J.L. and H. Steyn. 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*. Unpublished PhD thesis. Bloemfontein: University of the Free State. Available online: <http://scholar.ufs.ac.za:8080/xmlui/handle/11660/1918> (Accessed 7 July 2013).
- 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., K. Cillié, M. Coetzee, S. Ross and H. Zybrands. 2011. Ondersoek na die impak van 'n akademiese geletterdheidsintervensie op eerstejaarstudente se akademiese taalvermoë. *LitNet Akademies* 8(3): 487–506.
- Van Dyk, T., K. Van de Poel and F. Van der Slik. 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. and A. Weideman. 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., H. Zybrands, K. Cillié and M. Coetzee. 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. and S. Coetzee-Van Rooy. 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 and 4): 220–232.
- Van Wyk, A. & W. Greyling. 2008. Developing reading in a first-year academic literacy course. *Stellenbosch Papers in Linguistics* 38: 205–219.
- Weideman, A. 2003. Justifying course and task design in language teaching. *Acta Academica* 35(3): 26–48.
- 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.

Winberg, C., J. Wright, B.W. Birch and C. Jacobs. 2013. Conceptualising linguistic access to knowledge as interdisciplinary collaboration. *Journal for Language Teaching* 47(2): 89–107.

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. 2010. Some challenges and responses: Higher education in South Africa. *Discourse* 38(1): 24–36.

Zhu, W. 2009. Performing argumentative writing in English: Difficulties, processes, and strategies. *TESL Canada Journal* 19(1): 34–50.

Appendices

Appendix A: Example of a writing rubric

		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 (1998: 29–31; 2001: 40–44) 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 writing⁹.

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.’, or ‘Looking at its [this] background and current situation.’
13. 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.
14. Prepositions (missing or redundant)

⁹ Text structure and rhetorical quality are also investigated in the Storch and Tapper study using a more traditional rubric. However, the current article is primarily interested in the quantitative measures employed in the Storch and Tapper study. As rubrics are dealt with separately in the current article, the section on text structure and rhetorical quality is not included in this summary.

Lexical

15. Word choice. (Register errors such as ‘lots of’ were not included); e.g. ‘Many countries still out of [lack] responsibility’, or ‘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’.
16. Collocation. Erroneous expressions and phrasal verbs were counted as one error; e.g. ‘the key of the [to] success’, or ‘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)

17. Capitalisation
18. 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 five errors. Original text: Communication is a critical field for a successful project manager, how need to communicate his customers. Minimal reformulation: Communication is a critical field for a successful project manager, who (1) needs (2) to know (3) how (4) to communicate with (5) his customers. (The phrase was not reformulated as “who needs to communicate with his customers”, because from the context it was clear that the student wanted to convey the importance of a project manager knowing how to communicate.)

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 first or second 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									

	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
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									
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)									

	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
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. Using appropriate format and layout when typing assignments									

45. Which skills or strategies taught in the academic literacy intervention did you find most helpful in your other subjects?
46. Which skills or strategies do you think need to be focused on more in this academic literacy intervention?
47. Which skills 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). Questions 1, 24, 44, 45, 46 and 47 were added to the original questionnaire.

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 first and second 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				

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
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				
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. 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). Questions 1, 24, 44, 45, 46 and 47 were added to the original questionnaire.