Chapter 8: Implementation and evaluation of the cross-disciplinary intervention

8.1 Introduction

Whereas chapters 6 and 7 reported on planning, designing and evaluating a subject-specific essay-writing intervention, this chapter follows a similar design and procedure for an intervention with a broader disciplinary focus, and compares the results, using both qualitative and quantitative methods of data analysis and evaluation. First, an overview is given of the current debate regarding wider-angled approaches, followed by a description of the research methodology that was followed to measure the effectiveness of the cross-disciplinary intervention, and a presentation and discussion of the results.

8.2 Rationale and approach

As stated previously, the common-core versus subject-specific debate in language pedagogy has been going on for more than 20 years. In the context of teaching undergraduates to write academic essays, the most compelling argument from the side of common-core approaches might be that it is imperative for university students to move comfortably between the discourses of a number of academic disciplines. They need to "control a range of genres appropriate to each setting, and to handle the meanings and identities that each evokes" (Hyland 2009: 129). Thus, according to Bruce (2008: 34) there has been a movement away from discipline-based ESP course designs and methodology to a more "discourse and genre-based cross disciplinary approach". This trend, combined with the universal reality of undergraduate students being underprepared to engage in academic discourses (Johns 1995; 2002; Lillis 2001; Hyland 2004), and the researcher's desire to compare the effectiveness of narrow- and wide-angled approaches within a genre framework justified the design and evaluation of a cross-disciplinary academic writing intervention.

Following suggestions made by an external review panel that was appointed to evaluate the work of the Unit for Academic literacy, and reinforced by the research in progress,

the institution of a second-year module in academic writing was approved by Faculty in 2008, with commencement in 2009. The institutionalization of the module, in turn, increased the relevance of the research.

Limitations were that the module was not officially prescribed or recommended by any of the existing academic programmes in or outside the Faculty. A further constraint was that the alpha code UAL was assigned to the module, whereas all other modules offered by the Unit for Academic Literacy bore the code EOT, which may have had an impact on the visibility of the module (in an alphabetically organized prospectus). Finally, there was the added financial burden of R2000 (the cost of the module) added to students' annual programme fees. Although more than 30 students indicated interest, this price was too high for a "nice to have" that was unlikely to be covered by a bursary or student loan. Eventually only 14 students registered, of which 11 completed the module.

Another limitation was the researcher's lack of foreknowledge regarding the disciplines that would be represented by the students, and thus the specificity of the syllabus and the materials that could be designed in advance.

8.3 Design and implementation of the intervention

8.3.1 Respondents

Despite the constraints outlined above, 14 students registered for the module, of which 11 followed through. The attrition rate of 21% can be accounted for as follows: One of the students was an international exchange student from Germany who only attended seven weeks of the 14 week course; another indicated that she was interested only in political analysis, and that the content of the course was not entirely suited to her needs; and the third discontinued the module in the third week because of work load. The remaining 11 were registered for the following academic programmes: BA (2); BPolSci (5); BA Languages (3); BCom (1). The subjects which they were registered for included Accounting, Criminology, Economics, English, Journalism, History, History of Art, Political Sciences, Philosophy, Psychology, Sociology, and Visual Studies. Their sociodemographic profile could be summarized as follows: 2 white males with Afrikaans as their mother tongue; 1 white male with English as his mother tongue; 2 white females with Afrikaans as their mother tongue; 1 black female with Portuguese as

her mother tongue, 4 black females with an African language as their mother tongue, and 1 black male with an African language as his mother tongue.

8.3.2 Syllabus and materials

To facilitate comparison of the two interventions, and limit variability, the same broad syllabus structure was used as for the subject-specific module, focusing on the use of rhetorical modes, types of claims and types of support in developing an academic argument, while following the Teaching and Learning Cycle of the Australian genre school. Argumentation was given a more prominent role, and using rhetorical modes became one of the secondary threads. This was necessary because of less pronounced relationship between essay structure and primary rhetorical mode in subject-fields other than history. In addition, more emphasis was placed on stance and engagement, because of the lack of the history students' improvement in this area. Table 8.1 represents the presyllabus for the cross-disciplinary intervention.

Table 8.1 Presyllabus for the cross-disciplinary intervention

Study unit 1: Academic discourse(s) *The generic features of academic discourse are discussed with reference to authentic texts, followed by the study of texts from specific disciplines in the Humanities in order to emphasize the need for mastering the characteristic features of disciplinary discourses.* Themes • What is academic discourse? • Is there only one academic discourse?
Study unit 2: Modes of writing (text types) *The mastery of rhetorical modes is practised during a cycle comprising the exploration of excerpts from authentic academic texts, freewriting, explicit teaching of the lexicogrammatical features that characterize each mode, identification of frequently used modes, and independent writing of paragraphs or short essays.* Themes • Chronological writing: narratives, recounts and processes • Description • Comparison and contrast • Cause and effect • Exposition • Analysis

Study unit 3: Academic arguments: formulating claims

Examples of essays are analyzed to identify the types of claims contained in thesis statements. Students also formulate their own claims on the basis of given topics.

Themes

- Fact and opinion
- What is a claim/thesis?
- Positioning of the main thesis of an essay
- Types of claims (factual; causal; evaluative; recommendations)

Study unit 4: Academic arguments: invoking evidence

Examples of essays are analyzed to identify types of support and types of evidence, focusing on lexicogrammatical markers.

Themes

- Types of support (comparison; definition; well-chosen examples; statistics; appeals to audience needs; appeals to authority; addressing a counterargument)
- Using appropriate types of support for different types of claims

Each study unit comprised a set of outcomes and a learning component containing theory, model texts and a variety of authentic task types, some of which were done collaboratively in class, and some as homework tasks that had to be submitted for marks.

Course materials consisted of a 100 page study guide *cum* workbook, based on the presyllabus, a reader (hard copy) comprising 4 broadly focused articles on the theme selected for the content of the module, *viz. Poverty in Africa*, and a partially interactive Blackboard-based website. (This theme was suggested by lecturers from the departments who contributed materials for the writing task survey, because of its relevance across disciplines.) The website contained administrative information about the lecturer, the content and assessment as well as a calendar with important dates. Via a link to the library students had access to a variety of scholarly articles (for which copyright clearance was obtained) and web resources. The Discussion Tool allowed students to interact with one another on matters of common interest, and the Announcement Tool was used to alert students to important dates and events on the calendar. Additional class notes and the list of topics for the final examination were uploaded to folders on the home page for the module.

Although the workbook contained a selection of texts from a variety of disciplines in the humanities, it was realized that these might not be relevant to the core foci of the students who would register for the course, and that one would need to substitute some of the examples and exercises with more relevant materials through the course of the module. This procedure was regarded to be completely in line with the postmethod strategy of a basic presyllabus, which is adapted on the basis of feedback and learner needs.

Assessment was done in accordance with faculty regulations. The semester mark (progress mark) was based on continuous assessment of written homework tasks submitted and marked throughout the semester, and the exam mark, each contributing a weight of 50%. The 14 week intervention (two contact sessions per week) commenced in February 2009.

8.4 Quantitative evaluation

8.4.1 Method

All students enrolled for the module had to write a pretest and a posttest. The pretest assumed the format of a 50 minute in-class essay during the second week of the module. All participants received the reader (containing four articles on general aspects of poverty in Africa) a week in advance, and were requested to prepare for the pretest essay. They were allowed to use the reader as an in-class resource. The pretest did not count towards the students' final marks, as they had not received any tuition on essay-writing at that point. Before writing the pretest consent was obtained to use unattributed extracts from participants' essays as well as the analytic scores awarded by the raters.

The posttest comprised the summative evaluation of the module. The students were allowed to choose from a list of topics on various issues relating to poverty in Africa, which had been requested from the relevant academic departments. They were given one month in which to prepare for the essay exam. The planning, literature search, literature review, outlining, writing and reviewing had to be done without assistance from the lecturer in order to determine whether the scaffolding introduced throughout the module had equipped them to independently apply the linguistic and structural

principles they had acquired. Table 8.2 shows the topics that were chosen by the students:

Table 8.2 Essay topics chosen by students in the cross-disciplinary group

Topic (and description)	Subject-field	No. of students
To what extent was poverty an inevitable by-product of European colonialism in Africa?	History	1
Whose obligation is it to do something about poverty in society: the rich or the poor?	Philosophy	2
Whose obligation is it to do something about the moral problem of poverty: the poor or the government?	Philosophy	1
Analyze the poem "London" by William Blake (in the *Norton Anthology of Poetry*) OR "An abandoned bundle" by Oswald Mtshali (in the *Paperbook of South African Poetry* ed. Chapman) paying close attention to the way the poem depicts both physical and spiritual poverty.	English literature	1
Discuss how Boesman and Lena are dehumanized by poverty and racial discrimination in Athol Fugard's *Boesman and Lena*. Refer closely to the text throughout your discussion.	English literature	3
The policy gap and poverty.	Political Sciences	1
Evaluate the United Nations' Millennium Development Goals (MDGs) as a global strategy to arrest poverty, by referring to the MDGs' normative as well as practical contribution to the plight of the poor.	Economics	1
Famine and hunger are often associated with poverty. How can this be combated through policy initiatives?	Sociology	1

The exam was taken in the Computer Based Testing Laboratory of the Informatorium on campus. Students were allowed to bring into the venue a sheet of paper with five citations, not exceeding 100 words, which they could integrate in their essays. This concession was made to facilitate the assessment of their ability to engage with other authors. Students had to use the 2003 version of Microsoft Word, since not all of them were familiar with the 2007 version. The spell- and grammar-checker was disabled.

The pre- and the posttest essays were scored independently by the course designer, who also presented the generic course (Rater 1), and a part time lecturer with more than 20 years experience in teaching English literature, language and academic literacy, as well as a doctorate in Applied Linguistics from a reputable South African university (Rater 2). The assessment instrument was the same as for the subject-specific intervention (compare Table 7.1 in Chapter 7).

Rater 1 scored students higher than Rater 2 on both the pretests and the posttests: on average the pretests were scored 3.4% higher by Rater 1 than by Rater 2, and the posttests were scored 2.4% higher by Rater 1 than by Rater 2. The correlation coefficients of the scores of the two raters are 0.96 for the pretest and 0.97 for the posttest, therefore warranting the use of the average of the two raters' scores as a measure of each student's performance.

After the rating process had been completed, the two raters discussed their experience with scoring the essays. The second rater suggested that the formulation of certain items should be adapted with a view to future rating exercises, first because the relative weight of certain items was regarded to be either too high or too low, and second, to explain and clarify the scope of particular items, especially in cases where the rater would not be familiar with the terminology of certain paradigms in applied linguistics:

- Items 1 and 2 should be combined into one item, Use of source material, because of difficulty to make a clear distinction between Relevance of source data and Integration of source material. Because of the second rater's uncertainty as to the scope of these items, as well as his relative unfamiliarity with the sources that the students had used, he tended to award an average score of 4 on items 1 and 2.
- Item 5 (Development of main argument) should be explained by means of bracketed information such as (coherence and logic).
- Item 9 (Concord and tense) may be weighted too heavily, and its scope could be extended to "Use of verbs".
- For essays in the humanities item 11 (Technical and subtechnical lexis) might not be entirely relevant, first because certain subject-fields do not have as distinctive a nomenclature as others, and second because the use of subtechnical lexis overlaps with item 12 (Style). (Academic vocabulary may be seen as part of academic writing style.)

8.4.2 Presentation and discussion of results

The total score for each of the 11 respondents was converted to a percentage for ease of interpretation (compare Figure 8.1):

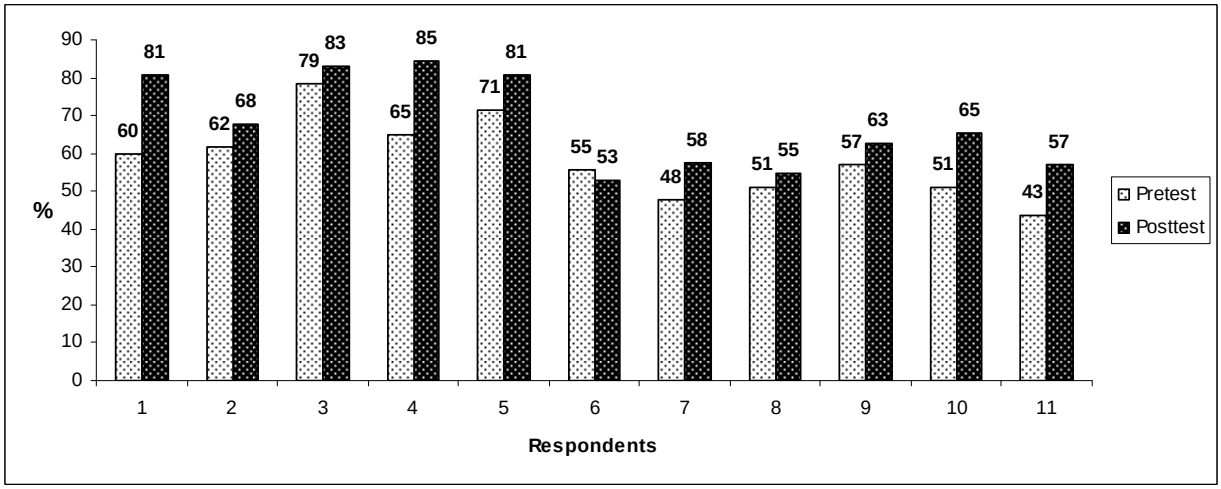


Figure 8.1 Comparison of pre- and posttest results of the cross-disciplinary group per respondent

The average improvement of the 11 respondents was 10%. With the exception of a single student – whose posttest score was only 2 percent less than her pretest score – all the students showed progress, with the largest improvement being 21%.

Figure 8.2 displays the average results per item after conversion to percentages:

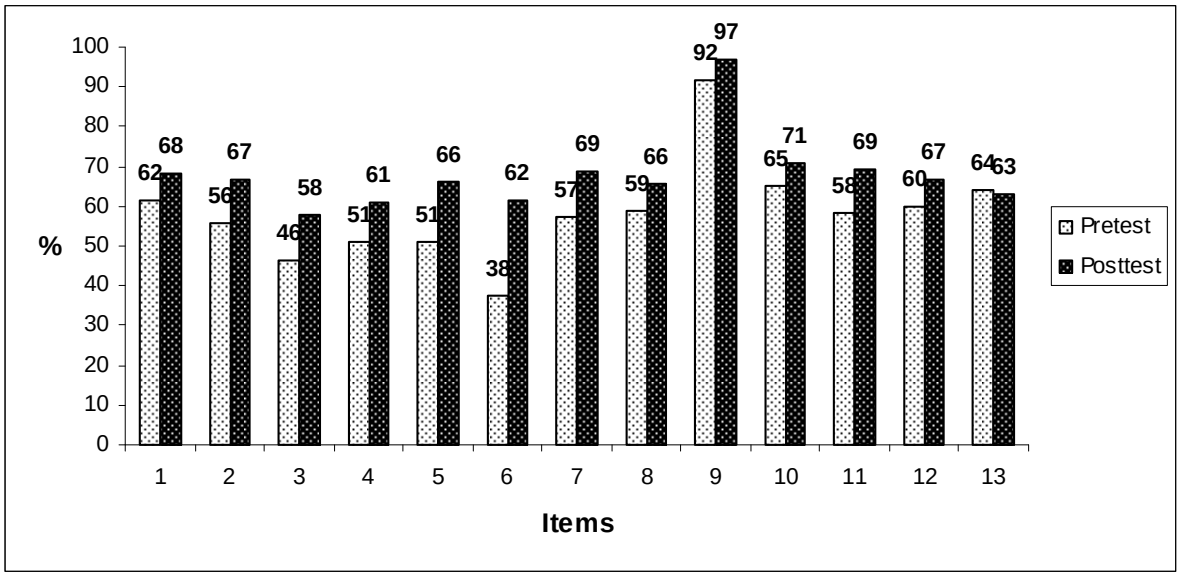


Figure 8.2 Comparison of pre- and posttest results of the cross-disciplinary group per item

Per item, all the posttest ratings were higher than the pretest ratings, except item 13, which was 1% lower. For item 6, the improvement was 24%, while the improvement

was between 10% and 15% on a further six items (items 2, 3, 4, 5, 7 and 11). On the remaining five items the improvement was more than 5%.

On the three primary dimensions of the analytic scoring instrument the improvement varied between 7% and 15%: Table 8.3 shows the mean improvement on the four dimensions of the instrument.

Table 8.3 Percentage improvement of the cross-disciplinary group per dimension

Dimension	Mean: pretest	Mean: posttest	Improvement
1. Use of source materials (Items 1-3)	54%	64%	10%
2. Structure and development (Items 4-7)	49%	64%	15%
3. Academic writing style (Items 8-12)	67%	74%	7%
4. Editing (Item 13)	64%	63%	-1%

8.4.3 Statistical analysis

The Wilcoxon signed-rank test (SPSS version 17; Williams, Sweeney, & Anderson 2009: 764-770) was again used to assess if the differences between the pre- and posttest ratings on each of the 13 questions comprising the instrument were significant. The Wilcoxon signed-rank test is a non-parametric test that is suitable for the analysis of small samples, as in the present case. The test indicates the probability of a significant difference between pre- and posttest ratings, and is appropriate for comparing data from the same participants – in this case the pre- and posttests written by each of the respondents who participated in the intervention.

The results presented in Figure 8.2 should be interpreted against the probability values obtained from the Wilcoxon signed-rank test on each of the 13 items, which are represented in Table 8.5. As in the case of the subject-specific intervention one-sided probability values (p-values) are reported, based on the hypothesis that students' skills would improve as a result of the intervention. P-values lower than 0.05 indicate that there is a significant improvement from the pre- to the posttest ratings at the 5% level of significance. Table 8.4 indicates the p-values for the four dimensions:

Table 8.4 One-sided p-values of the pre- and posttest ratings on the four dimensions of the cross-disciplinary intervention

Dimension	p-value
Dimension 1: Use of source materials	0.022
Dimension 2: Structure and development	0.003
Dimension 3: Academic writing style	0.004
Dimension 4: Editing	0.321
Overall	**0.001**

According to Table 8.4 the improvement between the pre- and posttest ratings is significant at the 5% level for three of the four the main dimensions of the scoring instrument. Only on dimension 4 (Editing) the improvement was not significant (p = 3205). In order to establish whether the p-values of all of the individual items were significant, ratings from the Wilcoxon signed-rank test was obtained for each of the 13 items, as represented by Table 8.5 below.

Table 8.5 One-sided p-values of the pre- and posttest ratings on the 13 items in the cross-disciplinary intervention, obtained from the Wilcoxon signed-rank test

Item	p-value
1	0.086
2	0.018
3	0.016
4	0.076
5	0.016
6	0.003
7	0.013
8	0.023
9	0.011
10	0.080
11	0.013
12	0.065
13	0.321
*16	0.002

*Item 16 was included as it shows the average improvement on all 13 items.

Overall, based on the total difference between the pretest and posttest scores (item 16), the improvement is significant with a p-value much smaller than 0.05. For the remainder of the items, the improvement between the pre- and posttest ratings is significant at the 5% level, with the exception of items 1, 4, 10 and 12, which are significant at the 10% level, and item 13, on which students did not show any significant improvement. This was to be expected, since the average percentage for this item was 63% in the pretest and 64% in the posttest. It should be noted again, however, that in the case of items 1, 4, 10 and 12 a larger sample might have resulted in significant improvement at the 5% level for these two questions as well.

Similar to the subject-specific intervention, an analysis of the pre- and the posttest essays was also conducted for the cross-disciplinary intervention, focusing on the following key resources within a Systemic Functional perspective: **Logical ideation**, representing the *ideational/experiential* function of language, **Appraisal,** representing the *interpersonal* function, and **Theme** representing the *textual* meaning.

8.5 Textual analysis of the essays

8.5.1 Method

Similar to the subject-specific intervention, the pre- and posttest essays of all the participants were tagged electronically for Logical ideation and Appraisal, using the categories and subcategories expounded in Tables 6.5 and 6.6 in chapter 6. Concordances were also built in the same way, using WordsmithTools version 4.0. Regarding thematic development a case study on the pre- and posttest essays of Respondent 1 was performed, comparable with the subject-specific intervention. Analogous to the subject-specific evaluation the focus was set on waves of known and new information at the clause level, using Butt *et al.*'s technique (2000: 143ff).

8.5.2 Presentation and discussion of findings

8.5.2.1 Ideational analysis

Figure 8.3 represents the summarized results of the ideation analysis:

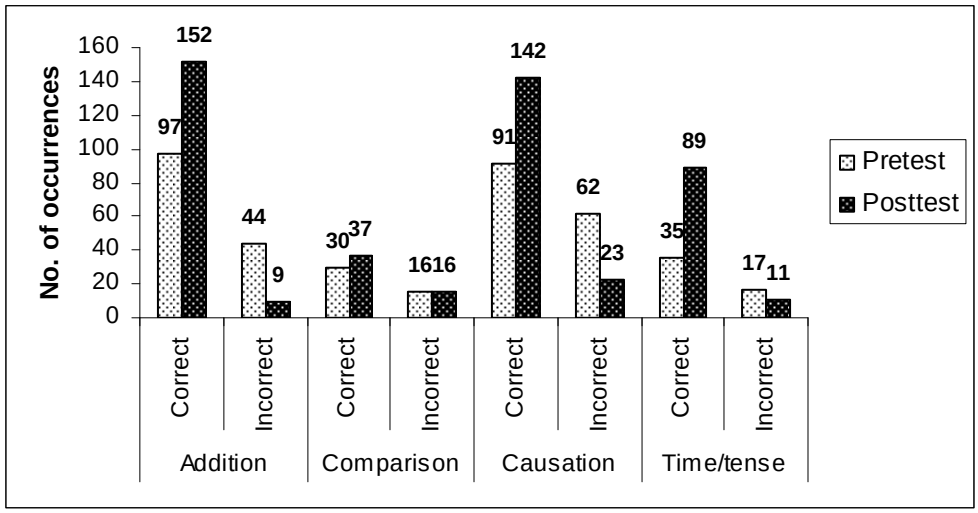


Figure 8.3 Logical ideation: comparison of pre- and posttest results in the cross-disciplinary intervention

According to the graph the cross-disciplinary intervention students used more markers of logical relationships in the posttest than they used in the pretest – in all four main categories. There was an increase of more than 50% in each category, except Comparison, where the improvement was 23%.

Analogous to the subject-specific intervention, significantly more tokens of Causation and Addition were used in the posttest than in the pretest. There was also a significant decrease in the number of errors in these categories between the pretest and the posttest (80% decrease in the number of Causation errors and 62% in the number of Addition errors). Upon further scrutiny it transpired that the posttest yielded more variety in the use of Causation resources (thus more variety in the representation of subcategories): Where the pretest yielded 7 correct usages of Condition, 4 of Means and 8 of Purpose (with 31 instances of Cause and 41 instances of Consequence), the posttest yielded 19 of Condition, 21 of Means, and 24 of Purpose (with 44 instances of Cause and 34 of Consequence).

Also noteworthy is the decrease of 35% errors in the Time/tense category. This decrease cannot be ascribed to an improved mastery of tense, but to an improvement in the use of other markers of temporal relationships. A possible explanation for the relatively few Tense errors in both the pretest and the posttest may be that the overarching topic for both the pretest and the posttest, *Poverty in Africa*, demanded less skill in moving back

and forth between present and past, than was demanded by the topics of the subject-specific intervention.

8.5.2.2 Interpersonal analysis

Appraisal

The Appraisal analysis of the cross-disciplinary intervention used the same three categories of analysis as the subject-specific evaluation, *viz.* Attitude (with subcategories Emotion, Judgement and Social Valuation), Engagement (divided into Attribute and Proclaim) and Graduation (split into Force and Focus). Figure 8.4 represents the summarized results:

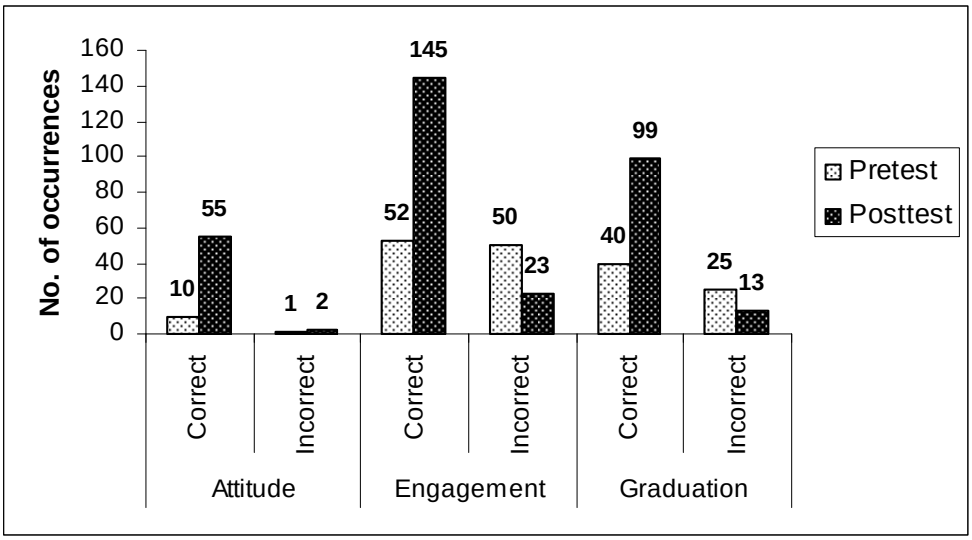


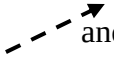
Figure 8.4 Appraisal: comparison of pre- and posttest results in the cross-disciplinary intervention

Figure 8.4 shows that on all three dimensions separately, there were significant improvements: On the Attitude dimension there was an increase from 10 to 55 (= 450%); the number of Engagement markers increased from 52 to 145 (= 173%), and the number of Graduation markers increased from 40 to 99 (= 148%). The steep increase in the number of Engagement markers in the posttest could possibly be ascribed to the lecturer *cum* researcher's emphasis on the importance of entering into debate with other authors. Even with 14% error on the posttest, it still proves worthwhile to teach students strategies of Engagement – even at undergraduate level. It is particularly encouraging that only 38 incorrect or inappropriate usages of any of the Appraisal resources occurred in the posttest.

8.5.2.3 Textual analysis

Theme and New

Theme analysis was conducted similar to the subject-specific intervention. In an analogous way Respondent 1's essays were selected for the case study, using the following symbols to categorize thematic bonds:

- bold vertical arrows ↑ and vertical bracketed arrows  to indicate strong thematic bonds with previous Themes;
- Oblique arrows  and oblique bracketed arrows  to indicate thematic bonds with previous News;
- non-bold vertical arrows ↑ and broken bracketed arrows  to indicate weak thematic bonds with previous Themes;
- oblique broken arrows  and oblique broken bracketed arrows  to indicate weak thematic bonds with previous News;
- the symbol **Ø** to indicate the absence of a thematic bond.

The quantified results of the Theme analysis of the pre- and posttest essays of Respondent 1 are given in Tables 8.6 and 8.7 below (compare Appendix G on CD for the full essays):

Table 8.6 Cross-disciplinary intervention: Pretest 1 (60%)

No. of words: 434
No. of paragraphs: 9
No. of clauses: 43

Strong bonds to directly preceding Theme	Strong bonds to earlier Theme (s)	Strong bonds to directly preceding New	Strong bonds to earlier New(s)	Weak bonds to directly preceding Theme	Weak bonds to earlier Theme	Weak bonds to directly preceding New	Weak bonds to earlier New(s)	No bonds
5	5	10	3	2	3	1	0	12
12%	12%	23%	7%	5%	7%	2%	0%	26%

Strong bonds: 23 (53%)
Weak and absent bonds: 18 (42%)

Table 8.7 Cross-disciplinary intervention; Posttest 1 (81%)

No. of words: 649
No. of paragraphs: 7
No. of clauses: 61

Strong bonds to directly preceding Theme	Strong bonds to earlier Theme (s)	Strong bonds to directly preceding New	Strong bonds to earlier New(s)	Weak bonds to directly preceding Theme	Weak bonds to earlier Theme	Weak bonds to directly preceding New	Weak bonds to earlier New(s)	Absent bonds
30	4	11	12	1	0	0	0	4
49%	07%	18%	20%	02%	0%	0%	0%	02%

Strong bonds: 57 (93%)
Weak and absent bonds: 5 (8%)

Although the Theme analysis indicated that the pretest contained more strong than weak bonds, there was still a remarkable improvement if compared to the results of the analysis of the posttest: Strong bonds increased from 53% to 93% and weak and absent bonds decreased from 42% to 8%.

In order to further decrease the probability that Respondent 1's improvement between the pretest and the posttest was due to chance, another respondent was randomly selected, *viz.* Respondent 7 (compare Appendix G). In her case the improvement in handling thematic development was less dramatic than in the case of Respondent 1: The pretest contained 18 (33%) strong bonds and 35 (65%) weak and absent bonds, whereas the posttest contained 25 (41%) strong bonds and 34 (56%) weak and absent bonds. In view of the fact that she improved by only 10% between the pretest and the posttest according to the analytic scoring (48% versus 58%), a more modest improvement in terms of her ability to handle thematic progression could be expected. It is important, nonetheless, to observe that the improvement is again present, and noteworthy.

8.6 Opinion survey

Similar to the subject-specific intervention an opinion survey was conducted, using the same questionnaire and the same procedures for recording the data and analyzing the results.

Figure 8.5 summarizes the average rating per student on each of the five dimensions of the construct.

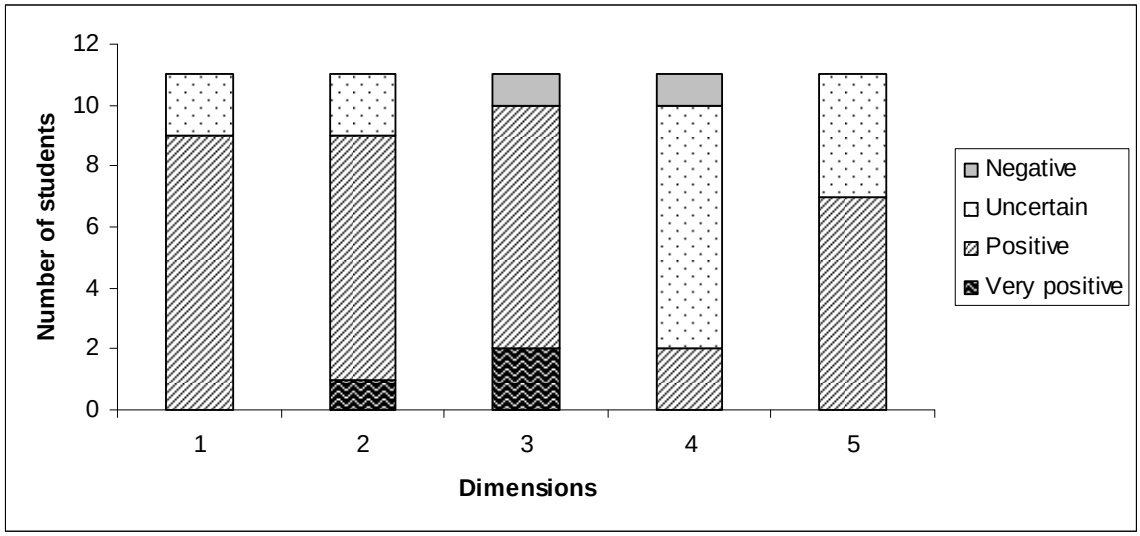


Figure 8.5 Students' opinions on the 5 dimensions of the construct: cross-disciplinary intervention

The students' general lack of commitment, and the fact that the lecturer had to reprimand certain individuals for their relative indifference, predicted a less favourable evaluation of the course. However, although few students chose the extremely positive option on the scales, their opinions regarding the course were still generally positive.

The dimension that evoked the most favourable responses was the Needs-driven syllabus (Dimension 3). From the raw data it could be determined that only 1 student "disagreed" that his/her expectations had been fulfilled (Statement 15) and that the most important questions he/she had had about essay-writing (Statement 17) had been fulfilled. He/she was "uncertain" as to whether the lecturer had been interested in addressing his/her personal needs (Statement 17).

Dimension 4 evoked the least favourable responses. For two reasons this finding was not surprising. As mentioned in the discussion of the subject-specific opinion survey in Chapter 7, students are generally skeptical about the potential of university courses to teach them critical thinking skills, and this perception was sustained by the outcome of that survey on Dimension 4. It would therefore be surprising if the students who

participated in the cross-disciplinary intervention would give a positive rating. Thus, the fact that only two students responded positively to this dimension, eight were uncertain and one was negative, did not come as a surprise. In order to gain insight into the responses to the individual statements comprising this dimension, a graph was generated (Figure 8.6). (Note again that the scales for Statements 23-25 were reversed for the statistical analysis to bring their polarity in line with that of Statements 21 and 22.)

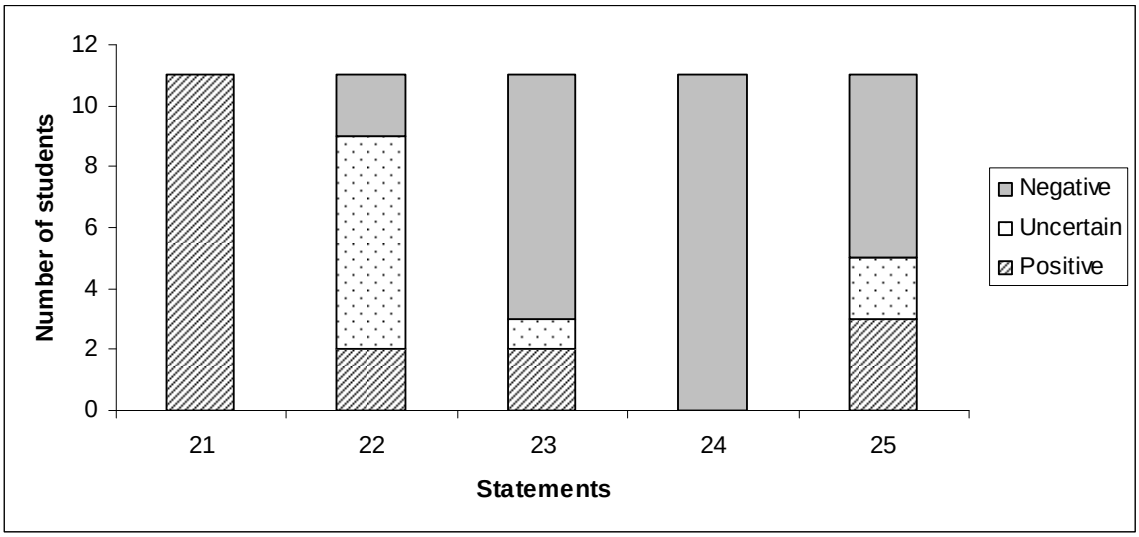


Figure 8.6 Responses to the concepts evaluated by statements 21-25 (Dimension 4)

The results indicated on the graph can be interpreted and explained as follows:

Statement 21: *It is empowering to know how to write in the genres valued by academics.*

Six students responded extremely positive (1), while the other five responded moderately positive (2). Since the statement was phrased positively the responses indicate a generally positive evaluation of empowerment through genre knowledge.

Statement 22: *If one of my academic lecturers says that it is forbidden to refer to myself ("I") in academic writing, I will take issue with him/her.*

Seven students were uncertain whether they would challenge a lecturer, two were moderately sure that they would, and two were very certain that they would not. The mixed responses were not surprising, since it became very clear during the intervention that subject-fields hosted divergent views in this regard. Furthermore, some tolerated individual differences, while others insisted on following the conventions laid down by

the particular discourse community. For instance, certain subject-fields almost prohibit self reference (History), others tolerate it if used in moderation (for example, Economics), while others encourage it (Philosophy). Thus, in a sense this question did not necessarily measure opinions, but rather knowledge of subject-field conventions.

Statement 23: *One should accept the content of textbooks and academic articles as true.*

Since this statement was phrased negatively a "positive" answer would have to be either 4 or 5 on the scale. Since eight of the students ticked either 1 or 2, almost two thirds seem to believe that the authority of prescribed sources should not be questioned or challenged.

Statement 24: *It is impossible to criticize one's own work.*

All 11 students ticked either 1 or 2, which means (seeing that the question has a negative polarity) that they are all convinced of the value of self-reflection. Of course, this does not mean that they necessarily always take the time to actively pursue this goal.

Statement 25: *Empowerment in tertiary education means that students should be allowed to write as they speak.*

The responses to this question were rather mixed, which might have resulted from the negative polarity. Strictly according to the answers only 45% of the respondents seem to harbour a misconception regarding an important objective of the intervention (6 students ticked 1 or 2), two were uncertain, one indicated moderate disagreement, and one indicated complete agreement. The swing of the pendulum in a negative direction does therefore not seem to be meaningful.

These explanations indicate that the slightly negative response to Dimension 4 was not a reason for too much concern. Only statement 23 seem to merit further investigation, but perhaps the first semester of the second year at university is still too early to have developed a critical orientation towards work produced by experts, and hopefully students will develop a more interrogating stance as they move closer to graduation.

Because of the importance to prove or disprove the measure in which genre-based interventions facilitated transfer, Dimension 5 was also teased out some more. It

transpired that all students responded positively or moderately positively to Statement 26 (*The course has contributed towards improving my writing ability in English*) and Statement 28 (*Much of what I have learned about essays I can also use when writing reports and other text types*). Furthermore, the majority (8 and 9 respectively) gave positively evaluating answers to Statement 27 (*I find it difficult to apply the principles we have learned in this course to writing tasks in other subjects*) and 29 (*Since I started this module my marks for written work in other subjects have improved*). Those who did not respond positively indicated uncertainty. These results are encouraging, particularly in the light of the less than desirable level of commitment and perseverance demonstrated by the students who registered for the cross-disciplinary module.

8.7 Author's reflection on the cross-disciplinary intervention

The varied performance of students in the cross-disciplinary intervention seems to derive, at least partially, from their reasons to register for the course. While some of the students who registered had a genuine desire to learn how to write academically in their chosen disciplines, others enrolled simply to accumulate credits towards their degrees. This reflection attempts to highlight the relationship between students' achievement in the essay-writing module and their perceptions about the instrumental value of their learning. It also emphasizes the dialectic relationship between learning to write and writing to learn.

Hyland (2009:124) claims that benefits of courses teaching students how to write in an academic way "are only perceived as such if students value what this literacy allows them to do". The present experience with administering the semi-generic intervention resonates with Hyland's claim, and echoes the findings of Lillis (2001:85), Lin (2000), Canagarajah (1999) and Ivanic (1998), *viz.* that certain students passively resist the assumptions and values which they are assumed to acquire. The fact that some students do obtain average to above average marks in content disciplines may reinforce this passive resistance. The following example is a case in point: After performing well below average on the third assignment in the cross-disciplinary module, one student remarked:

> Ma'am, I don't understand; in my other subjects I get good marks for my assignments, but in your class I fail. It's not like I don't like you or so; I enjoy your classes. You can see I always attend.

In contrast, students with a genuine desire to acquire the essayist literacy of the academy, and who experienced the benefits of applying this literacy in their content subjects, flourished. Respondent 5, for example, obtained high grades in secondary school, and matriculated with distinctions in Afrikaans and English. However, in philosophy, which is one of his major subjects at university, he obtained a just above average mark, and registered for the essay-writing module to improve his marks. Fortunately, for him, the essay topics contributed by the Philosophy Department were related to the philosophy curriculum. It is thus not surprising that this student's marks in the essay-writing module improved by 20% between the pretest and the posttest. His philosophy lecturer testified that his improvement in the content subject was about equal.

Thus, motivation alone is not sufficient for success. Subject-field knowledge is another prerequisite. The following anecdote is offered to further support this claim: Respondent 5, a student who is registered for a degree in Journalism, and who chose the same topic as the philosophy student (Respondent 4), *viz. Whose obligation is it to do something about poverty in society: the rich or the poor?* did not demonstrate the same improvement as the philosophy student between the pretest and the posttest. In fact, she scored four percentage points lower than he did in the posttest. Although her work was grammatically correct, her style, development of the main argument and selection of evidence remained "generic". The most plausible explanation is that although Respondent 5's proficiency in academic English was superior to that of Respondent 4 (based on her score in the pretest and her participation in class) she departed in her posttest essay from a zero knowledge base, since she had not been initiated into the "discourse of philosophy" through regular interaction with reading matter, lecturers and peers in the domain of philosophy.

This anecdote highlights a serious design error by the researcher. Students in the cross-disciplinary group were allowed free choices from the topics provided by the content departments. It was anticipated that they would choose topics relevant to their core university disciplines. However, some made their choices on the basis of familiarity, such as topics related to the literary works they had previously studied in English

Literature. To be specific, three students chose the drama *Boesman and Lena* by Athol Fugard, and one chose the poem *An Abandoned Bundle* by Oswald Mtshali, although only one student in the group was registered for a language programme. As confirmed by discussions at the conclusion of the module, some students' choices were motivated by convenience, rather than the ideals of the Vygotskyan Zone of Proximal Development. For two of the students (Respondents 8 and 9), unfortunately, convenience became a trap, because they deviated from the main focus, which in both cases was *the physical and spiritual effects of poverty in apartheid South Africa*, as portrayed by the particular author. When discussing their first drafts, it was difficult to convince these students, who had structured their essays according to literary elements instead of characteristics of physical and spiritual poverty, that their essays were "off topic".

Based on the researcher's teaching experience, combined with self-reflection, the following improvements for future interventions are suggested:

- Introduce extensive writing earlier in the semester.
- Build a corpus of authentic materials.
- Facilitate a close fit between students' core disciplines and their focus in an essay-writing intervention.

Introduction of **extensive writing** at an early stage is motivated by the empirical observation that students only became convinced of the implications of their lexicogrammatical, stylistic and structural choices after having written a full essay, which was after the 10th week of the 14 week module. A second reason for suggesting that full essays be written much sooner is the clear lack of engagement observed in the students when writing shorter assignments, such as paragraphs or parts of full essays. The haphazard and untidy way in which some of the shorter homework tasks had been executed supports the hypothesis that authenticity feeds into motivation, and motivation plays a major role in the quality of the output.

The solution supported by the majority of students during the post-intervention feedback session was to start writing complete essays very soon after the commencement of the course. This does not mean changing the content of the syllabus,

but in the context of a genre approach it suggests that joint exploration and joint construction in respect of each of the discourse skills emphasized in the course (rhetorical modes, making and supporting claims, thematic development, cohesion, and stance and engagement) should immediately be followed by independent writing of a full essay. In the assessment of the essays the primary focus should be on the particular skill or ability that the students had practised during the preceding week or fortnight.

Concerning materials design, **authenticity** has been confirmed as a core principle. The researcher relied heavily on exemplars from writing manuals published in the US and the UK (for example Barnet 2008; Oregon State University 1997; Richlin-Klonsky & Strenski 1994; Rosnow & Rosnow 1998; Schmidt 2005), but although some of the essays were good overall examples, not all of them were exemplary in terms of every aspect of the syllabus. In order to address these deficits future interventions could draw on essays written by local students who have successfully completed the intervention. Good examples could serve as model texts, whereas poorer attempts could be used to practise editing.

Finally, the designer of an intervention should ensure a close link between the texts and topics that are selected for writing purposes and the content that students have to learn in their core disciplines. It has been proven that writing helps them to master content, while at the same time content knowledge helps them to develop fluency and accuracy.

8.8 Conclusion

The results obtained from the quantitative evaluation indicate that students definitely benefited from the cross-disciplinary intervention. However, unlike the subject-specific intervention the improvement was not equal on the three primary dimensions measured by the analytic pre- and posttest assessment: On the dimension Handling of source materials the average improvement was 10%, on Structure and development they improved by 15% and on Academic writing style there was only 7% improvement. The most plausible explanation for the fact that the improvement in the use of source materials was moderate, or less than expected, may have been the fact that students did not need to study the content of the sources in depth for assessment in their content disciplines, and thus they might have been less motivated to engage with sources on the

broad topic selected by the course designer. Also, they did not necessarily consult the same sources for the pretest and the posttest, and thus did not become familiar with core resources. A possible reason for the slight improvement on academic writing style is that the respondents' grammar and vocabulary were on already at a fairly high level when they entered the course (67% on average). Furthermore, the intervention did not pay any specific attention to the improvement of grammar, and neither was style explicitly taught, except for brief pointers on issues of formality. The fact that the most significant improvement occurred on the dimension of structure and development was not a complete surprise, in that discourse structure, comprising thematic development at the level of the whole text (thesis and conclusion) the paragraph (topic sentence and paragraph development) and clause level (manipulation of Theme and New), is one of the dimensions that can be taught via templates and explicit instruction. Moreover, both stronger and weaker students are able to grasp the main principles and apply them.

The snapshots taken of students' performance on aspects of the three primary areas of meaning-making according to Systemic Functional Grammar indicates that explicit teaching of the grammatical resources for encoding these meanings does pay off. The fact that the students used significantly more Appraisal markers than the subject-specific intervention students is particularly meaningful, since the lecturer made a concerted effort at teaching Appraisal resources to the cross-disciplinary group.

According to the opinion survey, students were positive to moderately positive about the intervention. Although not all the respondents thought that their personal needs had been addressed, they were generally of the opinion that they had learned valuable skills, which they could apply in other contexts, and which had already stood them in good stead. However, the overall impression gained from the outcomes of the opinion survey and the personal experience of the course researcher was that the success of future interventions of this nature would depend, to a large extent, on the authenticity of the materials used and the ability of the classroom lecturer to engage students and ensure active participation.

Chapter 9: Comparison of the subject-specific and the cross-disciplinary interventions

9.1 Introduction

Chapters 6 to 8 reported on the design, development and evaluation of two genre-based writing interventions – one aimed at second-year students of history, and the other aimed at second-year students registered for a variety of subjects in the humanities. This chapter compares and evaluates the findings from the two interventions.

First, the statistical results of the two interventions obtained from the analytic scoring of the pre- and posttest essays are juxtaposed, followed by a statistical comparison of the improvement resulting from the two interventions. Subsequently, the findings of the SFL-based textual analyses are compared. Lastly, the results from the questionnaire surveys are statistically compared to give an impression of students' appraisal of the effectiveness of the respective interventions.

9.2 Comparison of the essay ratings

The evaluation of the subject-specific as well as the cross-disciplinary intervention pivoted on a comparison of the pre- and posttest essay scores, where a standardized analytic scoring instrument was used. The primary aim was to test the hypothesis that students' essay-writing abilities would improve significantly as a result of a genre-based writing intervention, irrespective of the disciplinary scope. The second aim was to establish the difference (if a difference should be found) between the effectiveness of narrow-angled and wide-angled genre-based interventions. For each intervention descriptive statistics were used to indicate the improvement per candidate, per item, and per cluster (dimension) of items. Thereafter statistical tests were conducted to calculate the probability that the improvement was statistically significant.

Table 9.1 below compares the improvement, per intervention group, on each of the four primary dimensions of the analytic scoring instrument (Use of source materials, Structure and development, Language and style, and Editing) as well as overall:

Table 9.1 Comparison of the two intervention groups in terms of their improvement on the four dimensions of the scoring instrument

Dimension	Mean: pretest		Mean: posttest		Improvement	
	S-specific	*Generic*	*S-specific*	*Generic*	*S-specific*	*Generic*
1. Use of source materials	51%	54%	69%	64%	18%	10%
2. Structure and development	56%	49%	74%	64%	18%	15%
3. Academic writing style	62%	67%	81%	74%	19%	7%
4. Editing	64%	64%	81%	63%	17%	-1%
					18%	**8%**

According to Table 9.1 there is a 10% "overall" difference between the groups in terms of their improvement as a result of the particular intervention. The table shows that the overall improvement of the subject-specific group was about equal on the three primary dimensions measured by the analytic pre- and posttest assessment (between 17% and 19%), while the overall improvement of the cross-disciplinary group was more moderate (8%), and also more variable: 10% on Use of source materials, 15% on Structure and development, 7% on Academic writing style and -1% on Editing.

According to the Wilcoxon signed-rank test (the non-parametrical equivalent to the paired T-test) both groups, individually, showed a significant overall improvement between the pretest and the posttest. Compare Table 9.2 below:

Table 9.2 The significance of the difference between the improvement of the two groups in the four dimensions of the scoring instrument

	Subject-specific	Generic
Dimension	**One-sided p-value**	**One-sided p-value**
1: Use of source materials	0.004	0.022
2: Structure and development	0.006	0.003
3: Academic writing style	0.003	0.004
4: Editing	0.008	0.321
Overall	**0.002**	**0.001**

Both interventions proved to be successful in their own right. The p-values for three of the four dimensions – Use of source materials, Structure and development and Academic writing style – were well below 0.05 for each group, and thus the improvement was statistically significant for each. Only on Dimension 4, Editing, did the improvement of the cross-disciplinary group not prove to be significant ($p = 0.3205$), which was predictable in the light of the fact that the performance of the group as a whole decreased by 1% between the pretest and the posttest.

In order to establish whether the **difference between the two interventions** (subject-specific and cross-disciplinary) was statistically significant, the Mann-Whitney U-test was applied. The Mann-Whitney U-test is the non-parametric equivalent of the independent samples T-test for assessing whether two independent samples of observations come from the same distribution, which is particularly useful for small samples. In statistical terms it assesses the ranked positions of scores in two different groups. If there are significant differences between the two groups, the p-value associated with the test statistic will be smaller than 0.05. The main finding was that overall, the subject-specific group performed significantly better than the cross-disciplinary group, as predicted by the simple comparison in Table 9.1. A **p-value** of **0.043** was obtained.

Because of the significance of the overall difference found between the subject-specific and the cross-disciplinary interventions, separate Mann-Whitney U-tests were run for each of the four main dimensions of the holistic scoring instrument. Table 9.3 shows the p-values for the four dimensions, as well as the overall value. Two-sided values are reported because one group was not necessarily expected to perform consistently better than the other.

Table 9.3 Two-sided p-values of the scores from the Mann-Whitney U-test

Dimension	p-value per dimension
DIMENSION 1: Handling of source materials	0.223
DIMENSION 2: Structure and development	0.809
DIMENSION 3: Academic writing style	0.020
DIMENSION 4: Editing	0.020
Overall	**0.043**

According to the separate Mann-Whitney U-tests, the subject-specific group did not perform significantly better than the cross-disciplinary group on every dimension. A significant difference was only found with respect to Dimension 3, Academic writing style, and Dimension 4, Editing (p-value, in each case = 0.02). For both these dimensions significant differences were expected on the basis of the simple comparison in Table 9.1. Since the value of the fourth dimension, Editing, was derived from a single item (item 13) a generalization can not be made. It can only be concluded that the subject-specific group succeeded much better than the cross-disciplinary group in improving their spelling and appropriate use of capital letters.

No significant difference was found with regard to Dimension 2, Structure and Development (p-value = 0.809). This was not surprising, because according to the tabulated comparison, the improvement of the two groups differs by a mere 3%: 18% for the subject-specific group and 15% for the cross-disciplinary group. According to my own belief, the basic principles of developing an argument at various levels of the text (the whole essay, paragraph and sentence) are largely subject-neutral, and can be taught and learned successfully through a combination of explicit instruction, model texts and sufficient exercise.

Similarly, no significant difference between the two groups in terms of Dimension 1, Use of source materials (p-value = 0.223) was indicated by the Mann-Whitney U-test. This finding might seem to be contrary to the result of the simple comparison in Table 9.1. A larger sample may result in a significant p-value.

Figure 9.1 shows the distribution of the data on the dimensions of the instrument comprising more than one ratable item (in other words Editing is excluded):

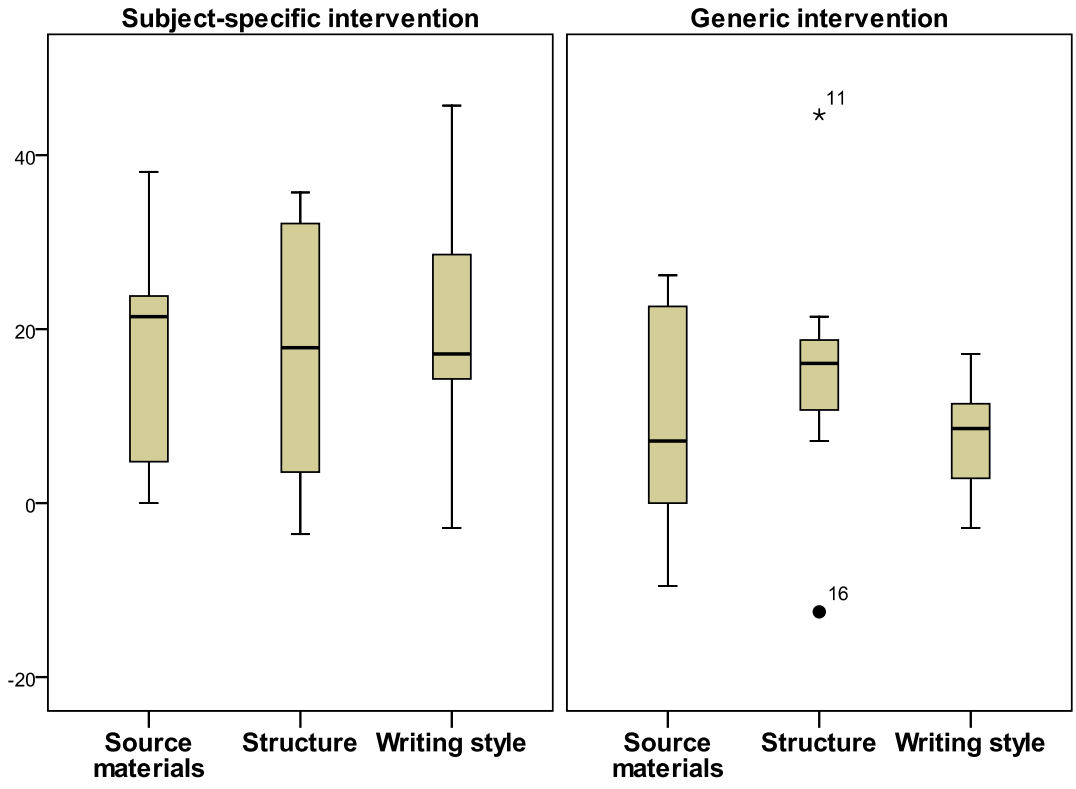


Figure 9.1 Box plots displaying the differences between the subject-specific and the cross-disciplinary intervention groups with regard to the three most important dimensions of the analytic scoring instrument according to the ranks assigned by the Mann-Whitney U-test

The box plots representing the data for Structure and development and for Writing style show that the cross-disciplinary intervention data are less spread out than the subject-specific intervention data. Furthermore, the middle 50% (between the 2nd and 3rd quartiles) overlap for both Use of source materials and Structure and development. As a matter of fact, for Use of source materials the distribution of the generic data constitutes a subset of the distribution of the subject-specific data. In the case of Academic writing style there is much less overlap: The middle 50% of the generic data clearly occupy lower ranks than the subject-specific data.

In order to establish whether individual items may have influenced the p-values on the main dimensions of the scoring instrument, Mann-Whitney U-tests were run for all 13 individual items of the holistic scoring instrument (compare Table 9.4). No Bonferroni corrections were made for the multiple testing (to avoid the inflation of the type I error

rate) as it had already been established that a significant difference existed between the two groups.

Table 9.4 Two-sided p-values of the raters' scores from the Mann-Whitney U-test per item

Dimension	Item	p-value per item	p-value per dimension
DIMENSION 1: Handling of source materials	1 Relevance	0.051	0.223
	2 Integration	0.349	
	3 Stance and engagement	0.654	
DIMENSION 2: Structure and development	4 Thesis statement	0.756	0.809
	5 Development of argument	0.863	
	6 Conclusion	0.557	
	7 Paragraph development	0.223	
DIMENSION 3: Academic writing style	8 Syntax	0.005	0.020
	9 Concord and tense	0.314	
	10 Linking devices	0.099	
	11 Lexis	0.114	
	12 Style	0.387	
DIMENSION 4: Editing	13 Spelling and capitalization	0.020	0.020
OVERALL	**16**		**0.043**

This more detailed analysis identifies specific items that may have exaggerated or diluted the p-values of the dimensions. In the case of dimension 3, Structure and Development, Syntax (item 8, with a p-value to 0.005, and thus significant at the 5% level) and Linking devices (item 10, with a p-value of 0.099, indicating significance at the 10% level) greatly influenced the p-value for the dimension as a whole. On the other hand, although the difference between the intervention groups regarding dimension 1, Use of source materials, was not significant according to the Mann-Whitney test ($p = 0.223$) the p-value for one of the three items comprising the dimension (item 1, Relevance of source materials) indicates a statistical difference between the cross-disciplinary and the subject-specific groups at the 10% level ($p = 0.051$).

It should be noted that the findings regarding the individual items were not surprising, and plausible explanations for significant differences (or a lack thereof) were not hard to find. Items 1 and 8 will be explored in more detail. With regard to Item 1 it can be argued that the history students actively engaged, quantitatively and qualitatively (in both their history classes and the academic literacy classes), with scholarly sources on a

specific theme, *viz. The history of Apartheid in South Africa*, with particular emphasis on *the Native Land Act of 1913*. They also became familiar with the core sources included in their history reader, which was also used for the essay-writing module. In contrast, the mixed group was exposed to fairly generic sources on the topic of focus, *Poverty in Africa*. They might have been less motivated than the subject-specific group to engage regularly with scholarly sources on this topic, since it was not necessary to internalize the content for assessment in their core modules. Furthermore, the students in this group were allowed to write their final exam essay on any of the topics provided by lecturers in the Faculty; and some of them chose topics that seemed to be interesting, but fell outside the focus of the academic programmes for which they were registered. For instance, one of the students, who was registered for a degree in Journalism, chose the topic *Whose obligation is it to do something about poverty in society: the rich or the poor?* This topic requires familiarity with philosophical ways of arguing. The student managed to structure her essay well and to invoke evidence from relevant sources, but she failed to exhibit mastery of the discourse of philosophy. Against this backdrop it is not surprising that the subject-specific group improved significantly more than the cross-disciplinary group on Item 1.

The p-value of Item 8 can be explained as follows: Although none of the interventions paid specific attention to the improvement of syntactic well-formedness, the subject-specific group had the advantage of becoming familiar with the historian's ways of formulation through extensive reading and writing in the discipline. During the course of the semester they wrote at least eight full academic essays on topics related to the history of Apartheid in South Africa. The respondents in the mixed group – with the exception of the two students who studied Philosophy – wrote only three full essays on aspects of poverty during the course of the semester-long essay-writing intervention.

Although plausible explanations can be found for the p-values of the primary dimensions, with specific reference to the impact of individual items, the findings raise questions regarding the validity of the construct underlying the scoring grid. More specifically, they raise questions about the researcher's (and other researchers') clustering of items in analytic scoring instruments for academic writing. Specific questions include:

- Can it be claimed that grammar (which might include syntax and cohesive devices), lexis and style constitute the construct Academic writing style?
- Is the ability to handle stance and engagement in any way connected to the ability to integrate facts and ideas from source materials in a composition, and the relevance of those facts and ideas to the topic at hand?

We now turn to the discourse analyses of the pre- and posttest essays for possible justification of the statistical data, but more specifically to find evidence that might assist course designers to adapt or refocus syllabi and/or teaching materials for future essay-writing interventions.

9.3 Text analysis of pre- and posttests

The discourse analyses that were performed on the essays were focused not so much on an overall impression of students' performance but were actually "enlarged detail" snapshots of students' abilities to handle key aspects of meaning-making in academic texts – as identified and described in the literature on Systemic Functional Grammar. Another aim was to explore the value of theory-supported discourse analysis in justifying rating scores.

9.3.1 Logical ideation

On the dimension of logical ideation (logical relationships between intra- as well as extra-textual concepts) the **subject-specific** group showed a large improvement in handling Causation. There was an overall increase from 74 correct usages in the pretest to 134 correct usages in the posttest (= 81% improvement). A moderate improvement was found in the Addition category (from 62 to 79 = 27%), and a slight improvement in the Time/Tense category. Temporal setting and Temporal sequence were mastered fairly well, already at the time of the pretest, and little improvement was demonstrated in the posttest. However, the number of tense errors decreased dramatically (from 43 in the pretest to 12 in the posttest = 72%). A possible explanation is that the history students had never been explicitly taught how to handle tense in historical writing (personal communication with the lecturer). It is likely that the explicit instruction and continuous feedback during the intervention assisted them in internalizing the system.

The **cross-disciplinary group** showed an increase of more than 50% on three of the Logical Ideation categories: Addition, Causation and Time/tense, and 23% on Comparison. Analogous to the subject-specific intervention, Causation and Addition were handled well, and errors also decreased significantly in this category (an 80% decrease in Causation errors and a 60% decrease in Addition errors). It was also encouraging that more variety occurred in their use of causation resources in the posttest: In addition to the subcategories Cause and Consequence, also Condition, Means and Purpose featured prominently in the posttest. Similar to the subject-specific intervention, Temporal relations and Tense were handled well by the cross-disciplinary intervention students in both the pretest and the posttest, but in contrast to the subject-specific intervention students (whose pretests contained many tense errors) the cross-disciplinary intervention students committed very few tense errors, even in their pretest essays. Only one tense error was recorded in the pretests and two in the posttests. This might be explained by the fact that in humanities disciplines other than history time does not play such a crucial role.

9.3.2 Appraisal

The **subject-specific** students improved inconsistently in their use of Appraisal resources. They showed the most marked increase (47%) in the Attitude category, which includes the subcategories Emotion, Judgment and Social valuation. This may be ascribed to their increased content knowledge, and thus their confidence in evaluating historical figures, institutions and events. The category of Engagement produced disappointing results, in that there was an overall decline from 73 to 38 correct usages. This was mostly due to a decline in the number of Attribution markers (from 44 to 13). The only plausible explanation is that an increase in students' subject-field knowledge – resulting from attending lectures, reading, studying and intensive writing on the history of segregation in South Africa – made them less dependent on sources when writing the posttest essay.

The students participating in the **cross-disciplinary intervention**, on the other hand, improved significantly in their command of Appraisal resources. In the Attitude category correct usages increased (from 10 to 55), in the Graduation category from 40 to 99, and in the Engagement category from 52 to 145. The steep increase in the use of

Engagement markers (Attribution, from 19 to 55 and Proclamation, from 33 to 94) stands in stark contrast to the decrease in the subject-specific intervention. Apparently, the emphasis that the lecturer for the cross-disciplinary intervention had placed on a command of Appraisal resources, and the increased amount of exercise in using these, paid off.

It is likely that the cross-disciplinary group's increased use of stance and engagement markers, as opposed to the slight and inconsistent improvement by the subject-specific students, contributed to the fact that no significant difference was measured on item 3 by the Mann-Whitney U-test.

9.3.3 Thematic analysis

The pre- and posttest essays of the first respondent in the subject-specific and the cross-disciplinary intervention respectively (henceforth Respondent S1 and Respondent G1) were sampled to analyze and plot thematic progression. In both cases improved capability to handle thematic progression was anticipated on the basis of the sizeable difference between the respondents' analytic scores on the pretest and the posttest: Respondent S1's overall score improved from 37% to 69%, and Respondent G1's score improved from 60% to 81%. Further predictors of improvement were the two respondents' scores on the dimension Structure and development, particularly on item 7 (Paragraph development). Respondent S1 scored 2 for this item on the 7-point scale in the pretest, and 5 in the posttest, whereas Respondent G1 scored 3 on the pretest and 6 on the posttest. Although the overall difference between S1's pre- and posttest scores (28%) was more impressive than the difference between G1's scores (21%), G1's scores fell into a higher bracket than those of S1, and thus it could be expected that the percentage of strong thematic bonds in G1's essays would also fall into a higher bracket than the number of strong bonds in S1's essays.

This prediction was borne out by the findings: In S1's essays the number of strong bonds (in relation to the number of weak and absent bonds as percentages of the total number of clauses) increased from 38% to 76%, whereas the number of strong bonds in the G1's essays increased from 53% to 93%. Conversely, the number of weak and

absent bonds in S1's essays decreased from 62% to 23% and in G1's essays from 42% to 8%.

These findings, which signify an impressive improvement in the case of the subject-specific as well the cross-disciplinary intervention student, are in line with the statistical finding of no significant difference between the two interventions on the dimension of Structure and development.

9.3.4 What the discourse analysis reveals

Although no grand generalizations can be made on the basis of these quasi-comparisons, there is a clear indication that both the subject-specific and the cross-disciplinary interventions afforded students tools and mechanisms to improve their academic writing. At least some of these resources must have been internalized to facilitate the improvement that took place between the pretest and the posttest.

In general, the students who took part in the subject-specific intervention became less reliant on sources, which might have impacted negatively on their explicit use of stance and engagement markers, but could have contributed to the enhanced relevance of the source materials (facts) they used in their essays. In contrast, the students in the cross-disciplinary group acquired a more marked command of stance and engagement than their subject-specific counterparts. They also demonstrated a more varied repertoire of cohesive devices. However, it is more likely that the intensified focus of the generic module on these lexicogrammatical devices (as a result of what the course designer had learned from the subject-specific intervention) had caused the improvement, and not the contextual focus (subject-specific or generic).

Based on the thematic analysis of two sample essays (one per intervention) students in both groups benefited from the intervention in terms of developing an argument systematically. This inference is supported by the statistical finding that there is no significant difference between the subject-specific and the cross-disciplinary group in terms of improvement on the dimension of Structure and development.

9.4 Opinion survey

In order to compare the results of the post-intervention opinion surveys, a Mann-Whitney U-test was performed for each of the five theoretical dimensions of the opinion survey, *viz.* (1) Staged and scaffolded teaching and learning model, (2) Purposeful social apprenticeship, (3) Needs-driven syllabus, (4) Critical orientation and (5) Skills transfer. For each of the two interventions the total score for the items comprising each dimension was obtained. The scales were reversed where necessary to facilitate uniform polarity. The spiderweb plot represented as Figure 9.2 shows the differences between the means of the responses of the two groups (after reversal of the scales with a negative polarity):

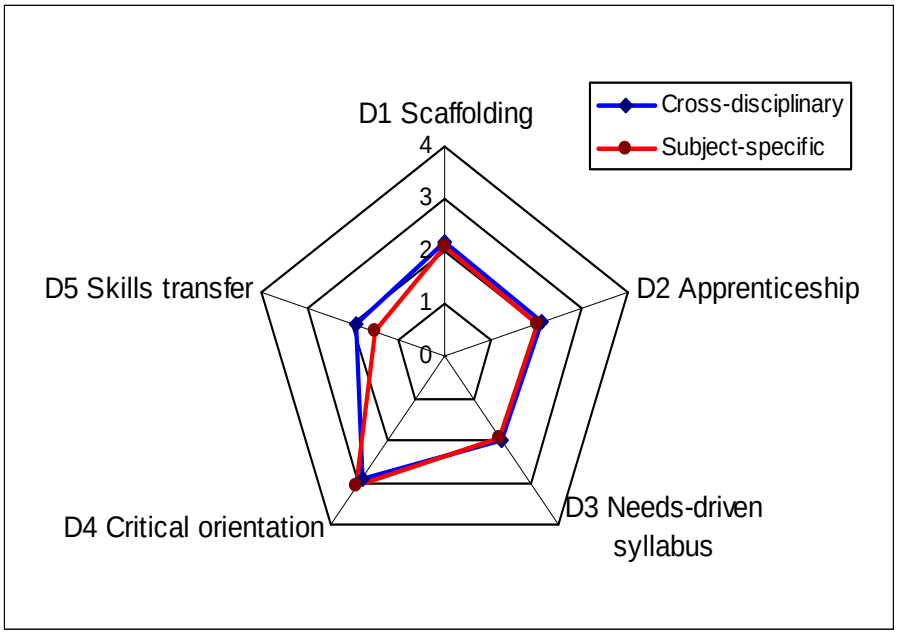


Figure 9.2 Spiderweb plot of the means of the subject-specific and the cross-disciplinary groups in the opinion surveys

From the graph it can be read that, on average, both groups felt reasonably positive about the way that a staged and scaffolded teaching and learning intervention assisted them in improving their academic writing skills (dimension 1); that both groups were, to a large extent, convinced of the positive effects of learning as a member of a discourse community (dimension 2); and also that the module had addressed their personal needs and goals reasonably well (dimension 3). On the other hand, both groups were uncertain as to the effect that the course might have had on their development of a critical

orientation (dimension 4). Although both groups were positive about the transferability of the skills they had learned (dimension 5), a predictably higher rating on this dimension (between positive and extremely positive) was obtained from the subject-specific group.

According to the Mann Whitney U-test (compare Table 9.5) the overall opinion of the two groups about the intervention did not differ significantly. As suggested by the differences in the statistical means for Skills transfer, a significant difference (at the 10% level) was found on this dimension:

Table 9.5 Two-sided p-values of the opinion survey findings regarding the 5 theoretical dimensions, obtained from the Mann-Whitney U-test

Dimension	p-value
1 Staged and scaffolded teaching and learning model	0.209
2 Purposeful social apprenticeship	0.260
3 Needs-driven syllabus	0.568
4 Critical orientation	0.130
5 Skills transfer	0.081
TOTAL	0.860

The subject-specific group was thus more inclined to think that the skills they had learned in the course were indeed transferable to other contexts. Further analysis of the data showed that although some students in the cross-disciplinary group were convinced that they could apply what they had learned to more than one discipline, others were much less positive about the transferability of the skills.

9.5 Conclusion

From the multifaceted comparison described in this chapter it can be concluded that both the subject-specific and the cross-disciplinary interventions were effective in their own right. In both cases there was a significant improvement in students' writing abilities between the pretest and the posttest: For the subject-specific intervention a p-value of 0.002 was obtained, while a p-value of 0.001 was obtained for the cross-disciplinary group.

A statistical comparison of the performance of the two groups reveals that the students who took part in the subject-specific intervention improved significantly more than those who took part in the cross-disciplinary intervention. Percentage-wise the subject-specific group improved by 19% overall, while the cross-disciplinary group improved by 8% – a difference which proves to be significant according to the Mann-Whitney U-test: $p = 0.004$. The improvement of the subject-specific group was also more consistent across the four dimensions of the scoring instrument than the improvement of the cross-disciplinary group.

Although both groups expressed fairly positive opinions about the intervention in general, the subject-specific group was significantly more positive than the cross-disciplinary group about the transferability of the skills they had learned. A p-value of 0.086 was obtained on the Mann-Whitney U-test, which means that the difference is significant at the 10% level.

These results indicate that genre-specific writing interventions can be effective, whether narrowly or more broadly focused. However, interventions that are more sharply focused on a particular discipline seem to be more effective, primarily as a result of enhanced motivation and more profound engagement with the subject matter through reading and writing with clearly delineated disciplinary foci.

Although language proficiency, especially grammar, might not noticeably improve through explicit teaching of lexicogrammar, the findings of this study indicate that a greater awareness of the lexicogrammatical resources can be facilitated through explicit teaching and tasks that make use of authentic materials.

Chapter 10: Conclusion

10.1 Introduction

The following research questions were formulated in the first chapter to address the issue of undergraduate students' inadequate academic writing abilities: (1) Can genre-based approaches be justified theoretically? (2) How effective are genre-based academic literacy interventions? (3) Which are more effective: specific or generic approaches? Question 1 relates to the input for and justification of the proposed applied linguistic design, while questions 2 and 3 relate to implementation and evaluation of two variations on a particular language teaching approach. Figure 10.1 shows how the research questions have been accommodated in the research design:

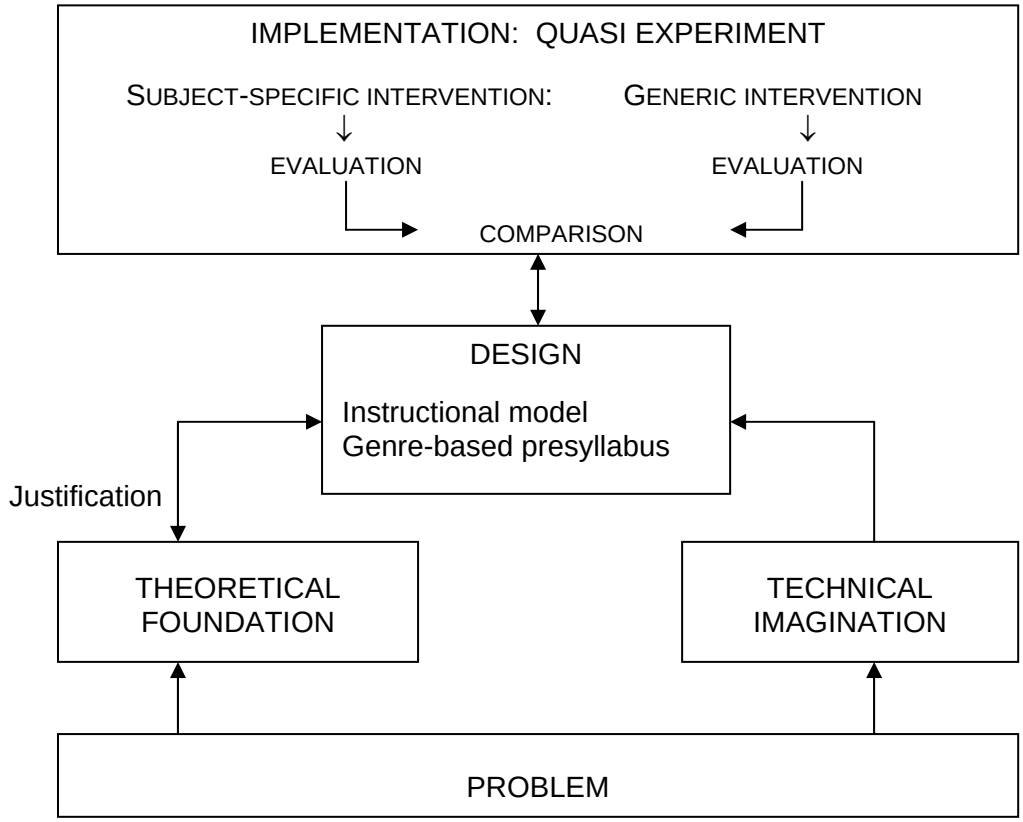


Figure 10.1 Summary of research strategy to address the research questions

This chapter attempts to indicate to what extent the research questions have been answered in order to make evidence-based recommendations for the design of future

academic writing interventions aimed at undergraduate students in the humanities. First, the theoretical justification of genre-based writing interventions is summarized. This is followed by an overview of the effectiveness of narrow-angled versus wide-angled interventions, and the significance of the difference between the two, as proven by the empirical research. Finally, some limitations of the study are briefly discussed, and recommendations are made regarding the application of the knowledge gained throughout the research process.

10.2 Theoretical justification

Genre approaches to teaching academic literacy have drawn from a diversity of linguistic, applied linguistic and language teaching theories. Figure 10.2, which should be read bottom-up, gives a schematic overview of the theories that underpin genre-based approaches:

<table>
<tr><td colspan="5">WRITING PEDAGOGY</td></tr>
<tr><td colspan="2">Skills-based approach</td><td>Practice-based approach</td><td colspan="2">Text-based approach</td></tr>
<tr><td colspan="5">LANGUAGE TEACHING THEORIES</td></tr>
<tr><td colspan="2">Multiliteracies
Critical literacies</td><td colspan="2">Communicative language teaching</td><td>Traditional Approaches</td></tr>
<tr><td colspan="5">APPLIED LINGUISTIC THEORIES</td></tr>
<tr><td>Post-modernism</td><td colspan="2">Constructivism</td><td>Extended Paradigm Model</td><td>Linguistic approach</td></tr>
<tr><td colspan="5">LINGUISTIC THEORIES</td></tr>
<tr><td rowspan="2">CDA
Multimodality</td><td rowspan="2">New Rhetoric</td><td rowspan="2">Cognitive Linguistics</td><td colspan="2">Systemic Functional Linguistics</td></tr>
<tr><td colspan="2">Sociolinguistics</td></tr>
</table>

Figure 10.2 Theoretical foundations of genre-based writing pedagogies

Among the linguistic theories, Systemic Functional Grammar is the theory that is regarded to have contributed most significantly to the theoretical grounding of genre pedagogies. SFL emphasises the systematic way in which language users make vocabulary and grammar choices in particular cultural and situational contexts. This

paradigm has had a profound influence on particularly the Australian (Sydney) genre school. Other linguistic theories that have been referred to for justification of specific features of genre approaches are Cognitive Linguistics and Critical Discourse Analysis. Cognitive Linguistics foregrounds genre knowledge: knowledge of content, communicative purpose, participant roles, discourse structure, and register. Critically oriented theories of language and other semiotic systems, such as Critical Discourse Analysis, add a political dimension to genre knowledge, *viz.* knowledge of power relations and institutional processes, and also emphasize the dialogic relationship between culture, cognition and semiosis. Among the genre schools it is particularly the New Rhetoric and the Sydney schools that are associated with CDA, because of their emphasis on social and intellectual empowerment through genre knowledge, as well as their encouragement of students and professionals to challenge the hegemonic power of conventional genres.

The **theory of learning** that best supports genre approaches is Constructivism. Genre-based approaches draw strongly upon the work of Vygotsky, in particular his Zone of Proximal Development. The ZPD is supported by two pillars, *viz.* cognitive and social apprenticeship, and scaffolding. Cognitive and social apprenticeship are linked to the rhetorical notion of learning as a member of a discourse community, while a scaffolded curriculum aims at initially providing strong peer and teacher support, and then gradually removing the support until the learner is knowledgeable and confident to construct full examples of the genre independently. Vygotskyan views feature prominently in all the so-called "post-process paradigms" in academic writing pedagogy.

The methodological input that genre-based approaches have received from **language teaching theories** derives particularly from Communicative Language Teaching. However, regarding the types of activities included in genre-based teaching programmes, genre-based pedagogy also draws from Traditional approaches and Critical Literacies approaches.

When narrowing down the focus to **theories of academic writing**, it is clear that genre approaches combine Text-based and Practice-based approaches: Text-based approaches

draw on the resources of linguistic analysis to understand the functional (rhetorical) and discipline-specific nature of writing tasks. Practice-based approaches emphasize the social and discursive practices through which disciplines constitute themselves.

The version of genre-based pedagogies adopted for the present research has relied heavily on Systemic Functional Linguistics with regard to drawing on the established conventions and values of academic disciplines, and making meaningful form-function choices. This version is in essence constructivist, in that the role of the learner as an active maker of meaning is emphasized, as well as the role of the teacher and peers as engaging in dialogue with the learner to create new meaning. The approach is overtly post-process, in that it is a considered combination of language teaching principles and techniques as well as classroom activities, with sufficient opportunity for critical reflection.

In the next section an overview is given of the design and evaluation of genre-based writing interventions drawing on the above theories. In particular, the question of effectiveness is addressed.

10.3 The effectiveness of genre-based approaches in general

Although a large number of empirical and quasi-empirical studies have been conducted to establish affinities between genres, text types and disciplines at tertiary institutions, it is believed that *in situ* research is a prerequisite for designing effective interventions – in this case genre-based academic writing courses for second-year undergraduate students of the humanities at the University of Pretoria.

To chart the target landscape, a survey of writing tasks was conducted. The results showed that the *academic essay* is the written genre most frequently required by lecturers of humanities disciplines, and that academic essays are made up of a variety of rhetorical modes. Those that feature most prominently are *discussion*, *explanation*, *description* and *(critical) analysis*. Subject-fields differ with regard to the rhetorical modes they prefer, the labels they use, and the way they combine different modes. On the other hand academic essays are structured in a fairly similar way across disciplines.

They typically comprise an introduction, body and conclusion, and develop an extended academic argument, supported by evidence. However, the nature of the evidence is subject-specific.

Against this backdrop it seems that both narrow-angled (subject-specific) interventions, with a close fit between the purposes and conventions of disciplinary communities, and more wide-angled (generic) interventions, which focus on one or more genres shared by a cluster of disciplines (such as the academic essay), could be effective. This is probably the reason why these two distinct approaches still exist within the domain of language pedagogy. However, few experimental or quasi-experimental studies have been conducted to prove the desirability (or the feasibility) of either of these intervention types.

The design of any genre-based intervention is ideally preceded by thorough contextual research. For the purpose of designing and evaluating a subject-specific intervention, in-depth research was conducted on the conventions of historical writing. History was chosen as the discipline of focus for a subject-specific intervention because the academic essay has been found to be the primary vehicle for undergraduate historical writing. The choice of history as the focal subject was also purposive and convenient, since the Department of Historical and Heritage Studies had expressed interest in the project and had been willing to offer its cooperation. Main findings were that the three main purposes of historical writing are (re)telling a story, understanding and explaining why things happened as they did, and evaluating events, structures and the writings of other historians. In historical texts these concepts have been lexicalized and grammaticalized in systematic ways. Concerning **time**, seven categories, straddling the boundaries of syntax and semantics, were defined to assist course designers and students in constructing and deconstructing time in historical texts, *viz.* sequencing time, setting in time, temporal process (phasing in time), (text internal) temporal organization, temporal modality, temporal duration and tense. Two primary ways of construing cause and effect were distinguished: sequential (chronological) causal relations between external events, and "simultaneous" mentioning of causes or effects. In terms of judgment or evaluation the most important categories for the historian are Attitude,

Graduation and Engagement, as distinguished in the Appraisal framework within Systemic Functional Linguistics.

In order to design a cross-disciplinary intervention research was conducted on the relationship between disciplinary purposes and writing conventions in a number of humanities disciplines, including philosophy, sociology, psychology, history of art and political sciences. Summaries were made of the most important conventions, and exemplars of essays and parts of essays were excerpted from these sources.

A basic genre-based presyllabus, comprising one or more cycles of exploration, explicit instruction, joint construction, independent construction and critical reflection, was adapted for subject-specific and generic purposes. Emphasis was placed on rhetorical modes, logical development of an argument from the thesis statement to the conclusion, and engagement with the authors of primary and secondary sources. However, the syllabi differed with regard to the specificity of the disciplinary focus, and thus also the themes of the materials and exercises.

10.4 The effectiveness of narrow-angled versus wide-angled interventions

The statistical analyses of the essay scores show that both narrow-angled and wide-angled genre-based interventions can be effective. The overall improvement of the students in both groups was statistically significant, although the size of the improvement differed across the four dimensions of the scoring instrument. Only on the dimension of Structure and development was the improvement of the two groups roughly similar (18% in the case of the subject-specific group and 15% in the case of the cross-disciplinary group). Thematic analyses of sampled pre- and posttest essays support the finding that both groups benefited from the instruction on structure and development: both students showed a sizeable improvement in their ability to develop an academic theme systematically.

Students from both groups were generally positive about the effect of the intervention on their academic writing abilities, and indicated that their personal needs had been more than adequately addressed. On the other hand, both groups were less positive

about their acquisition of critical thinking skills, which was not surprising in the light of the responses typically given by undergraduate students to questions about their acquisition of critical thinking skills. The only significant difference between the two groups was their perceptions about skills transfer. The subject-specific group was more positive, which probably went hand in hand with the fact that they engaged more with relevant subject matter and were more motivated.

Despite their more modest overall improvement in comparison with the subject-specific students, the cross-disciplinary group exceeded the researcher's expectations in terms of their mastery of Appraisal resources, particularly Attitude, Engagement and Graduation. The steep increase in their use of Engagement markers stands in stark contrast to the decrease in the essays of the subject-specific group. The improvement in the cross-disciplinary group's mastery of Appraisal resources should probably be ascribed to the lecturer's efforts in exposing the generic students more explicitly to these resources, and designing more appropriate classroom materials.

10.5 Limitations of the study

The main limitations of the study include (1) the relatively small sample size, (2) using the scores of only one rater for the subject-specific essays, and (3) the fact that the two interventions were not administered simultaneously. The small sample size may be seen to have impacted negatively on generalization. However, the statistical tests that were chosen (the Wilcoxon signed-rank test and the Mann-Whitney U-test) compensated for this limitation, as they had been designed for small samples. The non-parallel presentation of the interventions limits comparability because the syllabus and materials for the cross-disciplinary intervention were designed with some foreknowledge of what had (not) worked well in the subject-specific intervention. This could have influenced the significance of the statistical difference found between the two groups, as well as differences found in students' use of certain lexicogrammatical resources.

10.6 Summative remarks and the way forward

Although it would be dangerous to make grand generalizations on the basis of a quasi-experiment with fairly small samples there is a clear indication that genre-based,

scaffolded interventions do assist students in mastering the structural, conceptual and linguistic resources for meaning-making in academic discourse. On the basis of the findings it is believed that subject-specific interventions have a greater chance of succeeding than wider-angled interventions. Their greater success is primarily ascribed to the enhanced motivation that accompanies students' prospects of improving their achievement in content subjects. This prediction is underpinned by the finding that the superior performance of the subject-specific group is statistically significant. Furthermore, it seems that transferability of skills – or at least students' perception of transferability – is enhanced by extensive reading and writing with a particular thematic, and by extension, disciplinary focus.

Although narrow-angled interventions seem to be more beneficial than wide-angled interventions, such interventions may, however, be less feasible in that few tertiary institutions have the resources for offering dedicated writing modules – one for each discipline. This suggests research on alternative models for subject-specific teaching of academic writing, such as collaboration with content lecturers in a team-teaching or adjunct teaching context. It also points to the exploration of a combination of narrow-angled and generic designs in the same course.

Despite the less pronounced effects of wide-angled writing interventions, they do have some effect, and are therefore better than no intervention at all. This has been demonstrated by the significant improvement of the students on the generic course. Definite advantages of cross-disciplinary interventions are the opportunities they afford for making students aware of the dimensions along which subject-fields differ, and acquainting students with the conventions that tie in with the content, epistemology and philosophical underpinnings of a range of subject-fields. If capacity is available, a department, unit or centre with responsibility for teaching academic writing should conduct research on the relationships between subject-field purposes and writing conventions in a whole range of disciplines. It should also be considered to develop genre-based training courses for tutors, who might be Masters or PhD students in the disciplines where writing support for undergraduate students is desired.

In any event, attention should be paid to gathering, designing and developing authentic materials. It is desirable to compile a database of authentic model essays to demonstrate the successful application of essay-writing principles in specific subjects, at the level of the students, and relevant to the local (at least the South African) context. Published examples of good essays lack authenticity and are often not exemplary in every respect. In addition, students should ideally focus their writing on a particular discipline for the duration of a semester-long writing module, even within the boundaries of wide-angled modules. The greater effectiveness of the subject-specific intervention has shown that immersion into the content and materials of a specific discipline enhances engagement and encourages skills transfer. Finally, students should ideally be engaged in extended writing assignments from the beginning to the end of an intervention: the more text students produce, the more significant their improvement is likely to be.