

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

**Marking time: is there a differential effect on L2
written accuracy following focused or unfocused
written corrective feedback?**

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Education

in

Teaching and Learning

at

Massey University, Manawatu,

New Zealand.

Rachel Kim Hamlin

2013

Abstract

Marking time: is there a differential effect on L2 written accuracy following focused or unfocused written corrective feedback?

(Under the supervision of Dr. Karen Ashton and Dr. Alyson McGee)

Educational assessment in Brunei Darussalam is currently in a period of change as it transitions from an emphasis on traditional assessment methods to School Based Assessment for Learning (SBAfL). This research investigated whether traditional feedback in the form of direct focused written corrective feedback (CF) or direct unfocused CF produced differential effects on the accurate use of grammatical forms by Bruneian secondary school ESL learners. The results were considered with regards to the aims of SBAfL. Using two secondary school ESL classes totaling 38 students, two groups were formed: a Focused written CF group (n=19) and an Unfocused written CF group (n=19). The results of a pre-test, post-test and delayed post-test quasi-experiment indicated that there was no differential effect in the overall accuracy rate between either group. However, the unfocused group demonstrated a significant decline in their accurate use of the irregular past tense in the post-test, although this differential effect was not present in the delayed post-test. Overall, these findings suggested that both focused and unfocused written CF were of limited pedagogical value and raised questions about their use within an Assessment for Learning context.

Table of contents

Chapter 1: Introduction	9
1.1 The context of the study	9
1.2 Purpose of the study	12
1.3 Significance of the research problem	13
1.4 Organisation of the thesis	14
Chapter 2: Literature Review	15
2.1 Scope of the literature review	15
2.2 Feedback	16
2.2.1 Definition of feedback	16
2.2.2 Variability of effect of feedback	17
2.2.3 Feedback and culture	20
2.3 The process of writing and the place of feedback in L1 writing	21
2.4 The place of feedback in L2 writing	24
2.5 Arguments against written CF for L2 writing	27
2.6 Arguments to support written CF for L2 writing	30
2.6.1 Arguments to support written CF: studies including a control group	31
2.6.2 Arguments to support written CF: studies using subsequent writing	
Tasks	34
2.7 Summary of literature review	38
Chapter 3: Methodology	40
3.1 Research approaches	40

3.2 My research approach	42
3.3 Sampling and participants	46
3.4 Research design	48
3.5 Methods	50
3.5.1 Pre-test, post-test and delayed post-tests	50
3.5.2 Weekly written CF sessions	52
3.5.3 Written corrective feedback protocol	55
3.5.4 The questionnaire	57
3.6 Ethics	59
3.7 Role as the researcher	60
3.8 Triangulation of data	61
3.9 Validity	62
3.9.1 Internal validity	62
3.9.2 External validity	64
3.10 Data Analysis	65
3.10.1 The data collected	65
3.10.2 Establishing the nature of the test data	65
3.10.3 Analysing the test data	67
3.10.4 Analysing the student questionnaire	69
3.11 Summary of methodology	69
Chapter 4: Results	71
4.1 Pre-test, post-test and delayed post-test	71
4.1.1 Calculating writing accuracy scores	71

4.1.2 Testing for homogeneity of variance	72
4.1.3 Testing for normal distribution of data	73
4.1.4 Comparison of overall writing accuracy scores	75
4.1.5 Comparison of writing accuracy scores for each linguistic feature	77
4.1.6 Post-hoc analysis of the unfocused group's irregular past tense accuracy score	78
4.1.7 Summary of key results from the pre-test, post-test and delayed post-test	80
4.1.8 The feedback sessions	80
4.2 The questionnaire	82
4.2.1 Analysis of student perception ratings of usefulness of the written corrective feedback they received	82
4.2.2 Content analysis of student perception of preferred type of corrective feedback	82
Chapter 5: Discussion	84
5.1 Research Question 1: <i>Is there a differential effect on L2 written accuracy following focused or unfocused written CF?</i>	84
5.1.2 Similarities and differences of findings to other research	85
5.1.2.1 The systematic nature of focused feedback	85
5.1.2.2 Differences between secondary school students and university students	87
5.1.2.3 Length of time taken to write	90
5.1.3 The effect of written CF on individual linguistic features	90

5.1.3.1 Rule governed and non-rule governed language structures and written CF	92
5.1.3.2 Differentiation and written CF	93
5.1.4 Summary of findings regarding research question 1: <i>Is there a differential effect on L2 written accuracy following focused or unfocused written CF?</i>	95
5.2 Research Question 2: <i>What are students' perceptions of written CF?</i>	97
5.2.1 Students' expectations of written CF	97
5.2.2 Summary of findings regarding research question 2: <i>What are students' perceptions of written CF?</i>	100
5.3 Wider implications of the study	101
5.3.1 Focused or unfocused written corrective feedback?	101
5.3.2 Focused written corrective feedback and differentiation	101
5.3.3 Student perception and uptake of written corrective feedback	102
5.3.4 Focused and Unfocused written CF and School Based Assessment for Learning (SBAfL)	103
Chapter 6: Conclusion	105
6.1 Limitations of the study	107
6.2 Suggestions for future research	110
Chapter 7: References	112
Chapter 8: Appendices	126
Appendix 1: Pre-test, post-test and delayed post-test	126

Appendix 2: An example of student writing in the delayed post-test	129
Appendix 3: The first of the eight Mousedeer narratives used in the written CF Sessions	130
Appendix 4: Grammatical error tally sheet	130
Appendix 5: Graphs displaying linguistic scores for both the focused and unfocused groups which are not normally distributed	131

List of Tables

Table 1: Design of the study	49
Table 2: Levene test of homogeneity of variance	73
Table 3: Group Means and Standard Deviation for the Error Correction Tests	75
Table 4: Group Means for the Error Correction Tests	75
Table 5: Differences in medians for each linguistic feature	76
Table 6: Differences in medians for each linguistic feature	78
Table 7: Focused Group Irregular Past Tense Means for Error Correction Tests	79
Table 8: Focused Group's Accuracy Scores in the Error Correction Tests	80

Chapter 1: Introduction

This thesis examines the role of written corrective feedback (CF) in an English as a second language (ESL) context. In particular, it seeks to investigate any differential effects between focused and unfocused forms of written CF on English as a second language (L2) students' written accuracy. The study was carried out in a girls' secondary school in Brunei Darussalam.

1.1 The context of the study

In 1985, shortly after the tiny country of Brunei Darussalam in Southeast Asia gained independence from Great Britain, it introduced a Dwibahasa (dual language) Education system within which Malay (the National language) and English were to be of equal importance (CfBT, 2006). This has resulted in a dual language secondary school system with English being the medium of instruction for all but three of the subjects taught. The subjects which are taught in Malay are Malay literature, and two religious subjects. The use of English as the medium of instruction for the other subjects is intended to support language development to enable the students to sit the Cambridge 'O' Level examinations in year 11. These are the school-leaving examinations in several countries including Brunei, Singapore and Mauritius, and in Brunei they serve as the gate-keepers to future education and employment opportunities.

Traditionally, educational pedagogy in Brunei has relied heavily on strategies of rote learning and repetition (Scott and Fisher, 2002), with examinations and tests being the main source of assessment information. However, in an effort to change its education system to one which emphasizes assessment for learning in line with trends in other

countries in the region such as Hong Kong and Singapore (Ngan, Lee and Brown, 2010), the Bruneian Ministry of Education introduced a *National Education System SPN21* in 2007. An important element of this system has been the introduction of *School Based Assessment for Learning (SBAfL)* within both primary and secondary schools. Beginning in primary school, this programme has now filtered through to lower secondary school students in years seven to nine (ages 11 to 14) and this year the first batch of students taught under this system in Brunei will be in year ten. SBAfL has mandated the use of 'formal, standardized and moderated assessment tasks that familiarize teachers and students with Assessment for Learning (AfL) best practices' (Ministry of Education, 2012). A key element to this approach identified by the Ministry of Education has been providing feedback to students about their strengths and weaknesses, from information gathered through teacher, peer and self-assessment (Ministry of Education, 2012). By including practices such as self and peer-assessment within these standardized tasks, it is hoped that teachers will become familiar with assessment pedagogies which provide students with information about their next steps in the learning process. Within the English subject this means that year seven and eight students are required to complete a series of in-class assessments, each assessment taking a period of several weeks. In time the standardized tests will be phased out as teachers become more experienced at using AfL pedagogies, and are able to develop their own student assessment tasks. Thus far, assessment for learning strategies such as those described for lower school students in years seven and eight, have not been mandated for students from the upper school in years nine to eleven. However, it is expected that upper school teachers begin to use an assessment for learning approach

in their classes, and it is intended that professional development will be provided for them regarding this later in 2013.

Students begin learning English in Brunei at the age of five when they first enter school, and receive five hours a week of English language class, with all other classes being taught in the Malay language. From primary four onwards Mathematics and Science are also taught in English although a certain amount of code-switching between English and Malay does occur (Martin, 1996). By the time students reach secondary school, the medium of instruction for all classes except the two listed above is English. English language is taught as a separate subject for three hours a week with a heavy emphasis on the accurate writing skills required for the Cambridge 'O' Level examination.

Although AfL is the goal identified by the Ministry of Education, many current practices such as the heavy emphasis on teaching examination writing skills don't support this goal. Within English language classes in secondary school, the feedback most commonly provided to students comes in the form of written corrective feedback. It is mandatory for English teachers to collect a composition from students at least once a fortnight and to respond to it in the form of error correction (CfBT, 2006). From my own experiences and observations as an English teacher working in Brunei, this generally means that teachers cross out the students' incorrect word or phrase, and write the correct form above it. Despite being a compulsory part of the assessment programme, no research has been conducted within Brunei investigating the effectiveness of this practice in improving students' writing accuracy. Furthermore, as Brunei tries to implement an assessment for learning approach to education which provides students with feedback regarding how to bridge the gap between their current knowledge and

desired knowledge, it is relevant to consider the value of written CF at this time as it is one of the most frequent sources of assessment information provided for students in the English classroom.

1.2 Purpose of the study

The study was designed to explore any differential effects between focused and unfocused written CF. By using Bruneian secondary school students, it investigated the effects of focused and unfocused written CF on their English L2 written accuracy. It also considered how useful the students perceived the written CF they received to be. Providing written CF in response to students' writing is a well-entrenched teaching strategy within this educational context. However, as Brunei is currently introducing AfL strategies within the framework of its SBAfL policy, it is pertinent that written CF be examined at this time.

In light of the recent research (outlined in Chapter 2) which found evidence supporting the claim that written focused CF is more effective than unfocused CF (Ellis, Sheen, Murakami and Takashima, 2008; Bitchener and Knoch, 2010-a; Bitchener and Knoch, 2010-b; Sheen, Wright and Moldawa, 2009), and with consideration of the manner by which CF may be effected by culture (Sully de Luque and Sommer, 2000), this thesis investigates the following research questions:

1. Is there a differential effect on L2 written accuracy following focused or unfocused written CF?
2. What are the students' perceptions of written CF?

1.3 Significance of the research problem

The education system of Brunei Darussalam is in a time of immense change as in 2007 a new National Education System SPN21 (Ministry of Education, 2012) was introduced which strives to achieve a curriculum 'emphasising assessment for learning' (p.10). As part of this initiative, teachers have been encouraged to examine the ways they use assessment and feedback in the classroom. At the time this thesis research was conducted, I was employed as an English teacher within Brunei Darussalam for three years in an all-girls' government secondary school. This meant that I observed and experienced first-hand the changes the Ministry of Education is attempting to implement, as well as the actual pedagogical practices of teachers at the grass roots classroom level.

Within English language classes in secondary schools, feedback is most commonly given to students in the form of written CF on their writing. From my own observations, formative assessment techniques such as peer and self-assessment as emphasized by SPN21 are rarely utilized. This may represent a tension between traditional assessment techniques such as written CF which are favoured by teachers, and the types of AfL tasks now required by SPN21. It is pertinent therefore, given its prevalence of use, to investigate the effectiveness of written CF in order to contribute to the wider dialogue about effective feedback practice within Brunei Darussalam.

Another significance of this study is that written CF has not been investigated within Brunei Darussalam before. Furthermore, much of the previous written CF research has been conducted within a Western context. Therefore, an important contribution which this study can make to classroom feedback research is to add evidence regarding

differential effects between focused and unfocused written CF within a Southeast Asian Islamic context to the current body of research. This is important particularly as culture has been identified as a variable which can affect the way students respond to feedback (Sully de Luque and Summer, 2000).

1.4 Organisation of the thesis

This thesis consists of six chapters. This chapter, Chapter 1 has provided a background of CF research and stated the purpose and significance of the study. Chapter 2 outlines the context of the study, and then is devoted to reviewing the relevant literature on feedback to writing within both first language (L1) and L2 fields, before turning more specifically to discuss written CF. Chapter 3 presents the methodology of the research by describing the research approach, design and strategy used, as well as issues of ethics and validity. It will then describe the data analysis techniques of the study. Chapter 4 presents the results of the study, and these results will be interpreted and discussed in chapter 5. Finally, Chapter 6 will summarise the main findings and conclusions of the study. Limitations of the study and directions for future research will also be included in this chapter.

Chapter 2: Literature Review

This literature review summarises salient research and theory regarding the practice of providing CF about ESL students' writing.

This review first outlines the context of the study, before moving on to discuss relevant literature regarding theories of feedback. It will then more specifically discuss feedback on writing within the fields of both first and second language learning, and will conclude with a discussion of written CF literature within an ESL context.

2.1 Scope of the literature review

Comparatively very little educational research has been conducted in Brunei. A search of the online educational journal database ERIC revealed only 43 studies of English language learning in Brunei, and of those a large number were concerned with the bilingual education system of the country rather than English language teaching strategies and acquisition. No studies were found which related to the use of CF within Brunei despite the popularity of the practice. Therefore, within this literature review relevant research from different cultural contexts is discussed.

Additionally, very little research directly relating to the use of CF within a secondary school context has been conducted. The vast majority of studies have focused on university students as subjects. However, the practice of providing written CF about written accuracy described in the studies is similar (if not the same) as that I have observed frequently within a Bruneian secondary school. Thus, it was pertinent to include findings from different types of educational institutions within this review.

The literature summarised not only comes from a variety of cultural and institutional settings, but also from a number of academic disciplines. L2 writing theory can be described as a subset of SLA theory, which is also influenced by other disciplines such as English as a First Language (L1) writing theory and communication studies.

Therefore, relevant L1 writing research as well as L2 writing research from SLA is included in this literature review.

Having introduced the context for this study, and scope of literature reviewed, the focus now turns to a discussion of relevant aspects of the literature. Theoretical conceptualisation of feedback and relevant research will first be examined within a wider educational context, and then within an L1 and L2 context. Subsequently, specific evidence for and against written CF as an effective pedagogical practice will be discussed in order to gain a balanced understanding of the debate as it stands.

2.2 Feedback

This section will define feedback and summarise pertinent findings regarding its effectiveness as a teaching and learning tool.

2.2.1 Definition of feedback

Feedback is the process in which information about student output is provided in order to modify future learner understanding, motivation and/or behaviour (Kluger and DeNisi, 1996; Shute, 2008). For example, a teacher may comment about a student's use of adjectives within a composition, suggesting ways for improvement in subsequent compositions. Feedback is an instructional practice which can enhance both students' skills and motivation because it can lead to greater student efficacy (Brown, 2004;

Bruning and Horn, 2000; Kluger and DeNisi, 1996). CF is a type of feedback which is a teacher's reaction to a learner's errors and is a source of evidence for the learner about what they have not done well with regards to a desired standard (Adams, Nuevo and Egi, 2011). It may be given in a written form as is the topic of this thesis, or orally in the mode of teacher verbal recasts or corrections of student errors.

Feedback provides methods for reducing the gap between a student's actual performance and a targeted performance (Sadler, 1989). For example, being able to write using paragraphs may be an element considered necessary for good writing. A teacher may give a student feedback information about their use of paragraphs in a piece of writing with the goal of helping them achieve the goal of well-structured paragraphs in future writing.

2.2.2 Variability of effectiveness of feedback

Feedback is an important area to study as according to Kulhavy (1977) not all types of feedback are of equal educational value because feedback does not always lead to positive learning as it can be accepted, modified, or rejected. The type of feedback provided may also not be particularly useful. In fact, if the material being studied is completely incomprehensible to the student, providing feedback will probably have no effect on improving achievement as there is 'no way to relate the new information to what is already known' (Kulhavy, 1977, p. 220). In other words, if the gap between what is already known, the desired knowledge too great, and if the feedback does not bridge that gap one step at a time, it is difficult for the student to use feedback positively to

reach the goal. Therefore, we cannot merely assume the benefit of particular types of CF without careful consideration and investigation.

This variability of the effectiveness of different types of feedback was illuminated by Hattie (1999) in a synthesis of over 500 meta-analyses of research investigating factors influencing educational achievement. Twelve of these meta-analyses included specific information on feedback in the classroom. While there was some evidence that feedback can be powerful, the effect sizes demonstrated a large deviation, suggesting that some types of feedback are more powerful than others. Namely, students receiving feedback about a task and how to do it well had the greatest effect size, while praise, rewards and punishment had the lowest effect size. This means that feedback needs to be specifically task related if it is to be effective.

As well as being task related, another meta-analysis of studies of feedback interventions (Kluger and DeNisi, 1996) found that feedback was more powerful when it provided information on correct responses rather than incorrect responses, when goals were specific and challenging, and when task complexity was low. One reason they suggested to explain this was that if goals are not specific enough, the student's goal may in fact differ from that intended by the teacher giving the feedback, leading to student confusion about steps for future improvement. For example, if a teacher gives a student vague feedback about needing to improve paragraphs but doesn't give specific steps for how to achieve that, the student may not understand what is required of them and change their paragraphs in a manner which does not reflect the unspoken goals of the teacher. Thus, effective feedback needs to be both task related as suggested by Hattie (1999) and also specific.

Hattie and Timperley (2007) suggest feedback is effective when it deals with faulty interpretations, rather than a total lack of understanding. In fact, they go as far as to say that if there is a complete lack of understanding, feedback may even be damaging. If a student is completely confused about a concept, it is important for them to comprehend the basic constructs first, before being given feedback about how to achieve the highest level. Feedback about concepts which are not understood at all can create confusion and potentially have a de-motivating effect on students. Thus, it is possible that if students have no understanding of a particular linguistic feature which is important for English written accuracy, providing CF may have a negative impact. Students could become de-motivated by feedback which they do not understand, or it may be that the feedback is not specific and goal focused enough to lead to improvements in written accuracy.

The research presented in this section has highlighted both the potential benefits for students from receiving feedback, but also that certain types of feedback can have unintended negative effects. Feedback is most effective when it is task specific, goal related, and deals with faulty interpretations rather than a complete lack of understanding. Hence, it is important that classroom feedback practices are examined rather than their benefit merely being assumed. It is with this in mind that the focus of the literature review now turns to consider the relationship between feedback and culture.

2.2.3 Feedback and culture

In addition to the findings mentioned above regarding the characteristics of effective feedback, the manner in which students interact with feedback is also of importance. In order to improve performance, feedback has to be paid attention to (Anseel, Lievens and Schollaert, 2009). For this reason, several theorists interested in feedback interventions have argued that perception of feedback, and reflection on feedback are important aspects of the feedback process (e.g. Anseel et al, 2009; Hall, 2002). Student reactions to and reflection on feedback are vital to the acceptance and use of the feedback (Cawley, Keeping and Levy, 1998), and these may be affected by students' cultural characteristics (Sully de Luque and Sommer, 2000). In particular, researchers have found that people in hierarchically structured cultures (such as Brunei) are less trusting of supervisors' feedback than those from other cultures (Earley and Stubblebine, 1989).

Hofstede (1984) suggested that cultures differ on four basic dimensions: power distance; individualism; uncertainty avoidance; and masculinity. To date, much research regarding feedback has taken place in Western cultures which have low power distance and high individualism. Within Brunei, there is potentially a high power distance between the teacher and student, and a low level of individualism. Within the framework of Hofstede's model, where there is a perceived large gap between people in positions of power such as the teacher and students, while order and discipline are maintained, the lack of trust can mean that student uptake of feedback is low. Furthermore, societies with a low level of individualism tend to not like actions which impact orderliness or the status quo. This could affect students' perceptions of different types of feedback other

than those they are used to. If, as Hofstede suggests, collectivist societies such as Brunei do not like change, then this may impact upon the implementation of AfL techniques as required by SBAfL.

Therefore, cultural constructs and perceptions about CF are also relevant issues to consider if research findings are to be of real world value. While not the main focus of this thesis, it will be important to get student reflections on the CF they were exposed to in this research in order to get a clearer picture of the feedback process within this particular cultural context. Having established the nature of effective feedback, and highlighted the need to consider culture as a factor in any research, this discussion now turns to the process of L1 writing and feedback.

2.3 The process of writing and the place of feedback in L1 writing

The way that teachers give feedback on students' L1 writing has been the subject of much more research than that of L2 writing. As they are closely related fields it is worth summarizing some of the most relevant literature here before looking at findings related specifically to L2 writing. These will be discussed chronologically in order to trace the main developments in this area.

In a cognitive analysis of writing, Emig (1977) made a connection between writing and learning, claiming that writing is a learning activity that responds well to feedback because unlike speaking and listening, it is entirely a learned behaviour. People do not learn to write unless they are taught. Feedback is one of the methods by which writing is learned, which is made easier by the fact that writing also produces easily accessible evidence to give feedback to. Hayes and Flower (1980) further extended Emig's work

by arguing that writing is a process which includes planning, translating and reviewing. Writing is not a single act, rather it is made up of several distinct phases consisting of gathering together ideas about which to write, translating those ideas into writing on a page, and then reviewing the writing to see to what extent it communicates the desired meaning. Thus, they argued that it should be viewed as an extended process within which an integral element is reviewing the quality of the writing. This placed cognitive actions at the heart of writing, thereby emphasising the potential role of student interaction with external feedback in the writing process.

Writing as a process consisting of distinct stages has been very influential in research about improving student writing. Using a protocol analysis, during which students verbalised their thoughts while they were writing, Hayes and Flower (1983) expanded on their earlier research and found significant differences between the ways that novice and expert writers handled the writing process. Their findings suggested that skills involved in writing as a process such as planning are teachable, and therefore classroom feedback should focus on the writing process rather than surface level structures. Therefore, teachers have been encouraged to respond to overall problems of planning and content in student writing (Griffin, 1982).

Yet despite this encouragement to provide process-oriented feedback to students, Zamel (1985) found that many L1 teachers still treated what should be considered a first draft as the final product, giving feedback on both surface-level features and content areas to be developed. This, she claimed, made it difficult for students to know which type of error they should focus on, and hindered learning. In fact, Zamel noted that frequently the students attended to the surface-level features identified in feedback in

subsequent drafts in preference to the content areas to be developed about which feedback was also provided.

Others have been more critical of the process-oriented approach. For example, Horowitz (1986) argued that this approach to writing failed to prepare students for academic writing, is only suitable for some writers and tasks, and gives students a false impression about how their writing will be evaluated at University. Not all writing situations allow for a process of drafting and re-drafting in response to feedback beyond the classroom. Furthermore, he claimed that emphasizing the process-oriented approach ignores the fact that other approaches such as the correction of surface-level features can be equally successful.

Whether or not teachers respond to surface-level errors, or form-based areas for improvement, there is much support for the use of feedback in the writing process. A revised cognitive analysis of writing conducted by Hayes (2000) is similar to information-processing models of self-regulated learning (e.g., Butler and Winne, 1995; Winne, 1997, 2001). Both of these models of learning depict students as active agents involved in constructing their own knowledge, with external feedback emphasised as a tool in this process.

Thus, feedback given to writing is an important area worthy of investigation. Writing is an activity which provides evidence of a student's understanding of the English language, and as such it is possible for a teacher to provide feedback to writing. Feedback is imperative for students to engage with if they are to become active learners

and develop their English writing skills. Having established the importance of feedback in L1 writing, this review will now examine literature concerning its role in L2 writing.

2.4 The place of feedback in L2 writing

This section provides a general overview of the main theoretical positions regarding L2 language development and pedagogy, while the following sections provide a more detailed discussion specifically about the effectiveness of written CF at improving student written accuracy.

Traditionally, within L2 teaching, writing was seen only as a means to practice vocabulary and grammar, and all errors were corrected (Brown, 2007; Ferris, 2010; Ferris and Hedgecock, 2005). Other theorists argued that errors in language production are a natural part of a developmental process similar to that of a child acquiring a first language (e.g. Corder, 1967; Krashen, 1982). For example, a young child learning English as their mother tongue may still make grammatical errors such as calling two mice 'mouses'. As they mature their language skills develop and 'mouses' is replaced with the correct form 'mice'. The theory that L2 learning is also subject to a natural developmental process led to a debate between those favouring meaning-focused instruction, and those favouring form-focused instruction.

The meaning-focused argument developed at a similar time to the writing-process approach to writing instruction as outlined above. As the meaning-focused orientation to writing was applied to the L2 context, it was argued that grammar instruction and error correction should be minimised in order for students to develop their ideas through the process of drafting, receiving feedback and redrafting (Ferris, 2010). In a manner similar

to the process-oriented approach to feedback of writing in L1, Krashen (1984) advocated delaying feedback on errors until the last stage of editing.

Meaning-focused instruction emphasises the notion that like L1 acquisition, L2 acquisition occurs unconsciously and in natural stages. Some supporters of this argument claim that any overt attention to grammatical form is unnecessary, and error correction is ineffective (e.g. Krashen, 1981; Truscott, 1999). The best example of this philosophy in practice can perhaps be found in language immersion contexts.

Originating in Canada in the 1960s, immersion schools place students into an educational context surrounded by the target L2 in the belief that their L2 acquisition will develop naturally in a L2 rich environment (Cummins, 1998). Yet, despite many years of target language input, Canadian immersion program research suggests that students' production of L2 can still be grammatically inaccurate (Swain, 1985). This indicates that meaning-focused instruction does not necessarily lead to grammatical accuracy. Thus, form-focused instruction such as explicit grammar teaching and error correction could be an essential aspect of the development of L2 written accuracy.

Several L2 writing theorists have argued for form-focused instruction (e.g. Eskey, 1983; Horowitz, 1986; Silva, 1988). Form-focused instruction emphasised explicitly targeting linguistic structures in order to draw the student's attention to desired language elements (Ellis, 2001) and to 'help them make more efficient use of their limited exposure to the sounds, words, and sentences of the language they are learning' (Spada and Lightbown, 2008, p. 182). Grammar instruction is an essential aspect of this approach which can be defined as 'any instructional technique that draws learners' attention to some specific grammatical form in such a way that it helps them either to

understand it metalinguistically and/or process it in comprehension and/or production so that they can internalise it' (Ellis, 2006, p. 84). Written CF is one of the techniques by which language teachers have traditionally drawn learners' attention to grammatical form.

In the last few decades, L2 writing theory has alternated from favouring meaning-focused feedback to favouring form-focused feedback, until it finally settled on a middle ground accepted by most researchers (Ferris, 2010). While this middle ground recognises a place in language instruction for feedback on both meaning and form, it has generally been acknowledged that in contrast to L1 writing, L2 writers often need more instruction and feedback about the language itself (Myles, 2002). They may have limited knowledge of vocabulary, language structure and content, which can hinder their writing performance and thus L2 students require more form-focused feedback than L1 learners. Consequently, within a process-oriented model of writing, L2 students generally want more teacher guidance and feedback on form, particularly at the revision stage (Myles, 2002). The current viewpoint is by and large that L2 written instruction should include CF on form: exactly how much and in what form is still a topic of debate (Ferris, 2010). More research such as the current study is required to shed light on the effectiveness of CF on form within writing teaching programmes, in order to clarify how it might best be used.

With this in mind, a more detailed discussion of research investigating written CF of form within a L2 context follows, beginning first with arguments claiming that it is not an effective practice. Having examined research which failed to find any positive effects of written CF, possible reasons for this and evidence which does provide support for

written CF will then be discussed. It will conclude by presenting studies suggesting a differential effect between types of written CF, and on which knowledge this thesis hopes to build.

2.5 Arguments against written CF for L2 writing

In addition to the debate over meaning-focused feedback or form-focused feedback, several studies claimed to empirically support the position that written CF does not aid students' written accuracy. These will be summarised below beginning with the earliest relevant study, and concluding with the most recent.

Stiff (1967) studied the effect of teachers providing corrections in the margins, and at the end of a piece of writing on the compositions of 77 first year university students in the USA. Students were divided into three groups, receiving one of three types of correction: content-focused correction written in the page margins; form-focused correction written at the end of the composition; and a control group which received both the marginal and terminal correction. No statistically significant difference was found between any group, and Stiff therefore claimed that the placement of the feedback was not related to an improvement in student accuracy. However, the group which Stiff referred to as the 'control' group actually received written CF. Thus, at most, this study was able to claim no differential effects between the three configurations of written CF, as it may be that all three groups improved significantly more than if they did not receive any CF at all.

In an investigation including both meaning-focused and form-focused written CF Semke (1984) studied the journal writing of German foreign language students at a university in

the USA. In the 10 week study, students were divided into four groups either receiving direct error correction, feedback indicating the type of error with a code only, comments on content only, and a combination of direct correction and comments only. She found no significant difference in the accuracy of the four groups at the end of the ten week period. The only difference noted was that all three correction groups wrote more slowly than the comments only group, but this had no effect on accuracy.

Kepner (1991) analyzed student journal entries for Spanish foreign language students at another university in the USA. The students were randomly assigned to either a group which received written CF, or one which received a comment and no correction. Their writing accuracy was compared after a period of 12 weeks and no significant difference in errors was found. However, Truscott (2007), in a summary of CF studies, notes that students seemed to shorten and simplify their writing, possibly suggesting that students who received error correction were hesitant to use ambitious language structures in case they were wrong. This is comparable with Semke's (1984) observations that students in the correction groups wrote slower than the comments only group, perhaps due to hesitation regarding errors. The effect of written CF on the speed of students' writing is an aspect worth further investigation. Within an AfL framework it is important that students feel comfortable attempting more ambitious language structures, and are able to learn from the feedback about their errors rather than avoid them. This is particularly relevant to the current study as it draws to attention the potential tension between the current pedagogical practice of providing written CF, and the AfL philosophies emphasised by SBAfL.

In another investigation comparing form-focused and meaning-focused feedback, Sheppard (1992) compared the writing accuracy between two groups of students; one which received indirect error coding CF and another which received holistic comments. Students were asked to write seven compositions over a ten week period, which were then analysed based on students' accurate use of verb tense, punctuation and subordination. Students who received holistic comments made significantly greater improvement in writing accuracy than those who received error corrections. Furthermore, Sheppard also found that the CF group began to avoid complex language structures over time, possibly suggesting a negative effect of CF as found in the studies above. Thus, given the firmly entrenched status of written CF as a teaching practice in Brunei Darussalam, it becomes even more important to conduct research which investigates whether or not it has an effect on student written accuracy. This is particularly salient due to the emphasis of assessment for learning within the recent Bruneian education initiatives described in section 2.2.

None of the studies discussed above found any significant statistical difference in the accuracy of writing between groups of students receiving CF, and those receiving holistic or content focused comments only. Thus, it would seem that there is not adequate evidence supporting the well-entrenched language teaching technique of error correction. Truscott argues that if there is any benefit to grammar correction it is 'so small as to be uninteresting' (2007, p. 256), and more likely actually has a small negative effect as Sheppard's (1992) research seemed to suggest. Reviewing academic research supporting the use of CF, Truscott (1996) went so far as to argue for the abandonment of grammar correction from language courses altogether. He claimed that

researchers have paid too little attention to the side-effects of grammar correction such as the effect on student attitudes, motivation, the use of time, and energy in writing classes (Truscott, 1996). Indeed, Truscott (1999) claims that research which has been conducted tends to assume the benefit of grammar correction, while not actually answering the question of whether or not CF enables students to make greater gains in writing accuracy than no feedback.

In Brunei, the practice of providing written CF has become established despite no evidence suggesting that it is effective. In fact, it seems that written CF may lead to student avoidance of complex language structures (Kepner, 1991; Semke, 1984; Sheppard, 1992). Now, with the introduction of assessment for learning pedagogy within Brunei, research is required which may shed light on the usefulness of written CF within this context and whether or not written CF provides assessment information which aids student learning. While not providing adequate negative evidence against the use of written CF, the studies outlined above did raise questions about the efficacy of written CF, and further investigation in this area is required. This review now moves on to consider literature which claims benefits of CF on students' L2 written accuracy.

2.6 Arguments to support written CF for L2 writing

In contrast to the studies outlined above, several researchers have investigated the use of written CF and claim that it is of benefit in improving students' written accuracy. Their main findings are summarised below.

Bitchener and Knoch (2008) reviewed numerous studies which investigated the efficacy of written CF as opposed to writing practice alone. They divided these studies into

those with and without a control group. All five studies without a control group reported a significant improvement in grammatical accuracy after CF (Chandler, 2000; Ferris, 1995, 1997, 2006; Lalande, 1982). However, the absence of a control group raises questions about the validity of all of these studies, as improvement may have occurred through writing practice alone and increased classroom tuition alone, without any CF (Truscott, 1996). Without a control group, it is possible to investigate differential effects of different types of CF, however, it is not valid to claim that written CF is of more benefit than no written CF. It is salient to note that it was not possible within the current study to include a control group for reasons that will be outlined in section 3.3. Thus, the focus of the study is potential differential effects between types of written CF rather than comparing written CF with no CF.

Several studies which did include a control group have been conducted investigating feedback on form versus feedback on content, and the effectiveness of different types of written CF. It is to these this discussion will now turn.

2.6.1 Arguments to support written CF: studies including a control group

Fathman and Whalley (1990) studied feedback using intermediate level ESL college students in the US. Students were asked to write a composition describing a series of pictures. As well as having a control group, students were put into two other experimental conditions: a correction only group where the sole corrections were the underlining of grammatical errors; and a correction and content comment group. Corrected compositions were returned a few days later, and students were given half an hour to rewrite and improve their previous composition. Analysis of the accuracy of

students' writing showed an improvement in both the correction only, and the comment and correction groups. Fathman and Whalley concluded that feedback on both form and content, given separately or together, had a positive effect on students' rewriting.

Ashwell (2000) investigated the effect of written feedback throughout the writing process. He examined the effect of the order the feedback was presented on a writing process involving a first draft, second draft and final draft. Students either received content-focused feedback in the first draft, and form-focused feedback in the second draft, the same feedback in the reverse order, form and content feedback simultaneously, or no feedback at all. They were then asked to edit their compositions. Results showed no difference between feedback groups in the effect of the order of feedback, but all feedback groups outperformed the non-feedback group with regards to the formal accuracy of their writing. However, feedback did not seem to have an effect on the content of writing as improvements in terms of content were recorded for both feedback and non-feedback groups.

Thus, both Fathman and Whalley (1990) and Ashwell (2000) concluded that written CF in the form of grammar correction was of benefit to students' written accuracy. However, it may also be the case that one type of written CF is more effective in improving students' written accuracy than another. Written CF can be direct, in which the correct linguistic form is provided by the teacher, or indirect in which an error is indicated by the teacher but no correction is provided (Bitchener and Knoch, 2008). Also, within both of these categories CF exists on a continuum of more and less explicit feedback.

Ferris and Roberts (2001) examined how explicit indirect feedback should be. Students were assigned to one of three groups; CF using five common error codes, CF underlining the same five errors but providing no code, and no feedback. Students were asked to produce a 500 word composition at home, and then improve it after the experimental conditions had been applied. Ferris and Roberts noted an improvement in accuracy in both feedback groups which was significantly greater than that of the control group, although there was no difference in the degree of improvement between the feedback groups. This led them to argue that feedback does indeed help students to improve the accuracy of their writing, and the level of explicitness did not make a difference.

While Ferris and Roberts (2001), Ashwell (2000) and Fathman and Whalley (1990) seem to provide evidence for the use of CF in the language classroom, another factor needs to be taken into account when considering the relevance of the findings. None of the studies considered the effects of CF on new pieces of writing. Therefore, they were only able to show that CF can aid students in achieving greater writing accuracy in a second draft of the same composition (Ellis et al. 2008). Whether or not an improvement in writing accuracy transferred to gains in subsequent writing, cannot be deduced from these studies. As Sheen (2007) points out, '(to) claim that error correction results in learning, one must examine whether the improvement in revisions carries over to a new piece of writing or if the improvement is manifested on a post-test or delayed post-test' (p. 258). For this reason, it is important that the current study includes subsequent writing tasks within its design. This literature review will now consider existing research which did examine the effect of written CF on subsequent writing tasks.

2.6.2 Arguments to support written CF: studies using subsequent writing tasks

Bitchener, Young, and Cameron (2005) conducted a study investigating the effects of correction and teacher discussion on the development of English grammar prepositions, simple past tense and definite articles. Adult learners were divided into three groups. Group one had explicit correction, student-teacher conferences and 20 hours of English instruction a week. Group two had explicit correction and ten hours of English instruction per week. The third had no correction or conferences and only had four hours of English instruction per week. Despite the difference in class exposure, Bitchener et al. claim that students all had the same amount of instruction in grammar and writing per week. They found that CF led to improved accuracy in both the use of the simple past tense and articles, but not prepositions over a 12 week period. However, Bitchener et al. argued that CF did lead to improved written accuracy. It is worth mentioning as Truscott (2007) pointed out, that the control group received fewer instructional hours overall so the groups were not strictly comparable. It is possible that the written accuracy gains related to the extra instructional time provided, rather than the use of written CF.

Sheen (2007) studied the effects of written CF on the acquisition of English articles by intermediate ESL learners. Students were divided into a direct CF group (the researchers indicated errors and provided correct forms), a direct meta-linguistic group (the researchers indicated errors, provided correct forms and meta-linguistic explanations) and a control group (received no CF or explanations). Acquisition was measured in three different ways; an error correction test, a speeded dictation test, and a guided composition. The effects of the CF were statistically significant on all three

tests in comparison to the control group indicating that CF did contribute to the acquisition of English articles. However, this research focused solely on a single grammatical error (the articles 'the' and 'a'), which may have made it easier for students to process information.

This is the view of Sheen, Wright and Moldawa (2009) who after reviewing the findings of Sheen (2007) suggest that some previous investigations into CF have failed to show any effect because the CF 'was not sufficiently focused and intensive' (p. 559). By focussing on only one grammatical error (articles), they argue that CF was more effective than if it had been provided for a range of grammatical errors as in previous studies. The traditional approach to CF tended to provide correction on a wide range of errors in a wide range of students' writing. This 'unfocussed' CF may run the risk of overloading students' attention capacity (Sheen, 2007). Alternatively, Sheen suggests CF could follow a more focused approach by selecting a specific grammatical error noted in student writing, and focusing on this over a limited period of time. Thus, questions emerged regarding whether written CF should deal with all the linguistic features of written accuracy at once in an unfocused approach, or select specific features in a more focused approach (Sheen, Wright and Moldawa, 2009).

This question has been investigated by SLA research into oral CF, which found that CF was effective when it was focused on a specific linguistic feature repeatedly (e.g. Ellis, Loewen and Erlam, 2006; Lyster, 2004). This suggests that focused CF may potentially be more effective than unfocused CF in the written form too. Further investigation is therefore required to find out if this was also the case with written CF.

Furthering the research conducted by Sheen (2007), Ellis et al (2008) compared the effects of focused CF and unfocused CF on Japanese university students' accurate use of definite and indefinite articles in three narrative compositions. A focused group received corrections of article errors only, an unfocussed group received corrections of article errors as well as other errors, and a control group received no feedback. Assessment involved a pre-test, an immediate post-test and a delayed post-test. While both feedback groups significantly outperformed the control group, no difference in performance was noted between the focused and unfocussed groups. It should be noted however, that both focused and unfocused groups strongly featured definitive articles, and may not have been sufficiently distinguished from each other as to show difference between the two types of CF. Additionally, this study limited itself to analysing improvements in one grammatical feature, it did not take into account improvements which may have taken place over a range of grammatical structures.

Building on the work by Ellis et al. (2008), Sheen et al. (2009) investigated whether direct focused CF, direct unfocused CF and writing practice alone produced any effect on writing accuracy. Using intermediate ESL classes, they formed four groups; a focused written CF group, an unfocussed written CF group, a writing practice group and a control group. The grammatical target for the focused CF group was the definite and indefinite articles, whereas the target for the unfocused CF group included articles and four other grammatical features. Students completed two written compositions over a period of two weeks, and their progress was analysed using pre-test, post-test and delayed post-test data. The focused CF group showed significantly greater improvement than both the unfocused group and the control group. However, while the

focused group outperformed the control group four weeks later in the post-test, the unfocused group did not. This suggested that focused CF is more effective than unfocused CF, which showed no more benefit than writing practice alone.

Bitchener and Knoch (2010-a) further investigated the effect of focused written CF on two functional uses of the English article system (using 'a' and 'the'). Using advanced L2 learners at a university in the USA, they formed a control group and three treatment groups: one which received a written descriptive explanation; one which received indirect circling of errors; and one which received written descriptive feedback and oral form-focused instruction. The subjects were asked to describe events occurring in a picture of different social settings in a pre-test, post-test and delayed post-test. They found that all three treatment groups outperformed the control group in the post-test. However, this improvement in accuracy was only retained in the delayed post-test by the two direct CF groups (receiving written meta-linguistic explanation and those receiving the explanation and oral form-focused feedback). The indirect group receiving only circling to indicate an error did not retain the gains made in the immediate post-test. These results seem to corroborate the findings of Sheen et al. (2009) as in both the direct focused CF treatment groups retained gains in the post-test whereas those groups which received indirect CF did not.

The case for direct focused written CF was furthered by Bitchener and Knoch (2010-b). They conducted a 10-month study of the effects of written CF on English articles using 52 low-intermediate ESL students in Auckland. Assigned to four groups, the students received either written CF: direct corrective feedback, written, and oral meta-linguistic explanation; direct corrective feedback and written meta-linguistic explanation; direct

corrective feedback only; and the control group did not receive CF. In a pre-test, post-test and three delayed post-tests structure, the students produced five pieces of writing. Each of the treatment groups outperformed the control group on all post-tests and no difference in effectiveness was found between the three treatment groups.

Like Sheen et al (2009) and Bitchener and Knoch (2010-a), Bitchener and Knoch (2010-b) also found evidence to support the use of written CF within a focused linguistic context. Yet, in all three of these studies, and Ellis et al (2008), articles either featured heavily in analysis or were the only linguistic feature analysed. Whether or not focused written CF is an effective pedagogy to improve the written accuracy of other linguistic features cannot be confirmed by these studies, and more research is still required. Therefore, it is important to investigate a potential differential effect between focused and unfocused written CF for multiple linguistic features.

2.7 Summary of literature review

Educational practices in Brunei Darussalam are going through a period of great pedagogical change at the moment as the government implements its SPN21 vision for education. Feedback is one of the strategies which the AfL policy SBAfL emphasises, but as has been mentioned above, not all feedback has the same effect on learning. While some types of feedback have been shown to be beneficial to student learning, others have been found to have a detrimental effect (Hattie, 1999). Yet thus far in Brunei the firmly entrenched English language teaching technique of providing written CF regularly to students in response to their writing has not been examined.

Writing is a process which lends itself well to the use of feedback due to both its learned nature and the readily available evidence of a student's understanding which it provides. Within Brunei written CF is the type of feedback most frequently provided to ESL secondary school students. However, despite its common place in the classroom empirical findings are mixed as to its benefits. Most researchers now agree that both form and meaning should be attended to in a L2 teaching programme, yet the value of written CF is still a contentious issue. Notably though, recent studies such as Sheen et al (2009), Bitchener and Knoch (2010-a) and Bitchener and Knoch (2010-b) would seem to suggest that a focused approach to written CF rather than an unfocused approach may actually produce differential effects on L2 written accuracy. This is an area which requires further investigation, particularly within contexts such as Brunei Darussalam where the practice of providing written CF is not only widespread, but in fact currently government mandated.

Chapter 3: Methodology

This chapter will describe the methodological context of this thesis in terms of research approach, research design and strategy used, as well as addressing issues of ethics and validity. It will then describe the data analysis techniques used.

3.1 Research approaches

All research is situated within paradigms which consist of different ontological and epistemological assumptions 'about the nature of social reality and the way by which we can come to know this reality' (Blakie, 2010, p.9). The 'nature of social reality' can be described as our conceptualization of the nature of knowledge. The 'way' can best be understood as the 'procedures and activities for selecting, collecting, organizing and analyzing data' (Blakie, 2010, p.8). Assumptions made about the conceptualization of knowledge need to be clarified in order to understand methodological choices in terms of methods for data collection and analysis made in this research.

Creswell and Miller (1997) classify researchers' understandings of knowledge into four groups: quantitative, qualitative, ideological and pragmatic. Qualitative and quantitative approaches are seen as opposing paradigms. For example, the quantitative approach views knowledge as being external to the individual, and not based on subjective interpretation. Particular attention is given to ensuring that individual biases do not affect the knowledge under examination. However, the qualitative approach views knowledge as attained when people sense their world and give meaning to their 'senses through socially constructed interactions and discussions' (Creswell and Miller, 1997, p.37). The ideological approach views knowledge through a particular ideological lens, such as

feminism or socialism for example. However, of particular interest for the current study is the conceptualisation of the pragmatic approach, combining both quantitative and qualitative approaches to consider knowledge as existing both externally, and internally. Both objective, and subjective knowledge are acknowledged.

As is outlined further below, this conceptualisation of knowledge, combining both quantitative and qualitative approaches influenced the selection of the particular research method used. While quantitative research emphasises the counts and numbers of things (Berg, 2009), qualitative research refers to methods designed to investigate the 'meanings, concepts, definitions, characteristics, metaphors, symbols, and descriptions of things' (Berg, 2009, p.3). Supporters of the quantitative approach argue that it is the most objective, and therefore most accurate form of research because 'measurement enables us to transcend our subjectivity (Bradley and Shaefer, 1998, p108). However, supporters of the qualitative approach argue that its methods including observation, interview and in-depth investigations are 'more faithful to the social world' (Gergan and Gergan, 2000, p.1027) and allow data to emerge more freely from its natural context (Symonds and Gorards, 2010).

While previously thought to be too different epistemologically, and therefore incompatible (Symonds and Gorards, 2010), a mixed-methods approach utilising both qualitative and quantitative research paradigms has been identified as a third research paradigm (e.g. Howe, 1988; Teddlie and Tashakkori, 2003). The goal of a mixed-methods approach is to draw from the strengths of both the quantitative and qualitative approaches, and allow for researchers to 'mix and match design components that offer the best chance of answering their specific research questions ' (Johnson and

Onwuegbuzie, 2004, p15). With the combination of both objective and subjective knowledge sought in this study, a mixed-methods approach combining both quantitative and qualitative methods is the most suitable methodological choice. This means that a wider range of knowledge is able to be investigated than would have been the case if the research had been limited to either a quantitative or qualitative approach alone.

Thus, it is important that the particular ideological approach used in a study both informs and complements the manner in which knowledge is conceptualised, and the nature of the knowledge sought. The next section discusses the approach adopted for this research and explains why the paradigm adopted is the most suited to the research questions posed in this study.

3.2 My research approach

The current study conceptualizes knowledge in a manner best described as fitting Creswell and Miller's pragmatic approach (1997). With regards to the first research question '*is there a differential effect on L2 written accuracy following focused or unfocused written CF?*', it is possible that a differential effect of CF exists with regards to whether written CF impacts accuracy which holds objective truth for all L2 students. In this manner, the knowledge sought regarding the effects of written CF pedagogical practices is viewed as being objective and best investigated by using a quantitative approach.

While being a smaller part of this thesis, it is also important to consider how the knowledge sought may be culturally constructed. To this end, a qualitative research approach best addresses the second research question '*what are the students*'

perceptions of written CF?. Very little written CF research has been conducted outside of a Western context, and indeed none has been conducted within Brunei Darussalam. Hofstede and Hofstede (2005) theorised that 'cultural software' serves as a filter with regards to how individuals process information. It may be that Bruneian Islamic Malay students do not process CF in the same manner as those students from different cultures.

The dominant method which framed the research was the quantitative quasi-experiment (as defined below), with a questionnaire that included qualitative data providing additional information. The quasi-experimental design allowed for a direct comparison between focused and unfocused written CF in order to address the first research question 'is there a differential effect on L2 writing accuracy following focused or unfocused CF'. Experimental design is frequently claimed to be the most rigorous research method available to investigate causal relationships due to its high internal validity (Trochim, 2006). Internal validity is the extent to which the experimental design controls extraneous variables (Borg, Gall and Gall, 1993). In the case of the current study, the causal relationship under investigation is the relationship between written CF, and its effect on student written accuracy. An important aspect of controlling these variables in true experiments requires random selection of subjects and random assignment to treatment groups, thereby removing any systematic error that may be associated with using intact groups (Ross and Morrison, 1996). For this research random selection would have been unrealistic because it would have involved negotiating access to all secondary schools in Brunei and disrupting teaching programmes.

Thus, due to the natural school setting in which the research was conducted, random selection of study participants was not possible. Hence, the environmental context of the study precluded the use of a true-experimental design. An alternative to the true-experiment is the quasi-experiment. Like experiments, quasi-experiments are studies used to estimate the causal impact of an intervention. However, the major difference between the two is that quasi-experiments lack elements such as the random assignment of participants (Heppner, Kivlingham and Wampold.,1992). It is important to note however, that quasi-experiments 'should not be seen...as always inferior to true experiments...(they are sometimes) the next logical step in a long research process' (Fife-Schaw, 2006, 92, in Lee, 2011). This was certainly the case in this research, where the real life school context meant that a quasi-experiment was a practical research method.

It was not possible to use a control group in the current study due to the cultural context. Within Brunei Darussalam, the Ministry of Education legislate that all secondary school English language students must have their written work corrected by teachers at least once a fortnight. This means that it was not possible to withhold corrective feedback from any students for the length of the study and in any case it would have been unethical to do so. Therefore, investigating whether L2 written accuracy improved without the corrective feedback interventions is unfortunately beyond the scope of possibility for this study. The limitation here is that it will not be possible to investigate whether students who did not receive feedback would improve their written accuracy by writing practice alone. However, the lack of control group is acceptable for this smaller scale research which seeks to investigate any differential effects between focused and

unfocused CF, rather than comparing CF to no CF. Therefore, considering the lack of possibility of a control group, a quasi-experiment was a logical choice for the current study.

Another benefit of using a quasi-experimental design was that it was possible to build upon previous quasi-experiments conducted by Sheen (2007) and Sheen et al (2009). In this manner replication of the previous studies within different educational contexts could lend support for the concept of an external, generalisable principle (Lee, 2011).

The second research question was investigated through a questionnaire which included questions generating both qualitative and quantitative student responses. This type of questionnaire aims to determine the diversity of a particular topic, in this case student perceptions regarding CF. Qualitative questionnaire analysis is a useful tool through which to explore subjective meanings and experiences (Fink, 2003). Thus, it was a valuable tool within this research given that information was sought regarding students' subjective perceptions concerning CF.

This section has described why a mixed-method approach was the best methodological choice for this research. The discussion now turns to more specific details regarding the research design.

3.3 Sampling and participants

Statistically, selecting subjects using random probability sampling across the entire secondary school student population in Brunei may have helped to permit justifiable inference from a sample to a general population, and guard against bias. However, this was not possible in this context. Logistical obstacles related to funding, access and time limitations precluded the use of probability sampling in this research. As a foreigner in this educational context it is difficult to gain access to schools, and I would not have been released from my teaching duties to conduct research in other schools. However, I was fortunate that I work in Brunei and thus was allowed to conduct research within the school where I worked. In situations such as this where random probability sampling is not possible, the next best available option involves the minimization of the influences of selective forces as much as is possible within the smaller sample (Kane, 2002).

Non-probability convenience sampling which did not use random selection from the general population, and was based on available subjects was used in this study. Convenience samples rely on readily available subjects and are frequently used in educational research (Berg, 2009). In this case, the available subjects consisted of classes of year nine students, and a computer-generated random number list allocated subjects to either the focused or unfocused group.

Guidance given by the school principal suggested that it would be the least disruptive to use year nine students for this research as they were in the first year of a three year Cambridge 'O' Level course, and therefore were not required to sit any external examinations. As I am the teacher of three year nine classes, a decision was made in

consultation with the principal that my classes would be invited to take part in the research. The three classes were all from different ability levels (a high ability science stream, a mid-ability class, and a lower ability class) which allowed for a greater representation of students from the student population. In this manner, three clusters within the school were established.

Although initially 64 students from three classes elected to participate in the study, unforeseen circumstances where the middle ability class was required to perform in a National choir for His Majesty the Sultan of Brunei's birthday celebration meant that this class could not participate in the study. The remaining two classes both consisted of 19 students, so the entire sample size was 38. As the class which was forced to drop-out of the study had been randomly assigned to both the focused and unfocused group, their departure did not disrupt the groupings of the other students. Smaller sample sizes such as this mean that statistical inferences cannot be made to the general population. However, they can still form part of a valid and defensible methodology as they offer information about a particular section of a population which may also be relevant to other similar populations. Moreover, by combining the quantitative data with the qualitative snapshot of students' perceptions of written CF provided by the use of a survey, a wider breadth of information was gathered which gave an insight to the role of written CF within this particular cultural context.

The 36 participants were all female Bruneian Malay students aged 12-14. All had formally been learning English as an additional language at school for eight years. The students received three classes of one hour duration of English instruction each week. During the course of this research, all students received the same English instruction

from the same teacher, with an emphasis on oral language and communication skills.

During the two week time period between the post-test and delayed post-test, students were sitting their end of year school based examinations so no instruction was given.

While random selection was not possible from the entire student population for reasons mentioned above, random assignment of these participants into one of the two research groups was carried out in an attempt to control potential selective biases. Random assignment was achieved by computerized generation of random student numbers and assignment to either the focused or unfocused treatment group was based on those numbers. This meant that as far as possible, the group attributes for both treatment groups were roughly equivalent, and therefore any differential effect observed between groups could more likely be linked to the treatment, and not a characteristic of the individuals in a group.

3.4 Research Design

The dominant framework of this research design consisted of a 13-week design for the quasi-experiment with a questionnaire in week 10 of the study. The quasi-experiment utilised a pre-test, post-test, delayed post-test structure. Students received written CF for eight weekly writing tasks between the pre-test and the post-test. Table 1 shows the order of the quasi-experimental tasks and the questionnaire. Students were either in a focused written CF group which only received feedback on one particular type of linguistic feature (one of the copula 'be', articles, regular past tense, irregular past tense or prepositions), or an unfocused written CF group which received feedback on all five

linguistic features. Further details on the design of the weekly session and the pre, post and delayed tests, and the five linguistic features selected are provided in section 3.5.

Table 1

Design of the study.

	Focused CF <i>n</i> =19 Unfocused CF <i>n</i> =19
Week 1	Pre-test: Written narrative
Week 2	Written Task 1 Written CF on Task 1 Revised written Task 1
Week 3	Written Task 2 Written CF on Task 2 Revised written Task 2
Week 4	Written Task 3 Written CF on Task 3 Revised written Task 3
Week 5	Written Task 4 Written CF on Task 4 Revised written Task 4
Week 6	Written Task 5 Written CF on Task 5 Revised written Task 5
Week 7	Written Task 6 Written CF on Task 6 Revised written Task 6
Week 8	Written Task 7 Written CF on Task 7 Revised written Task 7
Week 9	Written Task 8 Written CF on Task 8 Revised written Task 8
Week 10	Post-test: Written narrative Questionnaire
Week 13	Delayed post-test: Written narrative

3.5 Methods

This section describes the structure, instruments and procedures used for the quasi-experiment and questionnaire.

3.5.1 Pre-test, post-test and delayed post-tests

A non-equivalent groups design was used, with a pre-test (week 1), post-test (week 10) and delayed post-test (week 13) structure. Non-equivalent group designs are structured like a pre-test, post-test experiment but lack the key feature of random assignment from the general population as a whole. Instead two groups are used which are thought to be comparable. In this case, as outlined in section 3.3, the subjects were students from two year nine classes at a Bruneian girls' secondary school who were randomly assigned to either a group receiving focused written CF, or unfocused written CF.

I designed these three tests, each of which consisted of a series of six pictures and key words. Three tests were used in order to investigate any possible differential effects following either written focused or unfocused CF over an extended period of time. The pre-test enabled base-line data to be collected, while the post-test allowed for a comparison of written accuracy scores immediately following the eight weeks of written CF. A delayed post-test was used in order to examine any delayed effects of the written CF. Each of these tests used a theme related to a common event in the students' lives (Appendix 1). The pre-test showed pictures of a student getting ready for the first day of school, the post-test showed pictures of an injury and a doctor's office, and the delayed post-test included pictures of going on a vacation. These themes, which were familiar to students, were selected in order to increase construct validity. As the classroom teacher

of the participants previously, I had observed these themes being repeated in student textbooks. Therefore, if an individual student had no personal experience of travelling away for a vacation, they did have experience of reading and writing about vacations. Each test was structured and organised in the same way, and the tasks were the same. This meant that I was able to establish some level of comparability across the three tests.

Pictures were used as story guides with only a few key words so that all students regardless of their individual English language ability could access and understand the theme. The six pictures were collected from a clip art site and were selected for relevance to the theme of each test. Students were also told that they were to use the picture stories as a guide, and that they could write an imaginative narrative. This enabled all students to write a short composition even if they lacked the vocabulary knowledge to write about the specific event depicted in the picture story. For example, if students did not know the vocabulary associated with playing basketball as is depicted in the post-test, they were still able to write a narrative about the same series of events for a sport which they were more familiar with. The tests were not designed to assess students' particular thematic vocabulary. Rather, the objective of the tests was to generate narrative writing in the past tense which enabled an examination of how the students used the five targeted grammatical features (copula 'be, regular past tense, irregular past tense, articles and prepositions). This allowed for a degree of comparability between the three tests because students were able to write a narrative composition of about half a page even if they did not have the vocabulary knowledge related to a particular theme of a test. The format, instructions and conditions of each

test were the same, and the desired outcome of each was the production of a past tense narrative composition. The use of regular and irregular past tense forms is an area which I have observed Bruneian students having difficulty with previously, and it is also an area which Sheen et al (2009) investigated with regards to the effect of written CF. In order to build on knowledge from that study, and investigate the effect of written CF on particular linguistic features which Bruneian students have difficulty with, the writing tasks used in the current study were designed to generate writing in the past tense. Moreover, students are used to similar activities in tests, so the process of writing a past tense narrative story based on a series of pictures is familiar to them and did not require teaching.

For each of the three tests the students were given the pictures, and were encouraged to discuss them with each other in pairs for five minutes. They were also asked by the researcher if they had any questions about the pictures, and if required clarification was given. A few students used this time to ask for a particular vocabulary word, but all were familiar with the process of writing a narrative based on pictures from school examinations and the task itself did not require further clarification. As the tests were not concerned with students' vocabulary knowledge, if a student asked for a particular word prior to the test they were provided with it. Students were then told they had 30 minutes to write their own narrative story of approximately two paragraphs.

3.5.2 Weekly written CF sessions

After the pre-test, and before the post-test, there were eight separate written CF sessions, e.g. weeks 2-9. During each of these sessions the students produced writing

which was then corrected using either focused or unfocused feedback. As in the pre, post and delayed post-tests, the writing tasks were past tense narratives. Students received written CF on their writing for these eight sessions, and any change in the written accuracy was then examined using the writing generated in the post and delayed post-tests. For the written CF writing tasks the students were given a story sheet (based on the mousedeer stories as described below) and were asked to read the story and discuss it in pairs. This writing task is of the type same as that used by Sheen et al (2009), the difference being that the stories used in the current study were from a Southeast Asian context rather than Western fables. It is also type of writing task which students are familiar with in class within this particular educational context. As these were stories that the students were already familiar with, these student discussions mainly concerned clarifying any unfamiliar words in English from the text. Once the students told the researcher that they understood the story, the story sheet was taken away from them, and they were instructed that they had 30 minutes to re-write the story in their own words.

The narrative tasks were based on the common South-East Asian tales of Mousedeer. These popular stories demonstrate the craftiness and intelligence of the mousedeer. In all of the stories, Mousedeer encounters a situation which he must use his craftiness to get out of. For example, in one story Mousedeer falls into a deep hole and cannot get out. To escape, he tricks an Elephant into getting into the hole to help him, and then Mousedeer climbs on the elephant's back to climb out of the hole, leaving the Elephant stuck in the hole. Mousedeer stories have been told and retold as folklore for generations, and are commonly known among the people of Brunei, Malaysia and

Indonesia (Carpenter, 1992), much as stories such as Cinderella are in New Zealand. Mousedeer was selected as the subject of the stories so that all students, regardless of their English language ability, could understand the story which they were to retell. As the researcher, I asked local Bruneians to tell me the versions of Mousedeer stories they knew, and also collected Mousedeer stories from the internet. Then I summarized the stories in my own words, in English, ensuring that they were not longer than half a page (Appendix 2).

Students from both experimental groups retold the same eight narratives in eight separate lessons (a different story each week), and received written corrective feedback from the researcher during the subsequent lesson. One group received focused corrective feedback for all of the eight writing sessions, while the other group received unfocused corrective feedback. During this lesson, students then re-wrote their stories taking into account the CF provided. All students were told on each occasion that the writing task would not affect their grade for the year. However, the students were told that at the end of the research project they would be given the results of all three tests, and they took the task seriously.

The procedure for each of the eight written CF sessions was as follows;

1. First, the teacher distributed the story sheet which was of no more than half a page in length, to the students and instructed them to read it silently.
2. Then the teacher read the story aloud in English to the class.
3. The teacher explained key words and discussed the story with the class.

4. The teacher collected the stories from the students and gave out blank writing paper.
5. The students were instructed to retell the story in as much detail as they could, in written form in English.
6. Written task 1. The teacher collected in the students' written stories.
7. Written CF on task. The teacher provided written CF to the stories, and returned it to the students during the next lesson (as researcher/teacher I saw these students for three lessons a week). No further feedback on the writing was given.
8. Revised written task. Students were given time to read the written CF, then instructed to re-write their story taking care to correct the errors indicated.

All eight steps for each story occurred within the same week. These stories were replicated for each of the eight weeks using eight different mousedeer stories.

3.5.3 Written corrective feedback protocol

Students were given one of two types of written CF for their narrative writing: focused or unfocused.

The protocol for giving focused CF consisted of reading each student's story and making note on a separate tally sheet (Appendix 3) of the number of grammatical errors for each of the targeted features (articles, copula 'be', regular past tense, irregular past tense and prepositions) .The targeted feature with the most errors became the linguistic feature I corrected on the student's actual writing. If two or more of the linguistic features had the same number of errors, I selected the first of these which occurred within the writing to correct. Then each of the identified errors in the student's writing

was underlined, and the correct form written above the incorrect word. For example, if a student wrote '*Then Sang Kancil run away*', and irregular past tense verbs were the grammatical feature selected for feedback, the word '*run*' would be underlined, and the correct form '*ran*' written above it.

The five targeted errors were selected in order to build upon previous research conducted by Sheen et al. (2009), which investigated the effect different types of feedback had on these particular grammatical features. Within their research, they explain that the features were selected due to a prediction that intermediate learners make errors in these features, and due to their easily identifiable nature within the written narrative context. As an English language teacher within Brunei Darussalam, I have also frequently noticed students making mistakes with these particular features. Therefore, I decided to use the same features in this research because these are features students frequently make errors with, and also because it enables a comparison of results between Sheen et al (2009) and the current study.

The protocol for giving unfocused CF consisted of reading the student's story and correcting all of the five targeted grammatical errors. Each error was underlined, and the correct form written above the incorrect word. For example, in the sentence '*Then Sang Kancil run to garden*' the irregular past tense error '*run*' would be underlined and '*ran*' written above it. The space between '*to*' and '*garden*' would also be underlined and the article '*the*' would be written above the space.

3.5.4 The questionnaire

A questionnaire was given to each student in week 10 of the study. Each student completed the same questionnaire, the only difference being an identifying dot on the questionnaires issued to the focused group to enable analysis by group ensuring anonymity. The questionnaire was administered in the classroom to all the students at the same time. This meant that the participants were all issued with the questionnaire while a teaching peer helping to administer the research was in the room. Students were encouraged to ask for clarification about the questionnaire if required. This was very important within an ESL context in order to ensure that students understood the questions asked of them in English, and it also ensured a high response rate. With ease of communication in mind, the questionnaire was also written keeping the language within it as simple as possible.

The questionnaire needed to allow for a variety of responses in an attempt to capture students' subjective perceptions of corrective feedback. Therefore, a short questionnaire using one rating scale question, and one open-ended question was a practical way to collect data within this particular cultural context. A rating scale was used in order to collect data in a structured way, which could allow for across group comparison. The open-ended question allowed students to provide a variety of subjective perceptions which would not have been possible with closed questions.

Schuman and Presser (1981) compared survey responses from closed and open type questions, and found considerable difference in the information they generated, despite

each question being about the same topic. They argued that this was caused by the use of unsuitable response categories provided in the survey. This was particularly salient within this study due to the different cultural backgrounds between myself as researcher, and the subjects. Had I provided a list of fixed responses for the students to select from in a question, or written closed questions based on predicted responses, I may have prevented true identification of student beliefs. Thus, an open-ended question was suitable for this study

In order to prevent the students from feeling pressure to respond to the survey in any particular way, a peer administered the survey. He instructed the students not to write their name on their questionnaire, fold it in half when it was complete and place it in a box. The students were made aware from the outset that their responses would remain anonymous.

The survey consisted of the following questions:

- How helpful was teacher feedback on your writing in improving your writing?

Not at all helpful

Quite helpful

Very helpful

1

2

3

4

5

6

- What kinds of teacher feedback do you think helps your English writing the most?
Why?

These two questions were used because they revealed information about student perception of the CF given during the research.

3.6 Ethics

The Massey University process for ethical approval was followed for this research, and ethical approval gained. Written permission was obtained from the Director of Education for Brunei, and the school principal to carry out this research within the school.

Information sheets detailing the aims and methodology of this study, and participants' rights written in both English and Malay were given to the students and their parents by a non-teaching peer in order to avoid students feeling pressured to participate. Written explicit consent was obtained from the students and their parents.

Informed consent was used and participants were informed that they could withdraw from the investigation at any stage. This is an important ethical consideration due to the 'emergent and unpredictable' nature of research in a real life context (Houghton, Casey, Shaw and Murphy, 2010, p.15). While I used one-off informed consent forms at the outset of the research, I could not guarantee that parents' and students' opinions of the project would not change. Moreover, as individuals have different perspectives about what makes up ethical behavior, Holloway and Wheeler (2002) argue that constant negotiation with participants is crucial. Consequently, participants had the autonomy to withdraw from the study at any time. However, no students besides from the class which were required to participate in a National level celebration withdrew from this study.

Another potential ethical issue considered was that of confidentiality. Pseudonyms and a selective description of details were used in order to protect the identities of the

subjects. If too many details are exposed, subjects risk exposure and embarrassment (Stake, 2000). Both the school and students were assured that every possible measure would be taken to maintain their confidentiality. Similarly, confidentiality of questionnaire responses was used to ensure that students didn't feel pressured to respond in a certain manner.

3.7 Role as the researcher

It was important that I considered my dual role of both researcher and teacher within this study. During the course of the research I was employed as an English teacher within the school where the study was conducted, and I taught English to the students who participated in this research. In order to both protect the students and ensure validity of the results, several safeguards were put in place. Firstly, a colleague explained the research to the students, taking particular care to inform them that participation was voluntary, and he provided them with information sheets at a time when I was absent from the class. He also collected the consent forms from the students so that they did not feel pressured to participate out of loyalty to me. The same colleague also administered the questionnaire in week 10 of the study which was kept anonymous. An identifying dot on the questionnaire distributed to students in the focused group allowed for group analysis of the results, whilst keeping students' individual identities anonymous.

I also ensured in the design of the study that when I analysed student writing I would be counting tangible grammatical features and errors, rather than using subjective

judgement which may have been unconsciously influenced by my relationships with the students.

3.8 Triangulation of data

The mixed-methods approach utilised in this research allowed for triangulation of data. Triangulation refers to use of multiple data-collection techniques designed to measure a single construct (Berg, 2009). Triangulation can lead to mutual confirmation of measures and validation of findings (Jick, 1983; Leedy, 2001; Mitchel, 1986). Some theorists (e.g. Creswell and Clark, 2007, Teddlie and Tashakkori, 2003) argue that using more than one method 'produced stronger inferences, answered research questions that other methodologies could not, and allowed for greater diversity of findings' (Denzin, 2010, p. 424).

Flick (2002) argues however, that triangulation can also be 'less a strategy for validating results...than an alternative to validation which increases scope, depth and consistency' (p.227). This is certainly the case within this research, as the use of a student perception questionnaire adds to the scope and depth of knowledge found within the quasi-experiment. These two research methods used together illuminated aspects of the research questions which one method alone would not have been able to. In particular, findings regarding the effects of both focused and unfocused CF are able to be examined in the light of student perceptions within this particular cultural setting.

3.9 Validity

3.9.1 Internal validity

Internal validity refers to the extent to which causal connection observations and interpretations drawn from data hold true and the extent to which variables which compromise the integrity of the research are minimised (Bryman, 2004). Cook and Campbell (1979) identified threats which exist when conducting one's own research. The threats which are salient to this research are discussed below.

The first threat to internal validity is outside influences (Cook and Campbell, 1979). It is possible that events other than those within the experimental condition contributed to the outcome of the study. This may be a greater threat within ESL research such as this as it is possible that students may review what they learn in the experimental condition outside the research context (Lee, 2011). For this reason, both the focused and unfocused groups were made up of students from different classes. This meant for example, that if one class was exposed to a particular learning situation in another English medium subject, the effect would be distributed between both the focused and the unfocused CF groups. Furthermore, as the English teacher of all of the students within the study, I followed the same lesson plan and used the same resources for both classes during the course of the research. Similarly, all students were issued with the questionnaire at the same time to prevent any discussion between them which may have influenced their responses. This involved timetabling a lesson with both classes at the same time so that all the students could complete the questionnaire at the same time.

Participant selection is also an important consideration in order to establish internal validity. Despite random assignment to groups from the classes used, the lack of random selection of subjects from the school population as a whole, means that it is possible for systemic bias to cause one group to be superior in some way to other groups. Therefore, a pre-test was used for both groups, and analysed in order to ascertain comparability between the groups. Pre-tests are often used in ELT research in order to statistically adjust the post-test scores between groups if any variance is found (Dornyei, 2007). However, in this case, Levene's Test (1960) revealed no variance between the two groups.

Cook and Campbell (1979) also claim that it is possible that familiarity to testing methods can lead to an improvement in test scores. In fact, Ellis (1994) argued that repeated measures may affect test scores in an unpredictable fashion. For this reason, a testing method was selected for the pre-test, post-test and delayed post-test which students have been familiar with for years. Writing stories based on sequential pictures has been used in national standardised tests at least since the students were in year 6, and in every year subsequently. In my own observations of teaching practice at the school in which the research was conducted, I have noted it used on many occasion, particularly in lower school examinations. Rather than test scores improving over the course of the tests due to familiarity with the test format, students at the outset in the pre-test were already very familiar with the format of the tests.

Mortality (participant dropout during the experiment) is another factor identified by Cook and Campbell (1979) which can affect internal validity. During this research it was necessary for one of the three classes involved to dropout due to National level

celebrations. This meant that the class involved did not have school lessons for one month as they were involved in choir practices. However, as both the focused and unfocused treatment groups' participants were randomly assigned from each of the classes, this withdrawal did not affect the proportional make-up of the CF groups. This meant that it was possible and practical to continue with the research using two classes instead of three. No other students withdrew during the research.

Thye (2007) suggest that the effects mentioned above (history, selection, test familiarity and mortality) would influence both experimental groups evenly. Added to this, as outlined above measures were taken to minimise any potential threats. Thus, any threat to internal validity should not affect the outcome of the research.

3.9.2 External validity

In order for research to establish external validity, it must be generalisable and replicable. The generalisability of any research findings involves making at least two inferences (Cornfield and Tukey, 1956). Firstly, it involves a statistical inference from an observed sample to the population from which it is drawn. Secondly, it involves subject-matter inferences from conclusions about the sampled population to the conclusions of interest for a wider population. It may be possible to draw an inference that findings of this study also hold for those like those observed i.e. other year nine students at this particular secondary school in Brunei. Generalizing findings beyond this small group requires that the 'universe be defined clearly enough that one can recognize elements belonging to it ' (Cronbach, Gleser and Rajaratnam, 1972). To this end, it is possible that the findings are representative for year nine students at other all girls' state schools

in Brunei Darussalam because cultural and contextual elements from this particular sample of students are also common to those students. Moreover, the results from this research are highly relevant to English language programmes for Muslim girls in neighbouring countries in South-east Asia such as Malaysia and Indonesia. This means that the findings of this research are able to add a different cultural dynamic to the findings of similar research such as Bitchener (2008), Ellis et al 2008 and Sheen et al (2009).

3.10 Data Analysis

3.10.1 The data collected

The pre-test, post-test and delayed post-tests for both the focused CF group and the unfocused CF group were scored for the accurate use of the five grammatical forms in the same manner to that of Sheen et al. (2009). The total number of correct and incorrect occurrences of the five target linguistic features were counted, as well as the total number of these forms. The latter was divided by the former to create an accuracy score for each student as also done by Sheen et al. (2009). Data was also collected in the questionnaire that students' completed during week 10 of the study in the form of rankings from one to six for the first question of the survey, and open-ended responses to the second question regarding preferred types of CF.

3.10.2 Establishing the nature of the test data

In order to establish the best way to analyse the data, it was first important to check the nature of the data as several statistical assumptions must be met before particular

statistical tests can be validly applied. Thus, it was vital to determine if these assumptions were valid.

Convenience sampling can lead to under-representation or over-representation of particular groups in a sample which can lead to biased results (Ross, 2005). For this reason, the accuracy scores of both the focused and unfocused corrective feedback groups in the pre-test phase were compared in order to ascertain parity prior to the intervention. The first assumption tested for was that of homogeneity of variance. This assumption means that the variances should be the same throughout the data. In other words, the variance of the outcome should be the same at all levels of all other variables. Analysis of variance (ANOVA) assumes that variances are equal across groups. Conover, Johnson, and Johnson (1981) verified empirically that Levene's test (1960) is one of the best tests for establishing homogeneity of variance with regards to power and robustness. Therefore the Levene's test was used to test each variable using an alpha level of 0.05. Homogeneity of variance was verified.

The next statistical assumption which was examined was that of normal distribution. If data is normally distributed, then it is evenly distributed around the mean in a very regular way. The rationale behind any hypothesis testing relies on having data that is normally distributed, so if this assumption is not met then the hypothesis testing is flawed (Field, 2009). Many statistical tests assume that the sampling distribution is normally distributed. The Kolmogorov-Smirnov test was used to compare the scores in the sample to a normally distributed set of scores with the same mean and standard deviation. These tests revealed that much of the raw data was not normally distributed which means that rather than being distributed around the mean in a predictable

manner, the data was skewed in one direction or another. In particular, the data revealed that in each group, there were about 5 or 6 students who were more proficient at particular linguistic features than the other students and this resulted in the data frequently being skewed to the right (Appendix 4). This data was a valid representation of the groups which were made up of students of differing English language proficiency, but it was important to analyse the data in a method which took into account the non-normal distribution.

There are several robust non-parametric tests which do not rely on the assumption of normal distribution (Field, 2009) such as Friedman's ANOVA and Wilcoxon signed rank tests. Given the nature of the data in this study, these tests were used in order to ensure the robustness of the analysis.

Having data which was not normally distributed meant that analysing the raw data using many of the standard statistical tests such as t-tests or ANOVAs are invalid without first transforming the data. However, whether or not the results of tests on transformed data are valid is a contentious issue. For example, in an extensive review Glass, Peckham and Sanders (1972) commented that 'the payoff of normalizing transformations in terms of more valid probability statements is low, and they are seldom considered to be worth the effort' (p.241). Thus, for the data collected in this study non-parametric tests which do not rely on normal distribution were the best choice.

3.10.3 Analysing the test data

Following the completion of the delayed post-test, two Friedman's ANOVA tests (Friedman, 1937) were conducted examining changes over time with respect to the

treatment intervention. Friedman's test is a non-parametric alternative to the repeated-measures ANOVA, and is used for testing differences between conditions when there are more than two conditions and the same participants have been used in all conditions (Field, 2009). A repeated measures design was used in order to investigate the changes in ranked median scores between the two groups over three points of time; during the pre-test, the post-test and the delayed post-test. In the case of this research, the null hypothesis investigated was that there would be no significant difference in the written accuracy score between the focused and unfocused feedback groups in each of the three measurement phases was tested.

Once the Friedman tests of the overall mean accuracy scores for both of the groups in each of the tests were completed, it was important to break the accuracy scores down into scores for each of the five linguistic features examined. This was in order to investigate if either CF method was more effective with regards to any particular linguistic feature. It was important to break the data down into the linguistic features so as to build on previous knowledge. Sheen et al (2009) found that students receiving focused CF demonstrated a significant improvement in their use of articles compared to an unfocused CF group. Thus, by analysing each linguistic feature separately it was possible to verify if the students in the focused CF group also showed an improvement in articles.

In cases where the overall effect from Friedman's ANOVA was not significant, no further analysis was necessary. However, for significant effects post-hoc testing was carried out using the non-parametric Wilcoxon signed-rank tests (Wilcoxon, 1945). This test is

used in situations where there are two sets of scores to compare from non-normally distributed data, and the scores come from the same participants.

3.10.4 Analysing the student questionnaire

The mean student rating from the first question of the exit questionnaire was also compared between groups, and tested for significance using a Mann-Whitney U test. This test is suitable for testing differences in non-normally distributed data between pairs from two conditions, when different participants have been used in each condition (Field, 2009).

The other question of the questionnaire was analysed using a constant comparative approach (Merriam, 2009). This approach involved inductively creating categories for coding simultaneously with comparing student answers from the questionnaire. In order to allow for a variety of student responses, expected responses were not identified at the outset. Rather, categories were created to capture student responses during the analysis stage. This allowed for student responses which may have been unanticipated. The student responses were read, reduced to type of feedback categories, and then analysed for thematic responses regarding their perceptions of the usefulness of the error correction, and the type of learning feedback students believed helped their writing improve the most (Hsieh and Shannon, 2005).

3.11 Summary of methodology

The current study conceptualises knowledge pragmatically (Creswell and Miller, 1997), as being both objective and best investigated using quantitative methods, and subjective requiring qualitative methods. This conceptualisation led to the selection of a

mixed-methods design consisting of a quasi-experiment with a questionnaire in its final stages. These were the most suitable data collection methods due to both the nature of the knowledge sought, and also the natural classroom setting of the educational context. Taking place in a Bruneian secondary school, the study consisted of 38 year nine female students who were randomly assigned to either a group receiving focused written CF or unfocused written CF. In order to ensure the safety of these participants, the Massey University process for ethics approval was followed, and my dual role as both researcher and teacher was carefully considered. The upmost care was also taken to ensure both the internal and external validity of the study. Data was collected by using a pre-test, post-test and delayed post-test procedure, and also from an anonymous student questionnaire administered after the post-test. This data was analysed to establish if it met the statistical assumptions of homogeneity and normal distribution. Once it was verified that the data was not normally distributed, the pre-test, post-test and delayed post-test data was analysed using the robust non-parametric statistical Friedman's and Wilcoxon signed-rank tests, while the student questionnaire was analysed using a Mann-Whitney U test and a constant comparative approach.

Chapter 4: Results

This section provides information about the data collected and then descriptive statistics for the overall written accuracy scores of both the focused and unfocused CF groups, and the accuracy scores for each of the five targeted linguistic features for each group. Following that the questionnaire results regarding the students' perceptions of CF will be interpreted and reported.

4.1 Pre-test, post-test and delayed post-test

4.1.1 Calculating writing accuracy scores

In order to answer the first research question '*Is there a differential effect on L2 written accuracy following focused or unfocused written CF?*' it was necessary to calculate writing accuracy scores for each student. This was done in order to compare any changes in accuracy from pre-test to post-test to delayed post-test and across the two groups – focused and unfocused. Using the student writing from the pre, post and delayed post-tests, an accuracy score for both the focused and unfocused groups was calculated for each of the five targeted linguistic features (articles, copula 'be', irregular past tense, regular past tense and prepositions). As mentioned in section 3.11.1, the total number of correct occurrences of each of these linguistic features was counted in the student writing, and then divided by the total number of times the linguistic feature was used in the writing task in order to calculate an accuracy score for each of the linguistic features for each student. The overall mean for the linguistic features combined was calculated in order to ascertain an overall accuracy score. These are presented in Table 3.

Having counted the raw data and created overall and individual linguistic feature accuracy scores for each student, the nature of the data was then tested to check if it met the statistical assumptions upon which the validity of many statistical tests rest. This was done in order to select the most suitable data analysis technique. The next two sections outline the testing conducted to check for the two key assumptions: the assumption of homogeneity of variance; and the assumption of normal distribution.

4.1.2 Testing for homogeneity of variance

Using the writing accuracy scores from the pre-test, homogeneity of variance was tested using Levene's test (1960) to ensure that the data from the focused CF group and the unfocused CF group were comparable. Standard statistical tests which analyse variance between data rely on the assumption that the different samples under comparison have similar variance, or in other words, the spread of dispersion of values must be about the same for each of the two groups. For each of the five linguistic features, the variances were compared using an alpha of 0.05. This allowed for a confidence level of 95% that any differences found were statistically significant, and could therefore be viewed as non-comparable groups. It is also the level at which an alpha is most commonly set, particularly in studies with small sample sizes (Noordzij, Tripepi, Dekker, Zoccali, Tranck and Jager, 2010). For this reason, this is the critical level used throughout this study for all statistical analysis. The significance column of Table 2 below displays the level of significance for each linguistic feature, and as they are all above the alpha of 0.05, homogeneity of variance between the focused and unfocused groups for each of the linguistic features could be established. This means that the spread of dispersion of the data for both groups is about the same showing that

that the groups are comparable. This is important as without this homogeneity of variance, we could not be confident that the observed mean differences were attributable to the effects of the treatment (i.e. focused or unfocused CF).

Table 2.

Levene test of homogeneity of variance

Linguistic Feature	<u>Levene Statistic</u>	<u>Sig.</u>
Articles	3.446	0.72
Copula 'be'	4.111	0.50
Regular past tense	0.154	0.697
Irregular past tense	0.060	0.808
Prepositions	0.030	0.862

4.1.3 Testing for normal distribution of data

Once homogeneity of variance was established, the accuracy scores for each of the linguistic features from the pre-test, post-test and delayed post-test for both the focused and unfocused groups were tested using the Kolmogorov-Smirnov test in order to check the statistical assumption of normal distribution. If data is normally distributed, it is evenly distributed around the mean in a bell-shaped curve. Violation of this assumption can lead to invalid statistics if parametric testing is used. A basic assumption of parametric testing is that the data being analysed is normally distributed. If this is not true, then it means that these tests would not accurately describe the data. Therefore, it was important to check this assumption in order to ascertain if subsequent analysis should use parametric testing, or non-parametric testing which does not assume normal distribution.

For reasons outlined in section 4.1.2, this test used an alpha of 0.05 as the level above which the data could be assumed to be normally distributed. Appendix 4 displays the individual linguistic scores for both the focused and unfocused group which was not normally distributed. It is possible to see in these graphs the manner by which the data differs from the bell-curve, and the existence of outlying data which skews the distribution. Of the 30 sets of data tested (each linguistic feature for each group, from each of the three tests), 16 sets scored below 0.05 suggesting that the data was not normally distributed. As mentioned in section 3.11.2, the data was skewed to the right suggesting that there was a group of about 5 or 6 students within the group who were more proficient with a particular linguistic feature. This is possibly due to the manner by which classes at this particular school are streamed based on their science examination results. One of the classes used in this study was a higher ability science class. These students were evenly distributed between the two experimental groups. From my own observations of these students, there seemed to be a small group within the class of high achievers who were proficient at both science and English. Thus, the non-normal distribution of the data was an accurate representation of the students within the groups, and needed to be analysed as it was rather than transformed.

Having found that much of the data did not have normal distribution, non-parametric testing which does not assume normal distribution was the best choice for data analysis for this study. The following section will report the results of the non-parametric testing of the pre-test, post-test and delayed post-test data.

4.1.4 Comparison of overall writing accuracy scores

The overall accuracy scores made up from combining all five linguistic features for each group were compared between the focused and unfocused groups for the pre-test, post-test and delayed post-test. The descriptive statistics are presented in Table 3, which displays the mean for each group in each test, as well as the standard deviation of the spread of data. The group means for each test are also depicted visually below in Table 4.

Table 3
Group Means and Standard Deviations for the Error Correction Tests

Correction Type	Pre-test		Post-test		Delayed Post-test	
	M	SD	M	SD	M	SD
Focused (n = 19)	.82	.11	.78	.15	.76	.15
Unfocused (n = 19)	.76	.15	.71	.14	.73	.15

Table 4
Group Means for the Error Correction Tests

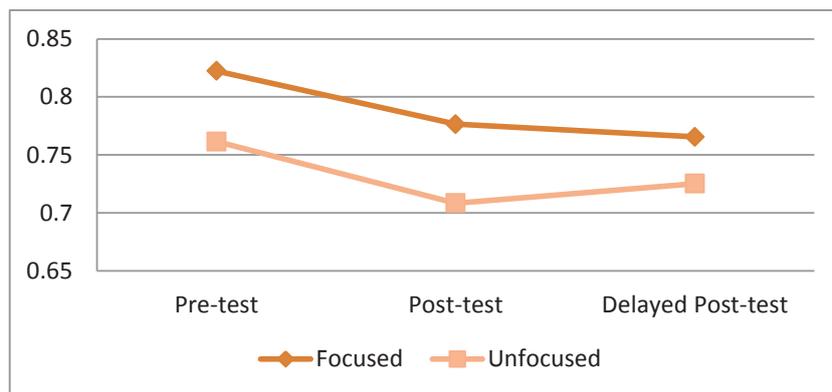


Table 4 presents the change in group mean for the error correction tests. While this graph displays an overall downward trend for both the focused and unfocused group from the pre-test to the delayed post-test, a Friedman test conducted to evaluate differences in medians among the overall accuracy scores for the focused group in the pre-test, post-test and delayed post-test found that this was not significant once again using an alpha significance level of 0.05. Thus, it is possible to say with a 95% level of confidence that the changes between the pre-test, post-test and delayed post-test accuracy scores are not statistically significant.

Similarly, another Friedman test evaluating differences in medians among the overall accuracy scores for the unfocused group in the pre-test, post-test and delayed post-test also was not significant. These results for both the focused and unfocused groups are presented in Table 5. The Friedman's test significance column displays the significance of the change between the three tests for each group. As both figures are above 0.05, this suggests that neither group's accuracy score changed significantly over the course of the study.

Table 5
Differences in Medians of the Overall Accuracy Scores

Group	<u>Friedman's test significance</u>
Focused (n = 19)	0.504
Unfocused (n = 19)	0.623

4.1.5 Comparison of writing accuracy scores for each linguistic feature

Having established no significant difference between the focused and unfocused groups' overall accuracy scores, the accuracy scores for the separate linguistic features were analysed using a series of Friedman tests. Again, an alpha of 0.5 was used as the critical level above which the result was assumed to be not statistically significant.

No significant differences were found for either group in their accuracy scores for each linguistic feature, with the exception of the unfocused group's use of the irregular past tense. For this particular linguistic feature, the unfocused group's accuracy score showed a significant change. Table 6 presents the results of the Friedman's tests for both the focused and unfocused groups for each linguistic feature. The Friedman's test significance column shows that all data reported is greater than 0.05, with the exception of the unfocused group's irregular past tense. This suggests that there was a significant difference between accuracy scores between two or more of the tests for this linguistic feature. However, further testing was required to investigate exactly which of the tests showed a significant difference to the others. This post-hoc testing is discussed in the next section.

Table 6

Differences in medians for each linguistic feature

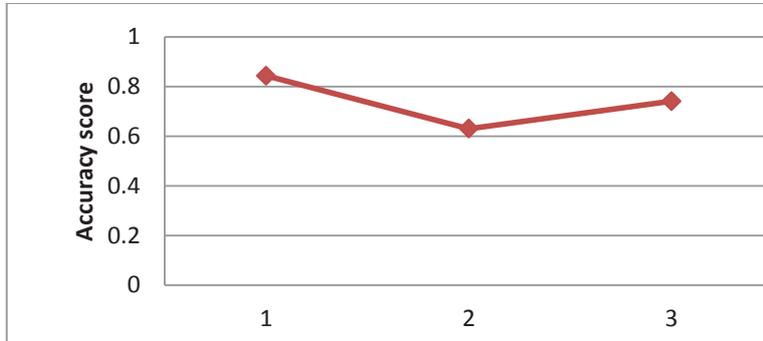
Linguistic Feature	Group	Friedman's test significance
Articles	Focused (n = 19)	0.115
	Unfocused (n = 19)	0.408
Copula 'be'	Focused (n = 19)	0.523
	Unfocused (n = 19)	0.253
Regular past tense	Focused (n = 19)	0.892
	Unfocused (n = 19)	0.113
Irregular past tense	Focused (n = 19)	0.899
	Unfocused (n = 19)	0.017
Prepositions	Focused (n = 19)	0.086
	Unfocused (n = 19)	0.368

4.1.6 Post-hoc analysis of the unfocused group's irregular past tense accuracy score

Having established that there was a significant difference between the accuracy scores from the pre-test, post-test and delayed post-test for the focused group's use of the irregular past tense, the mean test score was calculated and plotted to get a clearer picture of the overall data trend. Table 7 shows the mean trend for the unfocused group's use of the irregular past tense for the three tests. It displays the difference between the pre-test (test 1), and post-test (test 2) which is a downward trend suggesting that the accuracy score decreased following the CF received after the pre-test. However, after the post-test the scores had increased again by the delayed post-test (test 3) although they did not quite reach the pre-test levels.

Table 7

Focused Group Irregular Past Tense Means for Error Correction Tests



Key

1: Pre-test 2: Post-test 3: Delayed post-test

This difference was further investigated by using a Wilcoxon signed-rank test. This test is a useful way to examine paired samples with a repeat-measure design such as was used in this study (Green and Salkind, 2008). A Bonferroni correction was applied so all effects are reported at a .0167 level of significance. This was used in order to counteract potential of type I errors occurring due to multiple comparisons. A type 1 error occurs when it is believed that a difference exists when in fact there is none. To counteract this, the alpha selected for this study 0.05, was divided by the number of pairwise comparisons to be made, in this case three (the pre-test, post-test and delayed post-test) in order to determine the critical level of .0167.

The results of the Wilcoxon signed-rank test suggest that the students' use of the irregular past tense significantly changed from the time of the pre-test to the post-test. This represents a large effect size as it is above Cohen's (1969) benchmark of 0.5. However, the effect was not sustained by the delayed post-test. Table 8 displays the pair-wise comparisons made in the Wilcoxon signed-rank test. As can be seen in the

pre-test to post-test column, the level of significance is 0.004, which is below the .0167 level of significance.

Table 8
 Focused Group's Accuracy Scores in the Error Correction Tests

	Pre-test to post-test	Post-test to delayed post-test	Delayed post-test to pre-test
Z score	-2.897	-1.546	-1.477
Asymp. Sig (2-tailed)	0.004	0.122	0.140

4.1.7 Summary of key results from the pre-test, post-test and delayed post-test

Analysing the over group accuracy scores using a Friedman's test revealed no statistically significant difference between the written focused and unfocused CF. Once each linguistic feature was analysed separately, a negative effect was found for the focused group's use of the irregular past tense between the pre and post-tests.

However, this effect was not sustained by the time of the delayed post-test.

Having reported the results from the analysis of the pre-test, post-test and delayed post-test, this chapter will now turn to discuss the feedback sessions. This will be presented in the order of the two questionnaire questions.

4.1.8 The feedback sessions

During the eight writing sessions for which students were given feedback, particular errors for each of the five linguistic features were common among the student writing.

While the point of comparison for this research was the pre-test, post-test and delayed

post-test, it is useful to mention the types of errors which students were most frequently making, particularly as these errors were the same types of errors the students most commonly made in the pre-test, post-test and delayed post-test. Students frequently omitted the articles *a*, *an* or *the* required for definite nouns. For example one student wrote '*Then Kancil went to hole*' omitting '*the*' from before the noun '*hole*'. Omitting articles entirely was a more frequent error than using the wrong type of article. The copula *to be* was most frequently confused in the past tense form with *was* and *were* used incorrectly. An example from one student's writing which is indicative of the type of error most frequently used with this linguistic form was a misuse of *were* in the sentence '*Mouse deer were sneaky*' rather than correctly matching the subject *Mouse deer* with the verb *was*. When it came to the regular past tense, the most frequent error was students failing to use the appropriate tense. When they should have used the past tense they were using the present tense. For example, one student wrote '*then Kancil jump out of the hole*' rather than using the correct regular past tense verb *jumped*. However, the most common error I observed students making with regards to the irregular past tense was trying to conjugate it using the past tense suffix *-ed*. An example of this was when one student wrote '*Kancil foughted the crocodile*' using the regular *-ed* on the end of *fight*, rather than using the correct irregular form *fought*. Finally, frequent errors using prepositions were errors of using the wrong preposition rather than omitting prepositions. This can be seen in an example from a student's writing in the sentence '*Mouse deer jumped on the hole*'. Rather than using the correct preposition '*in*', the student selected the incorrect preposition '*on*'.

4.2 The questionnaire

4.2.1 Analysis of student perception ratings of usefulness of the written corrective feedback they received

In order to answer the second research question '*What were students' perceptions of written CF?*' the questionnaire responses were analysed. These are discussed in the order the questions were presented in the questionnaire beginning with the quantitative question requiring students to rank the usefulness of the CF they received, before moving on to the qualitative question regarding preferred types of CF.

A Mann-Whitney *U* test was used to investigate any difference between student rankings of the perceived usefulness of the written CF intervention received between the focused and unfocused groups. The results of the test were significant. The focused group had an average rank of 13.24, while the unfocused group had a significantly higher average rank of 25.76 suggesting that they perceived the written CF they received as being more useful than the focused group.

4.2.2 Content analysis of student perception of preferred type of corrective feedback

Content analysis of the open-ended survey question 'What kinds of teacher feedback do you think helps your English writing the most? Why?' was carried out using a process of 'reducing, thematitising, and theorising' the data (Casebolt and Hodge, 2010). Rather than anticipating the student responses prior to administering the questionnaire, in the interest of capturing students' subjective opinions the responses were collected first, and then categorized into the types of CF which emerged. This constant comparative

approach (Merriam, 2009) led to the appearance of two recurring themes, which perhaps due to students' limited experience with feedback types could be categorized as either focused or unfocused CF.

All 19 students from the unfocused CF group identified direct unfocused CF as being helpful in improving their English writing. One student mentioned 'if the teacher doesn't tell me where the mistakes are and what they are, then I can't learn from them', and another wrote 'I want the teacher to correct all my mistakes so that I can look at how it should be written'. Similarly, of the 19 students in the focused CF group, 15 identified direct unfocused CF as being helpful. One student wrote 'I want all of my mistakes fixed, not just some. I don't like knowing that there are mistakes in my writing that I can't see'. This theme of wanting to be made aware of where all their errors were was found in the responses of 11 students from the focused group, and 14 students from the unfocused group. However, four of the focused CF group did identify direct focused CF as being helpful, while none of the unfocused CF group did.

Having presented the results, the next chapter provides a discussion of them and their significance.

Chapter 5: Discussion

Within an AfL framework, assessment practices lie at the heart of teaching and learning. Through an effective AfL practice, teachers and students are both active participants in the learning process, constantly seeking ways to bridge the gap between existing and desired knowledge. Brunei Darussalam is currently in a period of transitioning from traditional teacher-centred methods of teaching, to student-centred AfL methods. As a result, traditional assessment and feedback methods are being called into question in light of the government's SBAfL initiative. With this in mind, the current study examined the effect of focused and unfocused written CF on L2 students' English written accuracy, as well as student perceptions of the feedback.

This chapter considers the results from the study with regards to the research questions. It begins by discussing the findings in relation to the first research question '*Is there a differential effect on L2 written accuracy following focused or unfocused written CF?*'. It then discusses the findings related to the second research question '*What are students' perceptions of written CF?*'. Following this, the implications of the study are examined within both a Bruneian and wider context.

5.1 Research Question 1: *Is there a differential effect on L2 written accuracy following focused or unfocused written CF?*

This question can be answered by looking at the results from the pre-test, post-test and delayed post-test for each of the two groups (focused and unfocused written CF). No significant difference was found in the overall rate of written accuracy including all five linguistic features (articles, copula 'be', regular past tense, irregular past tense and

prepositions) for either group from the pre-test to the post-test and delayed post-test.

This means that the type of written CF which students received did not impact on their written accuracy.

5.1.2 Similarities and differences of findings to other research

The results of the current study are in direct contrast with recent research which has suggested a differential effect between focused and unfocused written CF (Ellis et al, 2008; Bitchener and Knoch, 2010-a; Bitchener and Knoch, 2010-b; Sheen et al, 2009). Due to the findings of these studies and those investigating other types of CF e.g. Han (2002), it has been argued that focused written CF is a more effective pedagogical technique than unfocused written CF because learners are more likely to pay attention to corrections directed at a single error type, and are more likely to develop an understanding of the nature of the error (Ellis et al, 2008). Sheen et al (2009) argue that this is because when students receive unfocused written CF on a range of grammatical errors, they are unable to process the feedback effectively, and are unable to work out why they have been corrected. However, the findings of the current study did not support this claim.

5.1.2.1 The systematic nature of focused feedback

Sheen et al (2009) also found that students who received focused written CF on one grammatical feature, outperformed both an unfocused group and a control group in a post-test for overall accuracy for a range of grammatical features. One explanation they presented for this was that focused written CF is by its nature systematic, while unfocused CF can rely on teacher discretion which has been shown to often be arbitrary

and confusing e.g. (Long, 1977). They also suggested that a 'focused correction oriented them to attend to form in general, and this benefitted not just accuracy in the use of articles but accuracy in other structures as well. In other words, focused correction is clear and systematic and thereby induces learners to pay more attention to form overall' (p.566). Similarly, Bitchener and Knoch (2010-b) also argued that systematic focused written CF should be used, and teachers should negotiate with their students how many error categories will be targeted at one time, adding more as learning increases.

However, the findings of the current study did not provide evidence of any differential effects in the overall rates of L2 written accuracy between students receiving focused or unfocused written CF. Neither group improved their written accuracy following written CF. It should be noted, that the unfocused written CF used within the current study was particularly systematic and well-defined in nature. For each piece of writing, the students in the unfocused CF group received written CF for all of the five linguistic features each time they were incorrectly used. Other linguistic features with which they may have made errors were deliberately not corrected. A possible explanation for the lack of difference in the effects of the focused and unfocused written CF in the current study, is that by so tightly defining the unfocused written CF it became more systematic and thereby similar to the focused written CF in nature. It is possible that the difference between correcting five features within the student writing, and one feature was not sufficient to detect an effect. Within a classroom context, a teacher using unfocused written CF would also correct features such as punctuation, sentence structure and word choice. However, for the current study the linguistic features corrected were

limited to the five main grammatical features which I have observed students having difficulty with. This was done in order to focus on the interaction between written CF and grammar accuracy, and also to build upon the results of a previous study conducted by Sheen et al (2009). Thus, it is possible that unfocused written CF, even on a wider range of linguistic features than was included in the current study, if it was systematically issued, could have a similar effect to that previously found with focused written CF.

5.1.2.2 Differences between secondary school students and university students

Another way that the current study differs from previous studies which have found evidence supporting the use of focused written CF is that the students in this investigation were secondary school age. Previous studies have used university students as the subjects e.g. (Ellis et al, 2008; Bitchener and Knoch, 2010-a; Bitchener and Knoch, 2010-b; Sheen et al, 2009). While the level of language acquisition between the subjects in this study and those of previous studies may be comparable, the chronological age and language learning motivation may not be. Dornyei (2005) argues that language learning motivation contains two self-related concepts: the 'Ideal L2' self which contains one's ideal self-image conceptualizing the wish to be a competent L2 speaker, and the 'Ought-to L2' self which contains the 'attributes one believes one ought to possess in order to avoid possible negative outcomes' (p.106). Kormos, Kiddle and Csizer (2011) investigated the interaction between age and group-related differences in language-learning motivation. Their findings suggest that for secondary school students the 'Ideal L2' self and the 'Ought-to L2' self are unrelated constructs. In other words, secondary school students' view of the language learner they think they

should be, and how they see themselves ideally using the language in the future, are very different concepts. This is in direct contrast to university students and adult learners included in study conducted by Kormos et al (2011), where a link between the two constructs was found. Thus, the interaction between the 'Ideal L2' self and the 'Ought-to L2' self, the driving force of language learning (Dornyei, 2005), is different between secondary school students and older students.

This means that the secondary school students in the current study may not have had the same self-regulatory processes motivating their language learning as university students. In order to be motivated by the written CF given, students' 'Ought-to L2' selves need to conceptualise responding to this type of feedback as an important attribute of a language learner which both avoids negative consequences such as making errors, but also fits within their concept of the 'Ideal L2' self. If the students did not perceive the written CF to be a valuable learning tool which is related to how they want to use the language in the future, then this may have affected the value of the CF as there was no motivation for students to attend to it.

Students who attend university have also already had a degree of academic success in order to attain admittance to their courses, and thus have developed learning strategies which are at least partially successful. Secondary school students on the other hand characteristically have mounting defensive inhibitions (Brown, 1994) and teachers' approach to errors can increase students' inhibitions and language use anxiety (Fonseca and Toscano, 2007). It is possible that having their errors corrected in each piece of writing may have actually had a demotivating effect on the secondary school students in the current study, and that is the reason why their writing accuracy did not

improve with either the focused or unfocused written CF. Students' responses to the second questionnaire question '*What kinds of teacher feedback do you think helps your English writing the most? Why?*' seemed to support this. As well as identifying unfocused feedback as their preferred method, several students wrote comments such as '*I'm just not good at English*' and '*I keep making mistakes*'. It may be that the result of frequent error corrections was low English language self-esteem. This corresponds with Kluger and Denisi's (1996) findings that feedback is more powerful when it provides information regarding correct responses rather than incorrect ones. Written CF in the current study only provided negative responses to student writing, and this may have increased students' inhibitions and anxiety in a manner which older students are somewhat more resilient to.

Thus, no evidence was found in the current study that pedagogical strategies such as providing written CF which have been shown to have a positive effect on the language acquisition of university students, are also of benefit to secondary school students. However, the findings did call into question the relationship between written CF and student motivation. Secondary school students may not be able to link behaviours such as attending to negative written CF which their 'Ought-to L2' selves see as beneficial, with how they visualize their 'Ideal L2' selves. This relationship between these two concepts is an important factor influencing student language learning motivation (Dornyei, 2005). Without this link, as was evidenced by student comments in the current study, the L2 self-esteem of secondary school students may be negatively affected. This is of concern within the framework of SPAfL which aims to positively develop the students' attitudes towards learning (Ministry of Education, 2012).

5.1.2.3 Length of time taken to write

Another noteworthy finding was that during the course of the study, students began to take longer to complete their writing each week, despite the fact that the tasks were in the same format and their finished writing length did not increase. Whereas in the first writing task I noted that all students had completed their writing within the first half hour of the lesson, by the fourth and subsequent writing sessions many students were taking up to 45 minutes to complete their writing. This would seem to indicate that students were spending longer trying to become more accurate, perhaps as a sign of their language anxiety. Semke (1984) and Kepner (1991) also noted that students in CF groups wrote slower than students who did not receive CF. However, neither of these studies found any correlation between the speed of writing and accuracy. Yet producing language anxiety is something that should be avoided if students are to remain motivated to learn, and as will be discussed in more detail in section 5.3.5, this seems to be in direct contrast to the assessment for learning principles of the *SBAfL* educational initiative in Brunei.

5.1.3 The effect of written CF on individual linguistic features

Having discussed the results of the study which found no differential effect in the overall written accuracy of students following focused and unfocused written CF, consideration is now given to the results for each individual linguistic feature. One reason for analyzing each linguistic feature separately is that Bitchener (2008), Ellis et al (2008), Sheen (2007) have all argued that a problem with much previous CF research which has failed to find any positive effect of written CF (e.g. Stiff, 1967; Semke, 1984;

Kepner, 1991; Sheppard, 1992) has been its 'vague correct-all-the-errors approach' (Ferris, 2010, p192). Therefore, as well as considering any differential effect in the overall written accuracy of the students in this study (including the five linguistic features mentioned above), each feature was also considered separately in order to examine potential distinct effects.

In the current study, the only significant difference found between the focused and unfocused groups was in the post-test for the unfocused group's use of the irregular past tense. The group median for this linguistic feature went down from 1.00 in the pre-test to .56 in the post-test. This suggests that the performance of the unfocused group in the use of irregular verbs actually decreased following written unfocused CF and a large effect size of -.6647 was calculated using Cohen's classification system (1969). However, this difference was not sustained and by the delayed post-test the difference from the pre-test was not significant as the median had increased to .73. Thus, it would seem that unfocused written CF had a negative differential effect on the students' use of the irregular past tense in the short-term, but this was not carried forward to the delayed post-test at the end of the thirteen week study.

Aside from this, no significant differences were found between the focused and unfocused group medians for articles, the copula 'be', regular past tense, or prepositions. This is in direct contrast with several studies which have investigated the effect of written focused CF on students' use of the English articles system. For example, Bitchener, Young and Cameron (2005), Sheen (2007), Ellis et al, (2008), Sheen, Wright and Moldova (2009), Bitchener and Knoch (2010-a; 2010-b) all noted an improvement in students' written accuracy rates using articles following focused written

CF. Similarly, Bitchener et al (2005) reported an improvement in student use of the simple past tense following focused written CF.

5.1.3.1 Rule governed and non-rule governed language structures and written CF

Of the improvements in student written accuracy found in previous studies, all involved an improvement in the acquisition of simple rule governed linguistic forms like articles and the past simple tense (Bitchener, Young and Cameron, 2005; Bitchener et al, 2005; Sheen 2007; Ellis et al, 2008; Sheen, Wright and Moldova, 2009; Bitchener and Knoch, 2010-a; and Bitchener and Knoch, 2010-b). The negative effect found in the current study on the irregular past tense may be because the irregular past tense is not a rule-governed structure like articles and the regular past tense. Instead it involves the memorization of verb forms, which do not adhere to any logical rules. I have been unable to uncover any research which has found a positive effect of either focused or unfocused CF specifically on the irregular past tense. It is possible that the particular non rule-bound characteristics of the irregular past tense would benefit from different teacher input than those suitable for rule-governed structures. For example, within the Islamic Bruneian context, memorization through repetition is a technique commonly used for parts of the Koran, and teachers tend to rely on drill and memorization as key strategies (Cheong and Nuttman, 1995).

If we compare the nature of learning the irregular past tense as opposed to other linguistic features such as the regular past tense, it is possible that one reason for the unfocused group's written accuracy decreasing in the post-test was due to a dual-representation of knowledge (Marcus et al, 1992). This model explains the process of

learning the irregular past tense as memorising a series of separate cases, while learning the regular past tense involves learning to apply a rule.

It may well be that students would learn the irregular past tense more efficiently through pedagogical practices which mimic the memorization techniques which they are used to, and have had success with in other aspects of their education. While feedback is an essential part of an AfL curriculum, it is only one aspect of the teacher's toolbox to be used in conjunction with other pedagogies.

5.1.3.2 Differentiation and written CF

In the current study, students in the focused group only received feedback on the linguistic feature with which they were having the most difficulty. In this manner, the feedback they received was tailored to their specific needs. This highlights a potential incongruity between previous focused/unfocused written CF research to date, and actual classroom practice. Previous research (e.g. Bitchener, Young and Cameron, 2005; Sheen, 2007; Ellis et al, 2008; Sheen, Wright and Moldova, 2009; Bitchener and Knoch, 2010-a and Bitchener and Knoch, 2010-b) selected only one linguistic feature for the entire focused written CF group in their study. However, influential classroom practice theories recommend that teachers respond to errors according to student ability (Straub, 2000). There has been a move calling for teachers to differentiate their pedagogies to ensure that each student regardless of their academic abilities has equity of access to high-quality learning tailored to their particular needs (Darling-Hammond, Wise and Klein, 1999). In Brunei, classes of students are streamed by academic results in the sciences. This means that classes are made up of students with similar academic

abilities in science subjects, but vastly differing English language abilities. A blanket approach to correcting the same linguistic feature for every student in a class ignores individual student differences and needs. For example, it is unrealistic to only provide written CF for articles for every student in the class, when that may not be the knowledge gap for some of the students.

Truscott (1996) argued that the acquisition of some linguistic forms and structures has been shown to follow a natural order (Clahsen, Meisel and Pienemann, 1983), and written CF may be ineffective unless presented when the learner is ready. This is similar to Krashen's (1981) comprehensible input theory in which he also argues for a natural order in language acquisition which optimally occurs when the L2 input is one step beyond the learner's current level of linguistic competence. Due to this, Krashen argues that a teaching curriculum should offer a range of linguistic input so that each learner may receive the input which is at the optimal level for the linguistic feature which they are ready to acquire. This would seem to be in direct contrast with the previous studies on focused CF which only focused with the English articles systems (e.g. Bitchener, Young and Cameron, 2005; Sheen, 2007; Ellis et al, 2008; Sheen, Wright and Moldova, 2009; Bitchener and Knoch, 2010-a; 2010-b).

The current study highlighted a way that focused written CF could be applied within a differentiated classroom, with each student's writing corrected for the linguistic feature with which they were having the most difficulty. Consequently, some students did not receive focused written CF for some of the linguistic features that they were more competent with. This would seem to be more in line with the AfL policy being introduced to Brunei at the moment which emphasises catering to individual learners. However,

Krashen (1981) further argued that L2 teaching curriculum should not present one linguistic feature at a time in an attempt to present them to students in the order of acquisition. Rather, a range of teaching input should be presented and students would naturally respond to that which was aimed at the linguistic feature which they were ready to acquire. This view would seem to be more compatible with unfocused written CF which presents feedback to a range of linguistic features.

However, if the focused approach was effective at systematically targeting and improving one linguistic feature at a time, this should have affected the overall accuracy rate for student writing and a differential effect would have been noted between the overall accuracy rates for each group. Similarly, if unfocused written CF was effective at providing feedback which offered more of a chance of students being ready to naturally acquire a particular linguistic feature, then it should have been more effective at improving student written accuracy. As this was not the case for either type of written CF, no evidence was found to suggest any differential effect on student written accuracy following the use of written focused and unfocused CF.

5.1.4 Summary of findings regarding research question 1: *Is there a differential effect on L2 written accuracy following focused or unfocused written CF?*

With regards to the first research question 'Is there a differential effect on L2 written accuracy following focused or unfocused written CF?' the current study has found no evidence to suggest that one feedback type is more effective than the other.

Furthermore, the students' written accuracy did not improve at all, suggesting that the feedback had no effect on student writing within the 13 week timeframe investigated. A

possible reason for the lack of difference between the two types of feedback is that the unfocused written CF was delivered in a very structured and systematic manner, thereby rendering it similar to the focused CF. However, while this may account for the lack of a differential effect, it does not provide an explanation for why neither type of CF seemed to aid student written accuracy. It is possible that secondary school students do not respond to written CF which only provides negative feedback in the same manner as the University students who were the subject of previous investigations. It also seems that written CF provided for non-rule bound linguistic features such as the irregular past tense, may actually have a negative effect on language acquisition in the short-term. Given the memorization required to learn this linguistic feature, it may be that traditional methods such as rote learning and repetition which are used elsewhere in the Bruneian educational context, are more effective than CF. Likewise, while a one size fits all approach to the correction of different types of linguistic errors does not seem appropriate, neither does a blanket approach to providing all students in a class with CF for the same type of error.

The current study therefore, did not provide any evidence of a differential effect between focused and unfocused written CF, but it also calls into question the efficacy of providing written CF at all given that it did not lead to improvement in student written accuracy, and that it may not be in line with the goals of AfL. This discussion now turns to consider the second research question concerning student perception of the written CF they received.

5.2 Research Question 2: *What are students' perceptions of written CF?*

This question can be answered by looking at the results from the student questionnaire. In response to the first question of the questionnaire 'How helpful was teacher feedback on your writing in improving your writing?' a clear difference could be seen between the rating responses of the focused group to the unfocused group. The unfocused group rated the feedback they received as being significantly more helpful than the focused group did. The focused group had an average rank of 13.24, while the unfocused group had a significantly higher average rank of 25.76. This is perhaps not surprising as there is a long standing tradition in Brunei English language classrooms of unfocused CF, and as such it is the trusted status quo. Feedback is a social act (Lee, 2008) and students' preference for types of teacher feedback may stem from their past experiences whether or not they are beneficial to them (Alavi and Kaivanapanah, 2009).

5.2.1 Students' expectations of written CF

Furthermore, Hyland and Hyland (2006) argue that students are socially and historically situated beings who respond to what they see as valuable and useful, and people they see as credible. In the current study, students' previous experiences of feedback involved the correction of all errors, and it may have been that they saw that as the role of the teacher. When correction was not provided for all errors for the focused group, the students may have felt that they were not receiving as much feedback as their peers, and therefore that teacher and feedback strategy lacked credibility. For example, one student wrote on her questionnaire 'Why didn't you correct all my mistakes?', while

another wrote 'Please teacher correct all my wrongs'. These comments would seem to suggest that these particular students were not satisfied with the CF that they received.

Given the findings of recent studies which have suggested that focused written CF is a more effective strategy than unfocused written CF in improving students' written accuracy, it was important to try this strategy within this context. However, the findings of the questionnaire highlight the need for students to be active participants of any feedback approach, and it is important that they are made aware of the protocol and rationale behind any changes to feedback practices if the teacher and practice are to be viewed as credible. If students' have specific expectations of corrective feedback and their expectations are not met, language learning could be hindered (Schultz, 2002). This is particularly salient within an AfL context such as SBAfL which highlights the need for students to be engaged as active participants in the learning and assessment process.

The second question in the questionnaire 'What kind of teacher feedback do you think helps your English writing the most? Why?' gave further insight to why the majority of the students preferred unfocused written CF to focused written CF. All of the students in the unfocused group identified unfocused written CF as helpful, while 15 of the nineteen focused group members also identified unfocused written CF as helpful. Thus, overwhelmingly the students perceived that unfocused CF was more helpful despite neither group improving their written accuracy over the course of the study. These findings corroborate those of a study conducted by McCargar (1993) which revealed that in response to the statement 'Language teachers should correct every student error', all student groups except for the Japanese students agreed.

This is an important finding as student beliefs play a large role in motivation (Horwitz, 1988), and motivation is an important part of learning. Teaching activities such as feedback need to be perceived in students' minds as conducive to learning, and any discrepancies between teacher and student perceptions regarding the usefulness of a particular practice can be detrimental to learning (Schultz, 2002). Despite there being no indication of a differential effect between the use of written focused and unfocused CF in the current study, if a teacher uses focused written CF, but students do not believe that it is a useful practice, this could potentially disrupt any potential linguistic gains to be had. Within an AfL practice, whereby the feedback is actively used by the student in order to bridge a gap between current and desired knowledge (Sadler, 1989), the student should view the CF as a realistic and tangible next step in the learning process. Without this perception of the feedback, it seems unlikely that the students will attend to it.

A theme in many of the student responses reflected their desire to know where mistakes were, and a belief that in order to improve their written accuracy they must first be made aware of the mistakes. Several students also expressed discomfort at knowing that there were mistakes in their work but being unable to see them. Research suggests that in a quest for L2 accuracy, students are eager to have their errors pointed out by their teacher (Komura, 1999; Lee, 2005; Leki, 1991), and that they prefer direct corrective feedback of their errors (Radecki and Swales, 1998; Lee, 2005). These findings were confirmed by the results of this study. The students who received the written focused CF expressed a concern that they potentially had other errors in their writing of which they were not aware, and this seems to have caused them unease.

This highlights a tension between the students expressing a desire to have all of their errors corrected as has been their previous experiences with traditional unfocused written CF, and the demotivating effect that it appeared to have on the students as was discussed in sections 5.1.2.2, and 5.1.2.3. Feedback is more effective when it deals with faulty interpretations rather than a total lack of understanding, and as seems to have been the case in the current study, it can in fact damage student motivation if they do not understand it (Hattie & Timperley, 2007).

5.2.2 Summary of findings regarding research question 2: *What are students' perceptions of written CF?*

The students in the current study overwhelmingly perceived unfocused written CF as being more useful than focused written CF. This can be seen as a manifestation of the students as socially situated beings. Brunei has a long history of using unfocused CF, and it is what the students are used to. Therefore, despite the overall accuracy scores in this study showing no differential effect between the two types of feedback, the students preferred the status quo. No relationship between student perception of feedback type, and improvement in writing was found. The current study also revealed that a major reason for the students preferring unfocused written CF was that they believed that in order to improve their written accuracy, they must first be made aware of every mistake. This was in line with research from other Asian cultures (e.g. McCargar, 1993). This highlights a potential problem, as when student perceptions about effective feedback, and the feedback they receive differ, language learning can potentially be disrupted. Having discussed the findings regarding the two research questions for this study, the wider implications of these findings will now be examined.

5.3 Wider implications of the study

Generalisations cannot be argued based on data collected from a small sample of students in one secondary school. However, the implications considered in this section may be applicable to similar L2 English contexts.

5.3.1 Focused or unfocused written corrective feedback?

Findings from this study suggest that there is no differential effect between the use of written focused or unfocused CF (as defined in the current) study on students' L2 English written accuracy. Within Brunei, where written CF is a mandated practice, these findings indicate that there should be no particular requirement for one type of written CF over another. The fact that neither group showed any significant rate of improvement in their L2 English written accuracy also brings into question the efficacy of written CF as an instructional practice in general. However, whether or not written CF is more effective than no CF is beyond the scope of this investigation.

5.3.2 Focused written corrective feedback and differentiation

Using focused written CF which corrects the same linguistic feature for every learner is in direct contrast with the practice of differentiation. Teachers are being strongly encouraged to differentiate instructional techniques and feedback (Darling-Hammond, Wise and Klein, 1999), and this makes correcting the same linguistic feature for every student impractical in a differentiated class. Providing focused written CF which differentiates between students can be a complex procedure. A teacher must first have an awareness of a student's particular L2 ability and weaknesses before they are able to provide the CF on a particular linguistic feature. They must then monitor the student's

future use of the feature. The use of focused written CF in the current study has highlighted the need for a systematic approach to the use of focused written CF which cater for students' individual needs if it is to be in line with the concept of differentiation of feedback. Indeed, with the introduction of SBAfL to Brunei, which holds differentiation of feedback at its very core, it is even more important that all feedback is catered particularly to the needs of the individual learner.

An implication which was highlighted in the current study is the possibility of students not receiving written CF at all for linguistic features other than the targeted feature. This is an area which requires further consideration as potentially a student could be developmentally ready to acquire a particular feature (Krashen, 1981) and would therefore respond well to receiving written CF on it, but due to another feature being identified by the teacher as requiring focused written CF they may receive no teacher input about it at all.

5.3.3 Student perception and uptake of written corrective feedback

The results of the study showed that within this particular educational context, students perceived unfocused written CF as being more useful in helping them to improve their writing than focused written CF. As feedback is a social act (Lee, 2008), students' perceptions of written CF are greatly influenced by their previous experiences and expectations of it. Within Brunei traditional L2 English teaching methods have heavily relied on unfocused written CF, and this has meant that students have come to view this as being an effective practice. Added to that, L2 learners tend to want to have all their errors pointed out to them (Komura, 1999; Lee, 2005; Leki, 1991), and this finding was

replicated in the comments made by students in this study. Thus, consideration of the interactions of student perceptions of the written CF and their uptake of it, while beyond the scope of this study, is an issue which it has highlighted as a potential concern. It may be the case that in order for students to perceive focused CF as a useful practice, they need to know the rationale behind its use. Alternatively, it may take time for the practice to become viewed as normal while students are getting used to it.

5.3.4 Focused and Unfocused written CF and School Based Assessment for Learning (SBAfL)

In 2007 Brunei introduced a *National Education System SPN21* aimed at improving education within the Sultanate. *SBAfL* is a big part of this initiative, which emphasizes the use of assessment for learning best practices such as providing students with effective feedback about their learning. These changes have meant that examining current assessment practices is particularly salient within this context.

The findings of the current study highlight some issues regarding the use of written CF, which appear to be at odds with the principles of *SBAfL*. Firstly, as outlined above, students began to take longer to complete their writing, yet the length of their writing did not increase as the study progressed. This could mean that as they were becoming more anxious about their writing and so took longer trying to be more accurate. Yet their accuracy rates did not improve, suggesting that students were attempting to improve their writing without really knowing how. That can be an anxiety inducing situation for students, and can potentially have a demotivating effect which is in direct contrast to the goals of *SBAfL*. This is an area which requires further serious consideration.

Within SBAfL, feedback should be seen as a positive learning tool which illuminates the next step required to bridge a knowledge gap, not as something which produces anxiety. While written CF provides the corrections for students, no other explanation is given of particular grammar rules. They are expected to just understand why the error was wrong by looking at the correction. Yet, the results of this study indicate that this is not the case. Student writing did not improve after either the focused or unfocused written CF. Therefore, if written CF in this form is to be used within the context of SBAfL questions must be raised about how the students will receive the input they require from the assessment task in order to bridge their knowledge gap. This study has illuminated that written CF alone is not enough to aid student learning.

Added to this, there is a tension between what the students want from feedback, and the effect that this has on their motivation. While the students in the current study expressed a preference for unfocused written CF in the questionnaire, other comments they made in the questionnaire seemed to suggest that the students found the negative correction of so many errors demotivating. This perhaps demonstrates the challenges associated with introducing feedback practices into Brunei which may differ from those historically used.

The practice of providing written CF needs reconsidering within a SBAfL framework. This study has not found any benefit of either focused or unfocused written CF on student written accuracy. However, it has highlighted their demotivating effects and the manner in which they are at odds with the new direction of education in Brunei.

Chapter 6: Conclusion

In noting the increase of research regarding written CF since Truscott (1996) called for its abandonment from L2 programmes, Ferris (2010) wrote that 'it is safe to say that the benign neglect approach (Ferris, 1995) to accuracy issues in L2 writing has ended' (p.185). Written CF is no longer a practice with an assumed benefit despite a dearth of research investigating its effectiveness. Academic debate regarding the differential types and effects of CF has seen a resurgence of late, but as yet the debate does not seem to be one which has entered the classroom. However, as Truscott (1999) argued:

Teachers must constantly make decisions about what to do – and what not to do – in their classes. These decisions are necessarily made under conditions of uncertainty: research never puts an end to doubt.....the best we can hope for is that teachers will look seriously at the case against grammar correction, compare it to the case for correction, decide which is the stronger, and then incorporate that decision in their teaching (p.121).

Within Brunei Darussalam written CF has continued to be a government-mandated practice which as yet remains unchallenged in the classroom. In the light of educational changes in the Sultanate emphasizing assessment for learning pedagogies, surely now is a salient time to look at the place and type of written CF in Brunei.

In response to claims that focused written CF is more effective than unfocused CF (e.g. Ellis et al, 2008; Bitchener and Knoch, 2010-a; Bitchener and Knoch, 2010-b; Sheen et al, 2009), the results of the current study provide no evidence of a differential effect within a Bruneian secondary school. While the absence of a control group in the current study means that focused and unfocused CF cannot be compared to student accuracy rates following writing practice alone, the finding that students' overall written accuracy

did not improve after either type of CF does call into question the efficacy of the practice. In particular, unfocused CF seemed to have a negative effect on the use of non-rule bound linguistic features like the irregular past tense in the short-term, suggesting that written CF may be an unsuitable pedagogical practice for these types of language structures.

Furthermore, students in this study began to write slower following written CF. This could be an indication of a developing language anxiety as the students take longer writing in an attempt to avoid errors. While the limitations of the current study mean that these findings are far from conclusive, they are an indication that within this particular educational context, written CF could be a negative experience for students. This should be further investigated and considered as if this is indeed the case, then using written CF is incongruent with the philosophy of assessment for learning as emphasized by the SBAfL initiative in Brunei Darussalam.

Given the wide range of pedagogies available to a classroom teacher, consideration must be given to whether or not written CF is the best method for improving a student's use of non-rule governed linguistic features in particular, but also of the other linguistic features investigated in this study. At the very least, the current study has indicated that written CF should be supplemented by other teaching methods to advance language learning. It has demonstrated that relying on written CF as is the dominant method by which to improve student writing, is not an effective practice which may have negative effects on student motivation. Equally however, abandoning the practice altogether would potentially cause students' anxiety as it is counter to their socially and historically formed expectation of language learning. Therefore, it is pertinent at this time of change

in Brunei to reach a middle ground whereby written CF becomes less dominant as alternative AfL strategies are introduced and become normalised.

6.1 Limitations of the study

The scope of the research questions in this study were limited by the inability to have a control group of students who did not receive any written CF at all. The Ministry of Education in Brunei mandates that all secondary school students have one written composition corrected each fortnight, with feedback provided. Therefore it was not possible within this context to conduct a study with a control group withholding feedback, as all students must receive written CF each fortnight. It is entirely possible that the students' written accuracy could have improved without any written CF due to writing practice or natural phases of L2 acquisition alone. Hence, this study limited the research questions to investigating any differential effect between two types of written CF and student perceptions of it, rather than if written CF was an effective practice within secondary schools in Brunei.

Another limitation of the study was the short time period which it was conducted over. Due to constraints such as frequent examinations throughout the year in secondary schools in Brunei, and the limitations of conducting Masters level research within the timeframe of a year, the data collection phase of the study only lasted for 13 weeks. However, language acquisition is a process which takes a long period of time, and it is possible that in order to see the benefits of either type of written CF, a longer period of investigation would be required.

The small sample size used in this study was also a limitation which affects the ability to make generalisations about these results beyond its immediate context. Negotiating access to students within this school, and the need for one class of students to withdraw from the study due to other commitments necessitated the use of only two classes of students.

When the study was initially designed, the quasi-experiment investigating the first research question 'Is there a differential effect between focused and unfocused written CF?' was the main area of focus, with the second research question regarding student perceptions of CF being a minor subsidiary to this. For this reason the questionnaire only consisted of two broad questions. However, given the findings that students clearly preferred one feedback type over the other despite neither helping their written accuracy, it would have been of interest to delve further into these perceptions with the use of further questions.

Another potential limitation of the study lies within the comparability of the pre-test, post-test and delayed post-test. While each task was designed in order to generate a piece of student writing set in the past tense, in a format which they are used to in tests within this context, students were allowed freedom to write outside of the given topic if they found this easier. For example, the post-test depicts pictures of a person playing basketball who injures their elbow. Students were given the freedom to change details such as the sport played, or type of injury sustained to suit their own vocabulary knowledge and previous experiences. It is possible that this affected the extent to which the tests were the same for all students, and the use of more standardized writing tasks without the freedom of variability should be considered in subsequent studies. However,

if students were unable to change details such as the in the sporting example given above, then the tests may have been more of a vocabulary knowledge challenge than a test of writing grammatical accuracy.

Furthermore, while both the weekly writing tasks and the tests were designed to elicit narrative writing in the past tense, it is possible that the difference between these writing tasks and tests actually led to a different type of writing being elicited. Future investigations could consider utilising a closer match between writing tasks and tests.

Another difficulty encountered within this context concerns the difference between the type of writing students are expected to produce for the O level examination, and the type of writing they will encounter in their everyday lives. The narrative writing tasks used in the current study closely align with the requirement to write a narrative story in the Cambridge 'O' level examinations which students in Brunei sit at the end of year 11. However, in terms of measuring language acquisition as students are likely to use English within their everyday lives, narrative writing is not likely to be necessary. Therefore, it is possible that there may be a difference in the effect of written CF on student writing in a more naturalistic context rather than one which mimics an examination. This potential discrepancy is an area which may be worthy of further investigation.

Finally, the extent to which students may be able to memorise the story in the weekly writing tasks needs to be discussed. There is the possibility that students were able to memorise the words from the mouse deer story they read, and thus their writing was not a genuine reflection of their ability at that time. However, all the stories were selected

after discussion with Malay colleagues who felt that the stories were so well-known that most students would know them prior to reading them, and this seemed to be the case. Furthermore, these tasks were only used in order to generate writing which would be the subject of written CF. The testing phase used sequential pictures in order to avoid any possibility of memorisation. Nethertheless, future research could consider methods of generating student writing which do not allow for the possibility of memorisation.

6.2 Suggestions for future research

This study revealed no differential effect in the overall accuracy rates of students following either written focused or unfocused CF. However, this does not mean that differential learning was not taking place. It may be the case that the continuation of students making errors throughout the period of the study, were in fact necessary aspects of their processing and exploring the correct formation of the linguistic features. Therefore, a longitudinal study exploring differential effects between written focused and unfocused CF could perhaps uncover findings which this thirteen week study could not.

Another aspect for future research and consideration is an examination of the ways classroom teachers apply focused CF across a variety of linguistic features, with a particular interest in their use of differentiation. This is an aspect of written CF research which needs to be examined within the real life context of a diverse classroom, rather than imagining a homogeneity of English ability. Further investigation regarding systematic approaches to a variety of types of written CF would be beneficial in order to compare systematic methods of providing differentiation in a class.

Perhaps the most urgent area requiring further investigation within Brunei Darussalam given the current educational climate is the tension between existing and trusted feedback practices such as written CF, with the new pedagogies and philosophies emphasized by the SBAfL initiative. Students' and teachers' attitudes towards existing and new pedagogies can have a profound effect on their acceptance at the grass roots classroom level and how these are affecting the change implementation requires further attention. Added to this, the manner that techniques such as written CF can be used in conjunction with AfL methods in order to reach a middle ground of pedagogies should be considered.

The current study has in many ways raised more questions about written CF than it has answered. However at a period of great educational change in Brunei, perhaps it is a pertinent time to be asking questions. Therefore, I hope that this is the beginning of a dialogue regarding what has been a deeply entrenched language teaching practice. It is only through regarding our teaching practice with critical eyes that we will be able to discern the manner through which classroom pedagogies best enable our students for the future. As we move forward into the era of SBAfL in Brunei Darussalam, it is important that we are not just marking time blindly by using written CF without question. It is time to reflect on the assessment practices which will carry our students forward in the 21st century.

7. References

- Adams, R., Nuevo, A. and Egi, T. (2011). Explicit and Implicit Feedback. Modified Output, and SLA: Does Explicit and Implicit Feedback Promote Learning and Learner-Learner Interactions?, *The Modern Language Journal*, 95, Supplementary Issue, 42-63.
- Alavi, S. and Kaivanapanah, S. (2009). Feedback expectancy and EFL learners' Achievement in English, *Porta Linguarum*, 11, 99-114.
- Anseel, F., Lievens, F. and Schollaert, E. (2009). Reflection as a strategy to enhance task performance after feedback. *Organizational Behavior and Human Decision Processes*, 110, 23-35.
- Ashwell, T. (2000). Patterns of teacher response to student writing in a multi-draft Composition Classroom: Is content feedback followed by form feedback the best method? *Journal of Second Language Writing*, 9(3), 227-257.
- Berg, B.L. (2009). *Qualitative Research Methods for the Social Sciences*. Boston, MA: Allyn Bacon.
- Blakie, N. (2010). *Designing social research* (2nd ed.). Cambridge, England: Polity Press.
- Bitchener, J. and Knoch, U. (2008). The value of written corrective feedback for migrant and International students. *Language Teaching Research* 12(3), 409-431.
- Bitchener, J. and Knoch, U. (2010-a). Raising the linguistic accuracy level of advanced L2 writers with written corrective feedback. *Journal of Second Language Writing*, 19, 207-217.
- Bitchener, J. and Knoch, U. (2010-b). The contribution of written corrective feedback to Language development: A ten month investigation. *Applied Linguistics*, 31(2), 193-214.

- Bitchener, J., Young, S. and Cameron, D. (2005). The effects of different types of Corrective feedback on ESL student writing. *Journal of Second Language Writing* 14, 191-205.
- Borg, W. R., Gall, J. P. and Gall, M. D. (1993). *Applying educational research* (3rd ed.). New York, NY: Longman.
- Bradley, W. and Shaefer, K. (1998). Limitations of measurement in the social sciences. *The uses and misuses of data and models: The mathematization of the human sciences*. Thousand Oaks, CA: Sage.
- Brown, H. (1994). *Principles of Language Learning and Teaching*, New Jersey: Prentice Hall Regents.
- Brown, H. (2007). *Principles of Language Learning and Teaching* (5th ed.). White Plains, NY: Pearson Education.
- Brown, S. (2004). Assessment for learning. *Learning and Teaching in Higher Education*, 1, 81-89.
- Bruning, R. and Horn, C. (2000). Developing motivation to write. *Educational Psychologist*, 35, 25-37.
- Bryman, A. (2004). *Social research methods* (2nd ed.). Oxford, England: Oxford University Press.
- Butler, D. L. and Winne, P. H. (1995). Feedback and self-regulated learning: A Theoretical synthesis. *Review of Educational Research*, 65, 245-281.
- Carpenter, K. (1992). Kancil: From Mischief to Moral Education, *Western Folklore*, 51(2).
- Casebolt, K. and Hodge, S. (2010). High school physical education teachers' beliefs about teaching students with mild to severe disabilities. *The Physical Educator*. 67(3), 140-155.

- Cawley, B.D., Keeping, L.M. and Levy P.E. (1998). Participation in the performance appraisal process and employee reactions: A meta-analytic review of field investigations. *Journal of Applied Psychology*, 4, 615-633.
- CfBT (2006). *CfBT Teacher Handbook: Secondary Professional Aspects*. Bandar Seri Begawan: CfBT.
- Chandler, J. (2000), *The efficacy of error correction for improvement of accuracy of L2 student writing*. Paper presented at the AAAL Conference, Vancouver, BC.
- Cheong, D. and Nuttman, C. (1995). *Education Policy: a global or local agenda?* SEAMEO-INNOTECH conference on educational challenges. Manila, Philippines
- Clahsen, H., J. Meisel, and M. Pienemann. (1983). *Deutsch als Zweitsprache. Der Pracherwerb ausländischer Arbeiter*. Narr.
- Cohen, J. (1969). *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cook, T. D. and Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Chicago, IL: Rand McNally.
- Conover, W. J., Johnson, M. E. and Johnson, M. M. (1981). A comparative study of tests for homogeneity of variances, with applications to the outer continental shelf bidding data. *Technometrics*, 23, 351-36.
- Corder, S. (1967). The significance of learners' errors. *International Review of Applied Linguistics*, 5, 161-170
- Cornfield, J. and Tukey, J. W. (1956), Average Values of Mean Squares in Factorials, *Annals of Mathematical Statistics*, 27, 907–949.
- Creswell, J. W. and Clark, V. L. P. (2007). *Designing and conducting mixed methods research*. Thousand Oaks, CA: Sage.

- Creswell, J. and Miller, G. (1997). Research methodologies and the doctoral process. *New Directions for Higher Education*, 99, 33-46.
- Cronbach, L.J., Gleser, G.C., Nanda, H. and Rajaratnam, N. (1972). *The Dependability Of Behavioral Measurements*. Wiley, New York.
- Cummins, J. (1998). Immersion education for the millennium: What have we learned from 30 years of research on second language immersion? In M. R. Childs and R. M. Bostwick (Eds.) *Learning through two languages: Research and practice*. Second Katoh Gakuen International Symposium on Immersion and Bilingual Education (p. 34-47), Katoh Gakuen, Japan.
- Darling-Hammond, L., Wise, A. and Klein, S. (1999). *A license to teach: Raising Standards for teaching*. San Francisco, CA: JosseyBass.
- Denzin, N.K. (2010). Moments, Mixed Methods and Paradigm Dialogues. *Qualitative Inquiry*, 16(6).
- Dornyei, Z. (2005). *The Psychology of the Language Learner: Individual Differences in Second Language Acquisition*. Mahwah, NJ: Lawrence Erlbaum.
- Dornyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. Oxford, England: Oxford University Press.
- Earley, P. C. and Stubblebine, P. (1989). Intercultural assessment of performance feedback. *Group and Organizational Studies*, 14, 161-181.
- Ellis, L. (1994). *Research methods in the social sciences*. Madison, WI: Brown & Benchmark.
- Ellis, R. (2001). Investigating form-focused instruction. *Language Learning*, 51(Suppl. 1), 1-46.
- Ellis, R. (2006). Current Issues in the Teaching of Grammar: An SLA Perspective. *TESOL Quarterly* 40(1), 83-107.

- Ellis, R., Loewen, S and, Erlam, R., (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition* 28. 339-368.
- Ellis, R., Sheen, Y., Murakami, M. and Takashima, H., (2008). The effect of focused and unfocused written corrective feedback in an English as a foreign language classroom. *System*, 36, 353-371.
- Emig, J. (1977). Writing as a mode of learning. *College Composition and Communication*, 28, 122.
- Eskey, D. (1983). Meanwhile back in the real world...; Accuracy and fluency in second language teaching. *TESOL Quarterly*, 17, 315-323.
- Fathman, A. and Whalley, E. (1990). Teacher response to student writing: focus on form versus content. In Kroll, B. (Ed.), *Second Language Writing: Research Insights for the Classroom*. Cambridge: Cambridge University Press, 178-190.
- Ferris, D.R. (1995). Can advanced ESL students be taught to correct their most serious and frequent errors? *CATESOL Journal* 8, 41-62.
- Ferris, D.R. (1997). The influence of teacher commentary on student revision. *TESOL Quarterly* 31, 315-339.
- Ferris, D.R. (2006). Does error correction help student writers/new evidence on the short- and long-term effects of written error correction. In: Hyland, K., Hyland F., (Eds.), *Feedback in Second Language Writing: Contexts and Issues*. Cambridge University Press, Cambridge, 81-104.
- Ferris, D.R. (2010). Second Language Writing Research and Written Corrective Feedback in SLA: Intersections and Practical Applications, *Studies in Second Language Acquisition*, 32, 181-201.
- Ferris, D. and Hedgcock, J. (2005). *Teaching ESL composition: Purpose, process, and*

- practice*, Mahwah, NJ: Erlbaum.
- Ferris, D.R. and Roberts, B. (2001). Error feedback in L2 writing classes: how explicit does it need to be? *Journal of Second Language Writing* 10, 161-184.
- Field, A. (2009). *Discovering Statistics Using SPSS*. London: Sage.
- Fink, Arlene (2003). *The survey handbook*. Thousand Oaks, CA: Sage.
- Flick, U. (2002). *An introduction to qualitative research*. London: Sage.
- Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in The analysis of variance. *Journal of the American Statistical Association*, 32, 675-701.
- Fonseca-Mora, M. and Toscano-Fuentes, C. (2007). Fostering teenagers' Willingness to learn a foreign language, pp128-143. In Rubio- Alcalá, F. (ed.) *Self-esteem and foreign language learning*, New Castle: Cambridge Scholars Publishing.
- Gergen, M. and Gergen, K. (2000). Qualitative inquiry, tensions and transformations. In N. Denzin & Y. Lincoln (Eds.), *The landscape of qualitative research: Theories and issues* (pp. 1025-1046), Thousand Oaks, CA: Sage.
- Glass, G., Peckham, P. and Sanders, J., (1972). Consequences of failure to meet assumptions underlying the fixed effects analyses of variance and covariance. *Review of Educational Research*, 42(3), 237-288.
- Green, S. and Salkind, N. (2008). *Using SPSS for windows and Macintosh: Analyzing and Understanding Data*, Upper Saddle River, NJ: Pearson Education.
- Griffin, C.W. (1982). Theory of responding to student writing: The state of the art. *College Composition and Communication*, 33, 296-301.
- Hall, D.T. (2002). *Careers in and out of organizations*. Thousand Oaks, CA: Sage.
- Han, Z. (2002). A study of the impact of recasts on tense consistency in L2 output. *TESOL Quarterly* 36, 543-572.

- Hattie, J. A. (1999). Influences on student learning (Inaugural professorial address, University of Auckland, New Zealand) in Hattie, J., Timperley, H. (2007). *The Power of Feedback. Review of Educational Research, 1*, 81-112.
- Hattie, J. and Timperley, H. (2007). *The Power of Feedback. Review of Educational Research, 1*, 81-112.
- Hayes, J. R. (2000). A new framework for understanding cognition and affect in writing. In R. Indrisano & J. R. Squire (Eds.), *Perspectives on writing: Research, theory, and practice* (pp. 6-44). Newark, DE: International Reading Association.
- Hayes, J. and Flower, L. (1980) Identifying the organization of writing processes. In Gregg, Lee, Steinberg and Erwin (eds.) *Cognitive processes in writing: An Interdisciplinary approach*. Hillsdale, NJ: Lawrence Erlbaum, 3–30.
- Hayes, J. and Flower, L. (1983). Uncovering cognitive processes in writing: An introduction to protocol analysis. In P. Mosenthal, L. Tamar, and S. Walmsley, (Eds.), *Research in writing*. New York, NY: Longman. 206-220.
- Heppner, P.P., Kivlighan, D.M. and Wampold, B.E. (1992). Major research designs. In C. Verduin (Ed.), *Research designs in counseling*. Pacific Grove, CA: Brooks-Cole.
- Hofstede, G. (1984). *Culture's Consequences: International differences in work related values*. Beverly Hills, CA: Sage.
- Hofstede, G., and Hofstede, G.J. (2005). *Culture and organizations: Software of the mind*. 2nd ed. New York: McGraw-Hill.
- Holloway, I. and Wheeler, S. (2002). *Qualitative Researching* (2nd ed.), Oxford: Blackwell Publishing.
- Horowitz, D. (1986). Process not product: Less than meets the eye. *TESOL Quarterly, 20*, 141-144.

- Horwitz, E. (1988). The beliefs about language learning of beginning university foreign language students. *Modern language Journal*, 72, 283-294.
- Houghton, C., Casey, D., Shaw, D. and Murphy, K. (2010). Ethical challenges in qualitative research: examples from practice, *Nurse Researcher*, 18(1).
- Howe, K. R. (1988). Against the quantitative-qualitative incompatibility thesis or dogmas die hard. *Educational Researcher*, 17(8), 10-16.
- Hsieh, H. and Shannon, S. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9), 1277-1288.
- Hyland, K. and Hyland, F. (2006). Interpersonal aspects of response: Constructing and interpreting teacher written feedback. In K. Hyland and F. Hyland (Eds.), *Feedback in ESL writing: Contexts and issues* (pp. 206–224). Cambridge:Cambridge University Press.
- Jick, T.D. (1983). Mixing qualitative and quantitative methods: Triangulation in action. In J. Van Maanen (Ed.), *Qualitative Methodology* (pp. 135-148). Beverly Hills, CA: Sage.
- Johnson, R. and Onwuegbuzie, A. (2004). Mixed Methods Research: A Research Paradigm Whose Time Has Come, *Educational Researcher* 33(7), 14-26.
- Kane, M. (2002). Inferences about Variance Components and Reliability-Generalizability Coefficients in the Absence of Random Sampling. *Journal of Educational Measures*, 39(2), 165-181.
- Kepner, C.G. (1991). An experiment in the relationship of types of written feedback to the development of second-language writing skills. *Modern Language Journal*, 75, 305-313.
- Kluger, A . N. and DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback

- intervention theory. *Psychological Bulletin*, 119(2), 254-284.
- Kormos, J., Kiddle, T. and Csizer, K. (2011). Systems of Goals, Attitudes, and Self-related Beliefs in Second-Language-Learning Motivation. *Applied Linguistics*. 32(5), 495-516.
- Komura, K. (1999). *Student response to error correction in ESL classrooms*. Unpublished master's thesis. Sacramento: California State University.
- Krashen, S. (1981). *Second Language acquisition and second language learning*. Oxford: Pergamon
- Krashen, S.D. (1982). *Principles and Practices in Second Language Acquisition*. Oxford: Pergamon.
- Krashen, S. (1984). *Writing: Research, theory and applications*. Oxford: Oxford University Press.
- Kulhavy, R. W. (1977). Feedback in written instruction. *Review of Educational Research*, 47(1), 211-232.
- Lalande, J.F. (1982). Reducing composition errors: an experiment. *Modern Language Journal*, 66, 140-149.
- Lee, H. (2011). Experimental Methodology in English Teaching and Learning: Method features, Validity issues, and Embedded experimental design. *English Teaching: Practice and Critique*, 11(2), 25-45.
- Lee, I. (2005). Error correction in the L2 writing classroom: What do students think? *TESL Canada Journal*, 22(2), 1–16.
- Lee, I. (2008). Student reactions to teacher feedback in two Hong Kong secondary classrooms. *Journal of Second language Writing*, 17, 144-164.
- Leedy, P.D. (2001). *Practical Research: Planning and Design (8th ed.)*. Upper Saddle River, NJ: Prentice Hall.

- Leki, I. (1991). The preferences of ESL students for error correction in college-level Writing classes. *Foreign Language Annals*, 24, 203–218.
- Levene, H. (1960). Contributions to Probability and Statistics: Essays in Honour of Harold Hotelling, Olkin et al (eds), Standord, CA: Stanford University Press, 278-292.
- Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). New York: Academic Press.
- Lyster, R., (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition* 26, 399-432.
- McCargar, D. (1993). Teacher and student role expectations: Cross-cultural differences and implications. *Modern Language Journal*, 77, 192-207.
- Marcus, G. F., Pinker, S., Ullman, M., Hollander, M., Rosen, T. J. and Xu, F. (1992). Overregularization in language acquisition. *Monographs of the Society for Research in Child Development*, 57 (4), 1–182.
- Martin, P. (1996). Code-switching in the Primary Classroom: One response to the Planned and Unplanned Language Environment in Brunei. *Journal of Multilingual and Multicultural Development* 17(2), 128-144.
- Merriam, S.B. (2009). *Qualitative research: A guide to design and implementation*. San Francisco, CA: John Wiley and Sons.
- Ministry of Education. (2012). *Ministry of Education Brunei Darussalam: Strategic Plan 2012-2017*. Strategic Management Unit Department of Planning, Development And Research: Bandar Seri Begawan.
- Mitchell, E.S. (1986). Multiple triangulation: A methodology for nursing sciences. *Advances in Nursing Science* 8(3), 18-25.

- Myles, J. (2002). Second Language Writing and Research: The Writing Process and Error Analysis in Student Texts, *Teaching English as a Foreign Language E-Journal*, 6(2).
- Ngan, M., Lee, J. and Brown, G. (2010). Hong Kong Principals' Perceptions on Changes in Evaluation and Assessment Policies: They're Not for Learning. *Asian Journal of Educational Research and Synergy*, 2(1), 36-46.
- Noordzij, M., Tripepi, G., Dekker, F., Zoccali, C., Tranck and M., Jager, K. (2010). Sample size calculations: basic principles and common pitfalls. *Nephrology Dialysis Transplantation*, 1 of 6, doi: 10.1093/ndt/gfp732.
- Radecki, P. and Swales, J. (1988). ESL student reaction to written comments on their Written work. *System*, 16, 355–365.
- Ross, K., (2005). *Quantitative research methods in educational planning*. UNESCO Institute for Educational Planning.
- Ross, S. and Morrison, G. (1996). Experimental research methods. In D. H. Jonassen Ed.), *Handbook Of research for educational communications and technology* (pp. 1170-1213). Mahwah, NJ: Lawrence Erlbaum Association.
- Sadler, D. R. (1989). Formative Assessment and the Design of Instructional Systems. *Instructional Science* 18(2), 119-144.
- Schultz, R. (2002). Cultural Differences in Student and Teacher Perceptions Concerning The Role of Grammar Instruction and Corrective Feedback: USA-Colombia, *The Modern Language Journal*, 85(2), 244-258.
- Schuman, H. and Presser, S. (1981). *Questions and Answers on Attitude Surveys. Experiments in Question Form, Wording, and Context*. New York: Academic Press.
- Scott, R. and Fisher, D. (2002). *The Impact of an In-Service Course for Primary*

- Teachers*. Paper presented at the annual conference of the Australian Association for Research in Education. 1-5 December, Brisbane.
- Semke, H.M. (1984). Effects of the red pen. *Foreign Language Annals*, 17, 195-202.
- Sheen, Y. (2007). The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *TESOL Quarterly*, 41, 255-283.
- Sheen, Y., Wright, D. and Moldawa, A. (2009). Differential effects of focused and unfocused written correction on the accurate use of grammatical forms by adult ESL learners. *System* 37, 556-569.
- Sheppard, K. (1992). Two feedback types: do they make a difference? *Regional English Centre Journal* 23, 103-110.
- Shute, V. (2008). Focus on formative feedback. *Review of Educational Research*, 78, 153-189.
- Silva, T. (1988). Comments on Vivian Zamel's "Recent research on writing pedagogy", *TESOL Quarterly*, 22, 517-519.
- Spada, N. and Lightbown, P. M. (2008). Form-focused instruction: Isolated or integrated? *TESOL Quarterly*, 42, 181-207.
- Stake, R. (2000). Case Studies. In Denzin N., Lincoln Y., (Eds) *Handbook of Qualitative Research* (2nd ed), Thousand Oaks, CA: Sage.
- Stiff, R. (1967). The effect upon student composition of particular correction techniques. *Research in the Teaching of English*, 1, 54-75.
- Straub, R. (2000). The student, the text, and the classroom context: A case study of teacher response. *Assessing Writing*, 7, 23-55.
- Sully de Luque, M. and Sommer, S (2000). The Impact of Culture on Feedback-Seeking Behavior: An Integrated Model and Propositions, *The Academy of Management Review*, 25(4), 829-849.

- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition*. Rowley, MA: Newbury House, 235-252.
- Symonds, J.E. and Gorards, S. (2010). Death of mixed methods? Or the rebirth of research as a craft. *Evaluation & Research in Education*, 23(2), 121-136.
- Teddlie, C. and Tashakkori, A. (2003). Major issues and controversies in the use of mixed methods in the social and behavioural sciences. In A. Tashakkori and C. Teddlie (Eds.), *Handbook of mixed-methods in social and behavioral research* (pp. 3-50). Thousand Oaks, CA: Sage.
- Thye, S. R. (2007). Logical and philosophical foundations of experimental research in the social sciences. In M. Webster and J. Sell (Eds.), *Laboratory experiments in the social sciences* (pp. 57-86). San Diego, CA: Elsevier.
- Trochim, W., (2006). *Research Methods Knowledge Base (2nd. Ed)*, Cincinnati, OH: Atomic Dog Publishing.
- Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning*, 46, 327-369.
- Truscott, J. (1999). What's wrong with oral grammar correction. *Canadian Modern Language Review/La Revue canadienne des langues vivantes*, 55, 437-456.
- Truscott, J. (2007). The effect of error correction on learners' ability to write accurately. *Journal of Second Language Writing*, 16, 255-272.
- Wilcoxon, F. (1945). Individual comparison by ranking methods. *Biometrics*, 1, 80-83.
- Winne, P. H. (1997). Experimenting to bootstrap self-regulated learning. *Journal of Educational Psychology*, 89, 397-410.
- Winne, P. H. (2001). Self-regulated learning viewed from models of information processing. In B. J. Zimmerman & D. H. Schunk (Eds.), *Self-regulated learning*

and academic achievement: Theoretical perspectives (2nd ed., pp. 153-189).

Mahwah, NJ: Lawrence Erlbaum.

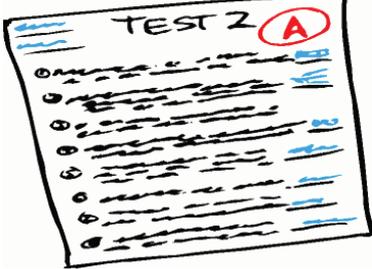
Zamel, V. (1985). Responding to student writing. *TESOL Quarterly*, 19, 79-101.

8. Appendices

Appendix 1: Pre-test, post-test and delayed post-test

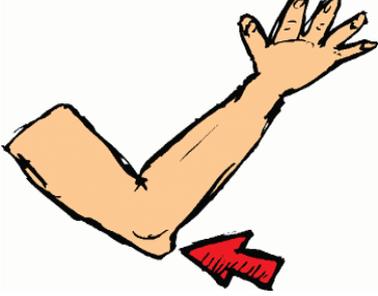
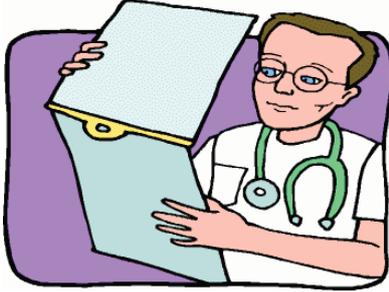
Pre-test

Task: You are to use the pictures and words below to help write your own two paragraph story. The pictures and words are a guide only. You must also use your own words in your story, and can be as imaginative as you wish.

 <p>The first day of school</p>	 <p>Nervous, excited, tired</p>	 <p>Brand new backpack</p>
	 <p>Difficult class work, a lot of writing</p>	 <p>Test results, relieved, proud</p>

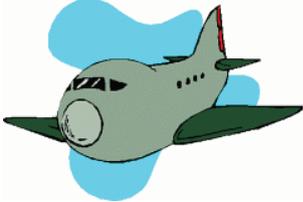
Post-test

Task: You are to use the pictures and words below to help write your own two paragraph story. The pictures and words are a guide only. You must also use your own words in your story, and can be as imaginative as you wish.

 <p>School sports tournament</p>	 <p>Goal, score</p>	 <p>Injury, elbow</p>
 <p>Hospital, doctor</p>	 <p>Broken arm</p>	 <p>Home to bed, recover</p>

Delayed post-test

Task: You are to use the pictures and words below to help write your own two paragraph story. The pictures and words are a guide only. You must also use your own words in your story, and can be as imaginative as you wish.

 <p>Holiday, plane, flying</p>	 <p>Flight attendant, safety procedures, scared</p>	 <p>Hotel, tired, excited</p>
 <p>Shopping</p>	 <p>Restaurant, hungry</p>	 <p>Home again, family</p>

Appendix 2: An example of student writing in the delayed post-test

Rara

This ~~coming~~ holiday Hbliday is so fun. I had ^{her} ~~my~~ holiday in Korea. ^{Her} My family and I went to Korea by plane. It was ^{the} ~~my~~ first time I rode a ~~at~~ plane. It was so fun but I almost throw up. ~~the~~ After we sat, ~~the~~ flight attendant told about the safety procedures and told us not to be panic ~~when~~ if the plane vibrate.

After we arrived at Korea, we went to hotel because we were ~~tried~~ exhausted. Rara and her family wanted to go to the mall to shopping and ~~try~~ tried to eat at the restaurant. I didn't wanted to come because I'm tired. After a few hours, Rara and her family came back. Rara told me that she ate kimchi. I felt jealous because I wanted to eat it too. The next week, we went home by plane again. This time, I didn't thraw up. I* really miss my family.

Appendix 3: The first of the eight Mousedeer narratives used in the written CF sessions

Kancil and the Farmer

One day, Kancil was trying to steal cucumbers from a farmer's field. The first time he was successful at stealing some cucumbers, but the Farmer got angry and put a scarecrow in the field. Kancil made fun of the scarecrow and said that it could not scare him. He kicked the scarecrow with his front leg, and the leg stuck to the scarecrow. The farmer had put glue on the scarecrow and although Kancil tried to escape, he was stuck.

When the farmer found Kancil, he laughed at him and put him in a cage to cook the next morning. When the farmer's dog came along, Kancil pretended to be very happy. The dog was confused and asked him why he was so happy about being locked in a cage and being eaten in the morning. The clever Kancil replied that the dog was wrong. Kancil was going to be the guest of honour at the feast in the morning. Dog was very upset, as he thought he should be the farmer's guest of honour. Kancil agreed and suggested that the dog exchange places with him. The dog agreed. He let Kancil out of the cage and got in the cage himself expecting that he would be guest of honour at the morning's feast. Clever Kancil ran away, and when the farmer came in the morning all he found in his cage was his dog.

Appendix 4: Grammatical error tally sheet

Group	Articles	Copula 'be'	Regular past tense	Irregular past tense	Prepositions

Appendix 5: Graphs displaying linguistic scores for both the focused and unfocused groups which are not normally distributed

Figure 1

Focused group's use of articles in the pre-test

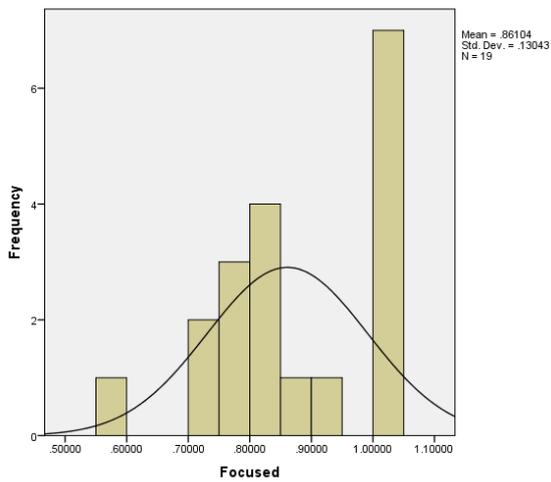


Figure 3

Unfocused group's use of articles in the post-test

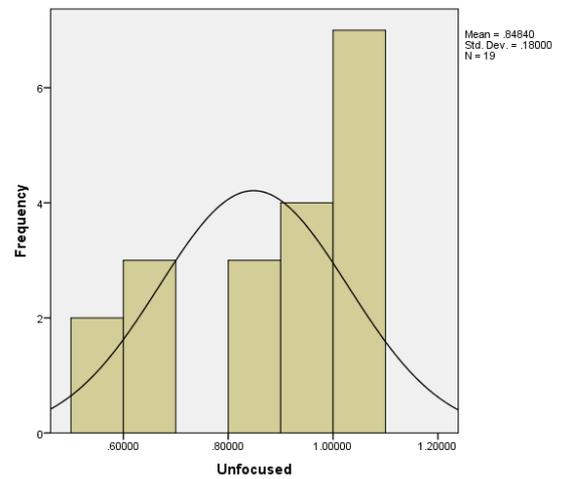


Figure 2

Unfocused group's use of articles in the pre-test

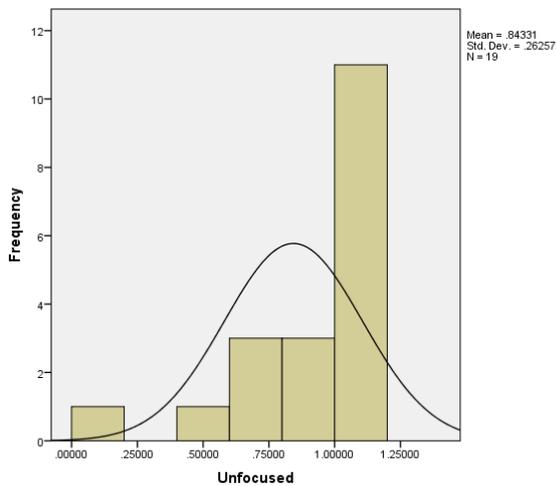


Figure 4

Unfocused group's use of articles in the delayed post-test

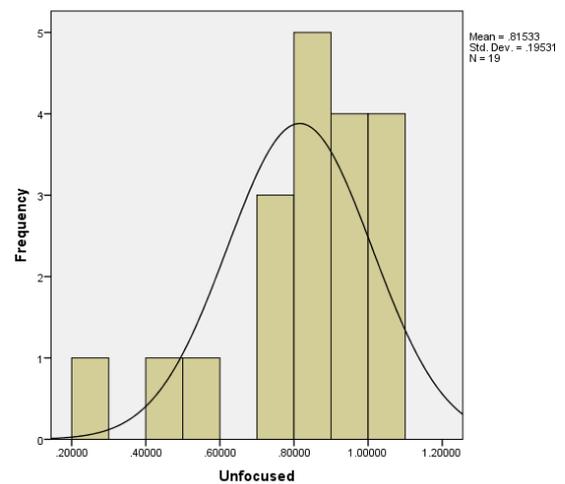


Figure 5

Focused group's use of the copula 'be' in the pre-test

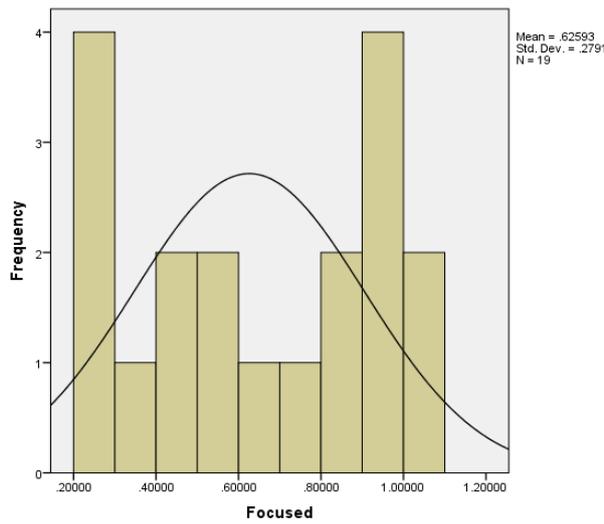


Figure 7

Focused group's use of the regular past tense in the post-test

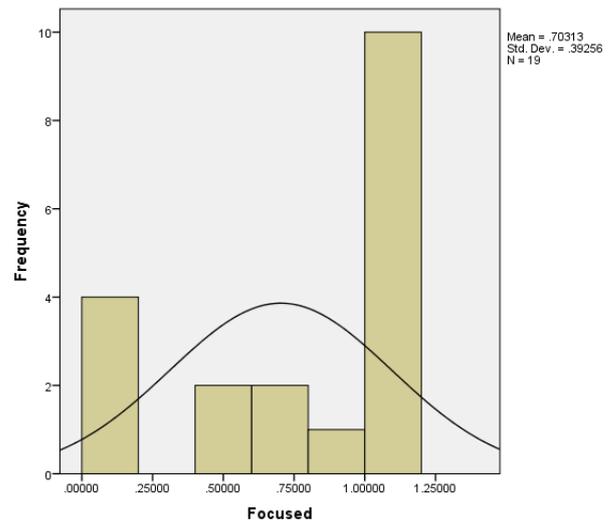


Figure 6

Unfocused group's use of the regular past tense in the pre-test

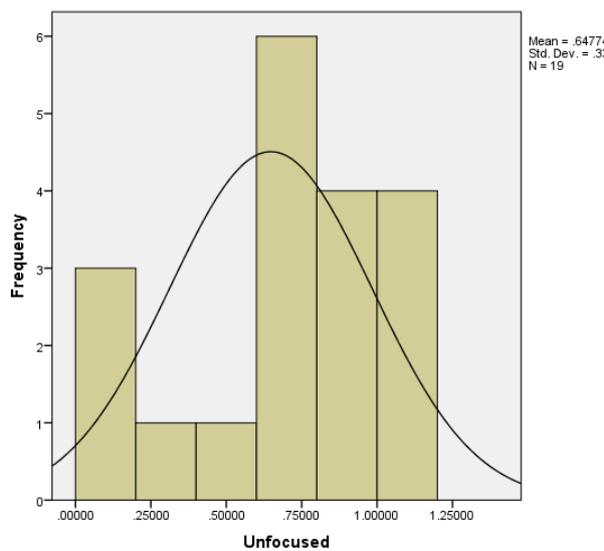


Figure 8

Focused group's use of the regular past tense in the delayed post-test

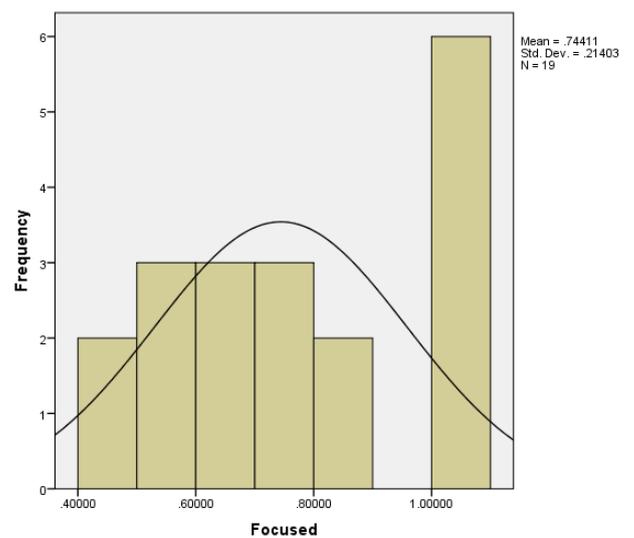


Figure 9

Focused group's use of the irregular past in the pre-test

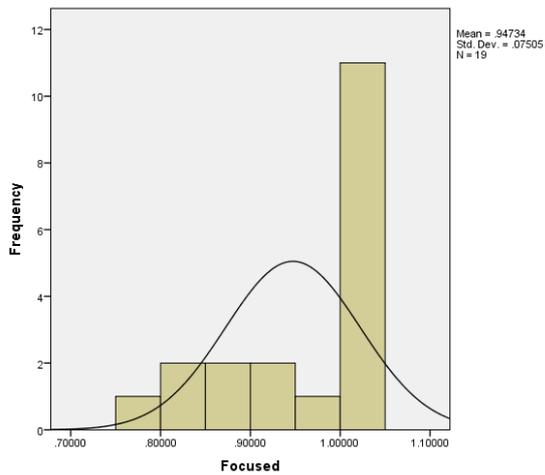


Figure 11

Focused group's use of the irregular past tense in the delayed post-test

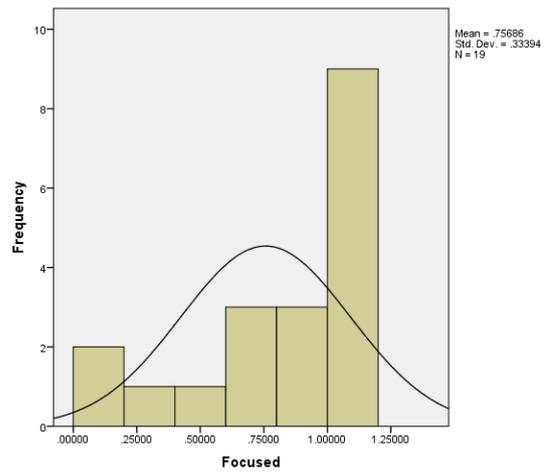


Figure 10

Unfocused group's use of the irregular past tense in the pre-test

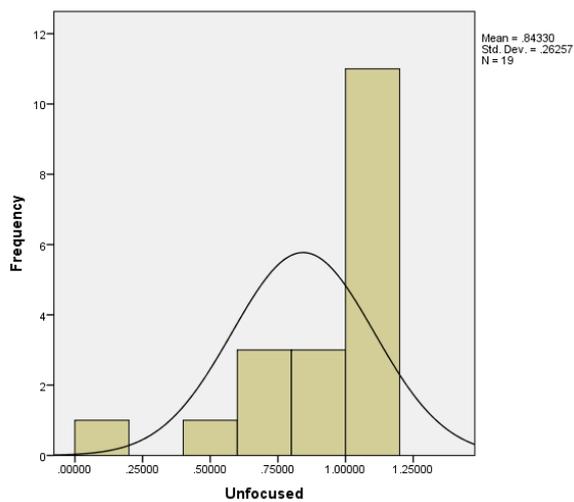


Figure 12

Focused group's use of prepositions in the pre-test

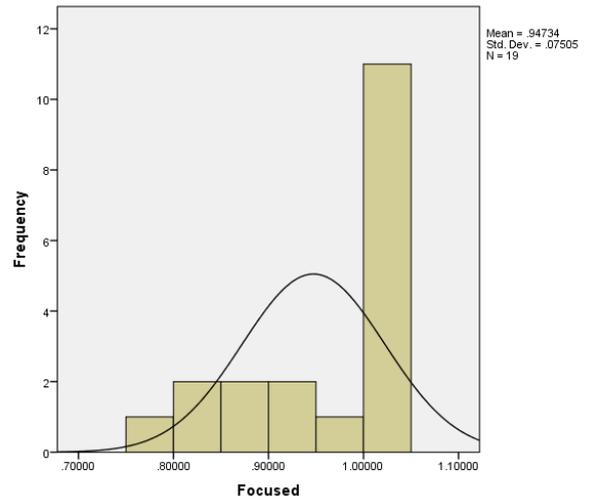


Figure 13

Unfocused group's use of prepositions in the pre-test

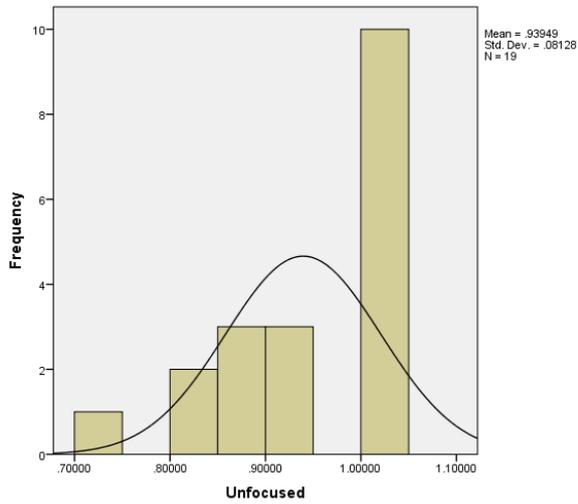


Figure 15

Focused group's use of prepositions in the delayed-post test

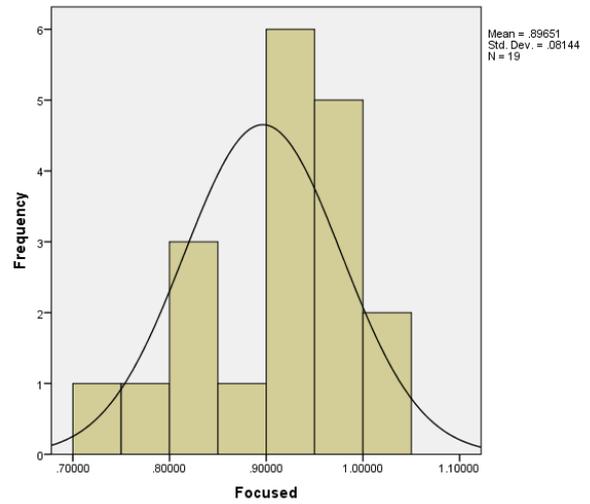


Figure 14

Unfocused group's use of prepositions in the post-test

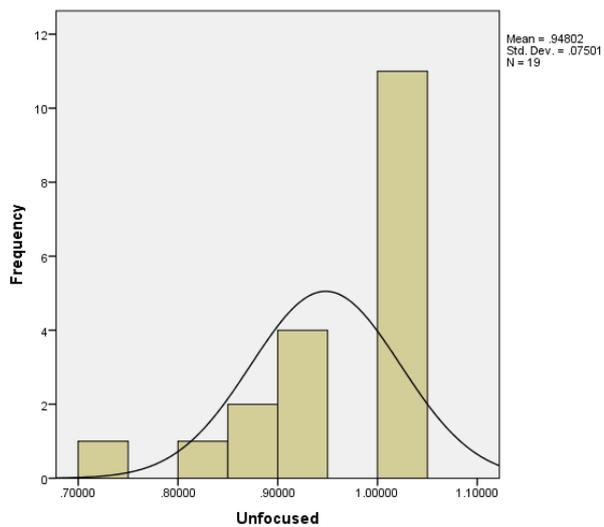


Figure 16

Unfocused group's use of prepositions in the delayed post-test

