



UNIVERSITI PUTRA MALAYSIA

***EFFECTS OF MODEL ESSAYS ON THE IMPROVEMENT OF IRANIAN
IELTS CANDIDATES' WRITING ABILITY***

MOHAMADREZA JAFARY

FPP 2014 13



**EFFECTS OF MODEL ESSAYS ON THE IMPROVEMENT OF IRANIAN
IELTS CANDIDATES' WRITING ABILITY**

By

MOHAMADREZA JAFARY

**Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia,
in Fulfillment of the Requirement for the Degree of Doctor of Philosophy**

June 2014

COPYRIGHT

All material contained within the thesis, including without limitation text, logos, icons, photographs and all other artwork, is copyright material of Universiti Putra Malaysia unless otherwise stated. Use may be made of any material contained within the thesis for non-commercial purposes from the copyright holder. Commercial use of material may only be made with the express, prior, written permission of Universiti Putra Malaysia.

Copyright © Universiti Putra Malaysia



DEDICATION

I Whole Heartedly Dedicate This Thesis To My Beloved Parents And Dear Supervisor For Their Incredible Endorsement And Care. I Also Wish To Express My Gratitude To Dear Yasi For Her Endless Inspiration And Support.



Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment
of the requirement for the degree of Doctor of Philosophy

**EFFECTS OF MODEL ESSAYS ON THE IMPROVEMENT OF IRANIAN
IELTS CANDIDATES' WRITING ABILITY**

By

MOHAMADREZA JAFARY

June 2014

Chairman: Nooreen Noordin, PhD

Faculty : Educational Studies

There are a number of studies on the efficacy of the common practice of most writing teachers in providing feedback to their students each having their own limitations or flaws. The present study, however, seeking to find some evidence on the effectiveness of providing students with corrective feedback moved one step forward by using a quasi-experimental approach and a qualitative study to provide a better understanding of how the model essay approach is more effective than reformulation as common tradition of writing instruction. Having been pretested for their English knowledge and their writing ability, 60 Iranian students were assigned into treatment groups (30) and control groups (30) in two different institutes in Tehran. The treatment group was provided with 10 model essays written by native speakers of English and were instructed to use them as corrective feedback. On the last session all participants were post-tested for their writing ability on IELTS academic writing Task 1 and Task 2 (descriptive and argumentative essays) but the treatment group was also interviewed for their experience on the analysis of the model essays and the prominent components which attracted them most when comparing their own works with the model essays they were provided with. In addition to this, their produced LRES (Language related episodes) were counted on four major categories defined by the IELTS academic writing rubric and their level of noticing was operationalized based on the counted LREs and text analysis. The control group had the same hours of instruction practicing how to write first the separate parts of and then whole essays paying attention to the main components identified by IELTS task 1 and 2 writing band descriptor ('task response,' 'coherence and cohesion,' 'lexical resource,' and 'grammatical range and accuracy'). Instead of model essays participants in the control group used reformulated text as a written feedback. The pre- and post-test writings of all participants were typewritten observing all the misspelling, punctuation errors, grammatical deviations, and use and usage problems in the original hand written copies before being given to 2 independent raters who scored them once for a total score and a second time to determine the sub-scores for the four essay components. The gain scores of the two groups were compared using a two way repeated measure ANOVA followed by post

hoc test (Bonferroni) and the results revealed why such patterns in the data and the statistical analysis were observed for the five scores. The treatment group was found significantly better for their total scores and for their sub-scores on the two components of task response, and lexical resources. The result of multi- mediator analysis indicated that the treatment positively affects all 4 sub-scales of IELTS writing rubric but only grammar and lexical resources significantly influenced the IELTS score. This, interestingly, matched with their interviews and produced LREs in the qualitative section of the research. Using model essays as a form of feedback in writing instruction actually improves learners' writing ability in both IELT academic writing task 1 and 2.



Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Doktor Falsafah

**KESAN ESEI MODEL MENGENAI MEMPERBAIKI PENULISAN
KEBOLEHAN IRAN IELTS CALON'**

Oleh

MOHAMADREZA JAFARY

Jun 2014

Pengerusi: Nooreen Noordin, PhD

Fakulti : Pengajian Pendidikan

Terdapat beberapa kajian mengenai keberkesanan amalan biasa penulisan guru yang paling banyak memberikan tindak balas kepada pelajar-pelajar mereka. Masing-masing mempunyai kelemahan atau kekurangan mereka sendiri. Kajian ini, bagaimanapun, ingin menemukan beberapa bukti tentang keberkesanan menyediakan pelajar dengan maklum balas pembedaan bergerak satu langkah ke hadapan dengan mencuba untuk menggunakan pendekatan yang lain untuk memberi pemahaman yang lebih baik bagaimana pendekatan esei sampel adalah lebih berkesan daripada tradisi biasa dengan menulis arahan. Setelah mengkaji tahap pengetahuan bahasa Inggeris mereka dan keupayaan penulisan mereka, 62 pelajar Iran telah diberikan kepada kumpulan-kumpulan pemulihan seramai 33 orang dan kumpulan kawalan seramai 29 orang dalam dua institusi yang berbeza di Tehran. Kumpulan pemulihan telah disediakan dengan 10 contoh esei yang telah ditulis oleh jurucakap berpendidikan asal bahasa Inggeris dan telah diarahkan untuk menggunakan mereka sebagai maklum balas pembedaan. Pada sesi terakhir selepas semua peserta diuji terhadap tahap keupayaan menulis, kumpulan pemulihan juga telah ditemuramah mengenai analisis contoh esei dan komponen utama yang paling menarik minat mereka apabila membandingkan penulisan mereka sendiri dengan contoh esei yang telah disediakan. Kumpulan kawalan mempunyai waktu pembelajaran yang sama bagaimana untuk menulis bahagian pertama dan kemudian esei keseluruhannya memberi perhatian kepada komponen-komponen utama yang dikenal pasti oleh IELTS tugas 2 bertulis band huraian ('sambutan tugas', 'kesepaduan dan perpaduan', 'leksikal sumber, 'dan' pelbagai tatabahasa dan ketepatan'). Pra-dan pasca-ujian tulisan semua peserta telah ditaip memerhatikan semua misspelling, kesilapan tanda baca, kesalahan tatabahasa, dan penggunaan dan masalah penggunaan dalam salinan tulisan tangan asal sebelum diberi kepada 2 penilai bebas yang memberi markah mereka sekali untuk markah keseluruhan dan untuk kali kedua untuk menentukan sub-markah bagi empat komponen esei. Keuntungan skor kedua-dua kumpulan telah dibandingkan dengan menggunakan pelbagai bebas t-ujian untuk lima markah. Kumpulan pemulihan didapati jauh lebih baik bagi jumlah markah mereka dan sub-markah mereka di kedua-dua komponen tindak balas tugas, dan kepaduan dan perpaduan. Ini, menarik, dipadankan dengan wawancara mereka yang mencerminkan komponen yang sama

seperti tempat tumpuan mereka. Menggunakan contoh esei sebagai satu bentuk tindak balas secara bertulis dapat meningkatkan keupayaan penulisan pelajar.



ACKNOWLEDGEMENTS

First and for most, I would like to extend my deepest praise to The Lord Almighty who has given me the patience, strength, purpose and courage to complete this project by his mercies. My deepest gratitude, goes to the following individuals who have provided contributions; my respectable supervisor Dr. Nooreen Noordin for her guidance, knowledge, suggestions and strong support during my studies. I am highly indebted and extremely grateful. I would like to extend my sincere appreciation to my supervision committee members, Assoc. Prof. Dr. Arshad Bin Samad and Dr. Roselan Baki for their Academic endorsement and graciously agreeing to be on my committee which was my exquisite pleasure. I also express appreciation to all my professors and lecturers during various courses just to name a few, Dr. Kamariah, Dr. Sahandri ,Dr. Ghazali and all my dear lecturers at Faculty of Educational Studies, UPM.



This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfillment of the requirement for the degree of Doctor of Philosophy. The members of the Supervisory Committee were as follows:

Senior. Lecturer. Noreen Noordin, PhD

Faculty of Educational Studies
Universiti Putra Malaysia
(Chairman)

Associate. Prof. Arshad Abd. Samad, PhD

Educational Studies
Universiti Putra Malaysia
(member)

Associate. Prof. Roslan Baki, PhD

Educational Studies
Universiti Putra Malaysia
(member)



BUJANG BIN KIM HUAT, PhD.

Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:

Declaration by graduate student

I hereby confirm that:

- this thesis is my original work;
- quotations, illustrations and citations have been duly referenced;
- this thesis has not been submitted previously or concurrently for any other degree at any other institutions;
- intellectual property from the thesis and copyright of thesis are fully-owned by Universiti Putra Malaysia, as according to the Universiti Putra Malaysia (Research) Rules 2012;
- written permission must be obtained from supervisor and the office of Deputy Vice-Chancellor (Research and Innovation) before thesis is published (in the form of written, printed or in electronic form) including books, journals, modules, proceedings, popular writings, seminar papers, manuscripts, posters, reports, lecture notes, learning modules or any other materials as stated in the Universiti Putra Malaysia (Research) Rules 2012;
- there is no plagiarism or data falsification/fabrication in the thesis, and scholarly integrity is upheld as according to the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) and the Universiti Putra Malaysia (Research) Rules 2012. The thesis has undergone plagiarism detection software.

Signature: _____ Date: 17 Sept, 2014

Name and Matric No: Mohandreza Jafary GS29066

Declaration by Members of Supervisory Committee

This is to confirm that:

- The research conducted and the writing of this thesis was under our supervision;
- Supervision responsibilities as stated in the university Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2012) are adhered to.

Signature: _____
Name of
Chairman of
Supervisory

Committee: **Senior Lecturer. Dr.
Noreen Noordin**

Signature: _____
Name of
Chairman of
Supervisory

Committee: **Associate. Prof. Dr.
Roselan Baki**

Signature: _____
Name of
Member of
Supervisory

Committee: **Associate. Prof. Dr.
Arshad. Abd. Samad**

TABLE OF CONTENTS

	Page
ABSTRACT	i
ABSTRAK	iii
ACKNOWLEDGEMENTS	v
APPROVAL	vi
DECLARATION	viii
LIST OF FIGURES	xiv
LIST OF TABLES	xvi
CHAPTER	
1 INTRODUCTION	1
1.1 Overview	1
1.2 Statement of the Problem	5
1.3 The Purpose of the Study	7
1.4 Research Questions	7
1.5 Research Hypotheses	7
1.6 Significance of the Study	8
1.7 Limitations of the Study	9
1.8 Definition of Key Terms	9
1.8.1 Model Essays	9
1.8.2 Corrective Feedback	10
1.8.3 Scoring Procedure	11
1.8.4 Noticing, understanding and awareness in SLA	12
1.8.5 Think Aloud Protocol Approach	12
1.8.6 Language Related Episodes	12
2 LITERATURE REVIEWS	13
2.1 Introduction	13
2.2 Theoretical Framework	13
2.2.1 Cognitive Theory and Noticing Hypothesis	13
2.2.2 Long's Interaction Hypothesis	14
2.2.3 Social Constructivist Theory	14
2.3 Conceptual Framework	15
2.4 Corrective Feedback	17
2.4.1 Truscott's Case against Grammar Correction	18
2.4.2 Feedback Types	21

2.4.3	Importance of Giving Feedback	25
2.4.4	Noticing in SLA	25
2.4.5	Noticing in L2 Writing Studies	26
2.4.6	Think-Aloud Protocol Approaches in L2 Writing Studies	27
2.4.7	Model Essay as a Feedback Tool	27
2.4.8	Two Types of Written Feedback: Reformulation and Models	28
2.4.9	What does Improvement in Writing Mean?	32
2.4.10	Teacher's Role in Responding to Students	33
2.4.11	Students' Reaction to Teachers' Comments	33
2.5	Writing Instruction History	34
2.5.1	Controlled Composition	34
2.5.2	Current-traditional Rhetoric or Product Approach	35
2.5.3	The Process Approach	35
2.5.4	English for Academic Purposes	36
2.6	The Status of the Present Study	36
2.7	Writing Assessment	37
2.7.1	Direct vs. Indirect Tests of Writing	37
2.7.2	Scales Differentiation	38
2.7.3	The Scoring Procedure	39
2.7.4	Reliability	43
2.8	Summary	46
3	METHODOLOGY	48
3.1	Introduction	48
3.2	Research Design	48
3.3	Sample Size	50
3.4	Sampling Procedure	50
3.5	Attrition	51
3.6	Instrumentation	51
3.6.1	IELTS	52
3.7	Reliability	54
3.7.1	Rater Reliability	55
3.7.2	Measures of Inter-rater Reliability	57
3.8	Validity	58
3.8.1	Internal Validity	59

3.8.2	External Validity	60
3.9	Procedure	60
3.10	Data analysis	62
3.10.1	Definitions of mediation and moderation mediators	63
3.10.2	The Nature of Mediator Variables	64
3.10.3	Interview	65
3.10.4	Text analysis method	65
3.10.5	Qualitative Data Collection Procedure	66
3.11	Summary	67
4	RESULTS AND DISCUSSION	68
4.1	Introduction	68
4.2	Rater Reliability	68
4.3	Demographics	69
4.3.1	Gender	69
4.3.2	Age	69
4.3.3	English Proficiency	70
4.4	Examining the Null Hypotheses	70
4.4.1	Effect of the Interventions on Mean score of Holistic for Task1 and Task2 in Experimental and Control Group	70
4.4.2	Effect of both writing instruction on Mean score of Task Response Score for Task1 and Task2 in Experimental and Control Group	74
4.4.3	Effect of Both Interventions on Mean Score of Coherence for Task1 and Task2 in Experimental and Control Group	78
4.4.4	Effect of Both Interventions on Mean score of Lexical in Experimental and Control Group	82
4.4.5	Effect of Both Interventions on the Mean Score of Grammar for Task1 and Task2 in Experimental and Control Group	86
4.4.6	Effect of Both Interventions on the Mean Score of IELTS in Experimental and Control Group	90
4.4.7	Correlation Between Holistic Score and 4 Sub-scales of IELTS Writing Rubric in Post-test	92
4.4.8	Outcomes of the Interview	94
4.5	Discussion	102
4.5.1	Research Question 1	104
4.5.2	Research Question 2	105
4.5.3	Research Question 3	106

4.5.4	Research Question 4	107
4.5.5	Research Question 5	108
4.5.6	Research Question 6	109
4.5.7	Research Question 7	109
4.6	Summary	112
5	SUMMARY, IMPLICATIONS, AND SUGGESTIONS FOR FURTHER RESEARCH	114
5.1	Summary of the study	114
5.2	Pedagogical Implications	114
5.3	Suggestions for Further Research	115
	REFERENCES	117
	APPENDICES	128
	BIODATA OF STUDENT	355
	LIST OF PUBLICATIONS	356

LIST OF FIGURES

Figure	Page
2-1. Conceptual framework characterising two approaches to essay writing instruction.	16
3-1. Research design.	49
3-2. Mediated relationship among variables (A=predictor; B=mediator; C=criterion/outcome).	63
3-3. Basic mediation chain.	65
4-1. Mean plot of holistic score of Task1 in both experimental (2) and control (1) groups.	72
4-2. Mean plot of holistic score for Task2 in both experimental (2) and control (1) groups	74
4-3. Mean plot of task response score for Task1 in both experimental (2) and control (1) groups	76
4-4. Mean plot of task response score for Task2 in both experimental (2) and control (1) groups	78
4-5. Mean plot of coherence score of Task1 in both experimental and control groups	80
4-6. Mean plot of coherence score for Task2 in both experimental (1) and control (2) groups	82
4-7. Mean plot of lexical score of Task1 in both experimental (2) and control (1) groups	84
4-8. Mean plot of lexical score of Task2 in both experimental (2) and control (1) groups	86
4-9. Mean plot of grammar score of Task1 in both experimental (2) and control (1) groups	88
4-10. Mean plot of grammar score of Task2 in both experimental (2) and control (1) groups	90
4-11. Mean plot of IELTS score in both experimental (1) and control (2) groups	92
4-12. Result of mediation analyses incorporating the model variables	94
4-13. Frequency of each LRE between two groups for Task1	99
4-14. Frequency of LREs at two level of awareness for both low and high proficiency in Task1	100

4-15. Frequency of each LRE between two groups for Task2	101
4-16. Frequency of LREs at two level of awareness for both low and high proficiency in Task2	102



© COPYRIGHT UPM

LIST OF TABLES

Table	Page
2-1. Features of Direct and Indirect Assessment	38
2-2. Advantages and Disadvantages of Analytic Scoring.	41
2-3. A Comparison of Holistic and Analytic Scales in Terms of Six Qualities of Test Usefulness (Weigle, 2002, p. 121).	42
3-1. The studied groups and the number of participants.	50
4-1. The Rater Reliability Indices for Rater 1 and 2.	69
4-2. Checking gender preexisting differences	69
4-3. Comparing the age differences between the control and experimental groups.	69
4-4. Comparison of English proficiency differences between the control and experimental groups.	70
4-5. Descriptive statistic (Mean, SD) of holistic score for Task1	71
4-6. Holistic mean difference between Experimental and Control Groups in pre, post and delayed posttest for Task1.	71
4-7. The difference of Task 1 holistic mean scores between tests in Experimental and Control Groups for Task1.	72
4-8. Descriptive statistic (Mean, SD) of holistic score of Task2.	73
4-9. Holistic mean difference between Experimental and Control Groups in pre, post and delayed posttest for Task2.	73
4-10. Holistic mean score difference between the tests in Experimental and Control Group for Task2.	73
4-11. Descriptive statistic (Mean, SD) of task response score inTask1	75
4-12. Task response mean difference between Experimental and Control Groups in pre, post and delayed posttest for Task1	75
4-13. Holistic mean score difference between tests in Experimental and Control Group for Task1	75
4-14. Descriptive statistic (Mean, SD) for task response score in Task2	76
4-15. Task response mean difference between Experimental and Control Groups in pre, post and delayed posttest for Task2	77

4-16. Holistic mean score difference between tests in Experimental and Control Group for Task2	77
4-17. Descriptive statistic (Mean, SD) of coherence score of Task1	79
4-18. Coherence means difference between Experimental and Control Groups in pre, post and delayed posttest for Task1	79
4-19. Coherence means score difference between tests in Experimental and Control Group for Task1	79
4-20. Descriptive statistic (Mean, SD) of coherence score of Task2	80
4-21. Coherence means difference between Experimental and Control Groups in pre, post and delayed posttest for Task2	81
4-22. Coherence means score difference between tests in Experimental and Control Group for Task2	81
4-23. Lexical mean difference between Experimental and Control Groups in pre, post and delayed posttest for Task1	83
4-24. Lexical score mean difference between tests in Experimental and Control Group for Task1	83
4-25. Descriptive statistic (Mean, SD) of lexical score of Task2	84
4-26. Lexical mean difference between Experimental and Control Groups in pre, post and delayed posttest for Task2	85
4-27. Lexical score mean difference between tests in Experimental and Control Group for Task2	85
4-28. Descriptive statistic (Mean, SD) of grammar scores of Task1	86
4-29. Grammar means difference between Experimental and Control Groups in pre, post and delayed posttest for Task1	87
4-30. Grammar means score difference between tests in Experimental and Control Group for Task1	87
4-31. Descriptive statistic (Mean, SD) of grammar score for Task2.	88
4-32. Grammar means difference between Experimental and Control Groups in pre, post and delayed posttest for Task2	89
4-33. Grammar means score difference between tests in Experimental and Control Group for Task2	89
4-34. Descriptive statistic (Mean, SD) of IELTS scores	90
4-35. IELTS means scores difference between Experimental and Control Groups in pre, post and delayed posttest	91

4-36. IELTS score difference between tests in Experimental and Control Group	91
4-37. Correlation between holistic score and four sub-scales of IELTS writing rubric.	92
4-38. Result of mediation analyses incorporating the model variables (post-test).	93
4-39. Descriptive statistics of LERs of high and low proficiency groups for Task1	98
4-40. Frequency of LREs at two level of awareness for both low and high proficiency in Task1	99
4-41. Descriptive statistics of LERs of high and low proficiency groups for Task2	101
4-42. Frequency of LREs at two level of awareness for both low and high proficiency in Task2	102

CHAPTER ONE

INTRODUCTION

1.1 Overview

The history of ESL composition teaching can be viewed as a succession of methods and approaches to L2 writing, “a cycle in which particular approaches achieve dominance and then fade, but never really disappear” (Silva, 1990, p. 11). The four most influential approaches since 1945 have been the current-traditional rhetoric, controlled composition, English for academic purposes and the process approach.

Controlled composition or what is sometimes referred to as ‘guided composition’ seems to be rooted in Charles Fries’ oral approach, the originator of the audio-lingual method of second language teaching. Notions underlying the approach include the claims that language is primarily speech, rooted in structural linguistics, and learning is considered as habit formation, taken from behaviorist psychology. As such, writing is of secondary concern as reinforcement for oral habits (Silva, 1990).

Writing in this approach acts like a handmaid of other skills and in no way can it take precedence or priority as a major skill to be developed (Rivers, 1981), and it must be “considered as a service activity rather than as an end in itself”. The reader to such pieces of work is the ESL teacher playing the role of a proofreader or an editor not necessarily interested in the quality of ideas expressed but concerned with formal linguistic features (Silva, 1990).

A reaction to controlled composition approach in the mid-sixties came in the form of an almost new approach called current-traditional rhetoric. It was in response to students’ needs for producing extended written discourse. It was suggested that controlled composition was not enough and there was more to writing than making grammatical sentences. This approach was proposed to fill in such a gap. Current-traditional rhetoric’s primary focus was on how to make larger stretches of discourse especially paragraphs by paying a lot of attention to contrastive rhetoric theory. As a result, more pattern drills at the rhetorical level rather than at the syntactic level was called for. In this approach, attention was given not only to elements of a paragraph (topic sentence, supporting sentences, conclusion sentences, and transitions), but also to different options for its development (illustration, exemplification, comparison, contrast, etc.). Essay development was also of major focus.

The introduction of the process approach was a reaction to the shortcomings of the previous two methods. Many believed that none of them could foster thought or its expression since controlled approach was totally irrelevant to such a goal, and the linearity and prescriptivism of the current-traditional rhetoric discouraged creative thinking and writing. As such, the composing process was seen as a “non-linear, exploratory, and generative process whereby writers discover and reformulate their ideas as they attempt to approximate meaning” (Zamel, 1983). In classroom context, this approach entails providing a positive, encouraging and collaborative environment in which learners, with the ample time and little interference they receive, can work through their composing process. Unlike previous approaches, the teacher’s role is to

help students develop viable strategies for getting started, drafting, revising, and editing.

Though the process approach may appear as a perfect method to some, it has faced many criticism. One major portion of that comes from the fact that this approach does not adequately address some central issues in ESL writing. The process approach ignores the variations in writing processes due to differences in individuals, writing tasks, and situations; language proficiency; level of cognitive development; the development of schemata for academic discourse; and the insights from the study of contrastive rhetoric (Reid, 1984).

Another question raised about this approach is whether it realistically prepares learners for academic purposes. It creates a class atmosphere which bears little resemblance to the situations in which they will have to function in future. It also ignores certain types of key academic writing tasks such as essay exams (Horowitz & Daniel, 1986a). He believes that this approach gives students a false impression of how university writing will be evaluated. In other words, the process approach operates in a sociocultural vacuum.

An alternative to process approach with primary emphasis on academic discourse genres and nature of academic writing tasks is the EAP approach. It aims at helping to socialize the students into the academic context and ensuring that learners fall within a range appropriate and acceptable writing behaviors which are approved and dictated by the academic community (Horowitz & Daniel, 1986b). Its instructional methodology is said to seek to recreate the conditions under which actual university writing tasks are carried out.

Whatever approach used by teachers to teach writing in a writing class, providing students with the right form of corrective feedback in responding to the pieces of writing they produce is an indispensable part of the course. However, providing students with written corrective feedback has always been controversial. While some claim that it is ineffective and even harmful (Truscott, 1996, 1999, 2007), others (Ferris, 1999) argue for the practice.

Knoblauch and Brannon (Knoblauch & Brannon, 1981) and Hillocks (1986) reviewing lots of research findings of different sorts suggest that teacher correction has little effect on students' writings. He also states that "the available research suggests that teaching by written comments on compositions is generally ineffective.

Truscott (Truscott, 1996) reviewing a number of research concluded that grammar correction in L2 writing classes should be abandoned. He gave a number of reasons for that: (a) Research evidence shows that grammar correction is ineffective; (b) this lack of effectiveness is exactly what should be expected, given the nature of the correction process and the nature of language learning; (c) grammar correction has significant harmful effects; and (d) the various arguments offered for continuing it all lack merits.

One type of reasoning he presents encompasses the practical problems with error correction. First of all, he argues that many teachers may fail to notice errors (A. D. Cohen & Cavalcanti, 1990). Even if they do recognize an error, they still may not have an idea of the correct use of the grammatical point. They may know an error has

occurred, but they might not know exactly why it is an error. As such they may not be able to explain the principle to the students. Even if teachers give good explanations, still students may not understand them or even if they do, they may not be motivated enough to take care of that. Finally even the “students who do try to write in accordance with the corrections they receive may not maintain their motivation to do so for long” (Cohen & Cavalcanti, 1990, p. 352).

However, Truscott’s claim has not been without criticism. Ferris (1999) was the first to challenge his ideas. The first criticism given by Ferris relates to Truscott’s definition of ‘error correction’. She believes that Truscott has defined error correction in the vaguest of terms. Truscott (1996) defines it as “correction of grammatical errors for the purpose of improving a student’s ability to write accurately.” He also states that “correction comes in many different forms, but for present purposes such distinctions have little significance.” Ferris disagrees with him in that she believes there is a lot of research evidence that “effective error correction, which is selective, prioritized, and clear, can and does help at least some students” As such in considering if grammar correction is ‘effective,’ it is crucially important to know what sort of error correction we are discussing (Ferris, 2004). The second main criticism offered by her is that Truscott in his review of the literature on the effectiveness of grammar correction has under- or over-stated the findings and claims of the original studies to fit his own argument.

The corrective feedback that students may be provided with may come in different forms. These different types have received a lot of attention in the past decades and each study carried out in this field dealing with feedback, has in one way or another included and dealt with one especial type or combination of a limited number of them.

The two types of feedback that have received much attention from researchers interested in the issue of corrective feedback in writing are ‘direct’ and ‘indirect’ corrective feedback. ‘Direct’ is defined as the provision of the correct linguistic form or structure by the teachers above the students’ linguistic error (Ferris, 2003). It may include the crossing out of an unnecessary word or phrase, or the insertion of the missing word.

On the other hand, indirect corrective feedback is the one in which, in some way or another, is indicated that an error has occurred without explicit attention drawn (Ferris, 2003). It may involve one of the following forms: underlining or circling the error; recording in the margin the number of errors in a given line; or using a code to show where the error has occurred and what type of error it is. The teacher does not correct the error, but it is the students’ task to do the job.

Another distinction usually made regarding corrective feedback in language writing courses is the one between ‘selective’ and ‘comprehensive’ corrective feedback. In comprehensive approach (the most common practice) teachers correct all (or at least a range of) the errors in learners’ written work. This approach is also sometimes called the ‘unfocused’ approach. On the other hand, in selective or focused approach, specific errors are selected to be corrected or dealt with, and other errors are ignored. A highly focused corrective feedback will focus on a single type of error (Ellis, 2009a; Ellis, Sheen, Murakami, & Takashima, 2008).

Another issue about which there is not so much evidence is the level of explicitness required for the error feedback provided for students. “Specifically, when teachers mark student errors, do they need to indicate the type of error the student has made, or is it adequate for the teacher to simply underline or circle an erroneous form, leaving it to the student to diagnose and correct the problem?” (Ferris & Roberts, 2001, p. 162).

Teaching a skill cannot be separated from assessing that skill. However, when it comes to writing, unlike other skills, subjectivity on the part of the assessors reveals itself in a more vivid manner. There have been different approaches to the assessment of written compositions. Any scale used in the task of assessing students’ performance in writing, implicitly or explicitly, represents the theoretical basis upon which the test is founded (McNamara & Candlin, 1996). Using a scoring scale entails having some criteria for the evaluation of the written products against which the raters have to assess the compositions. In other words, they should not judge a learner’s performance against those of others. In this manner assessment is based on scoring rubrics of criterion-reference nature rather than norm-reference. Students’ essays are usually rated on the basis of a scoring rubric which can be of different types: holistic, analytic or trait-based.

“Holistic scoring is based on the view that there are inherent qualities of written text which are greater than the sum of the text’s countable elements”(Hamp-Lyons, 1990). Based on the overall impression of a script, raters assign a single score to the text in this type of scoring. It is more practical and less expensive than analytic scoring (Jean Chandler, 2003). It is also intended to focus reader’s attention on the strength of the writing rather than the deficiencies (White, 1985). However, it is of little diagnostic value and the scores obtained are sometimes difficult to interpret (Jean Chandler, 2003).

In analytic scoring, instead of giving a single score to the whole text, each aspect of writing, i.e., vocabulary, grammar, organization, cohesion and so forth is given a separate score. As a result, a more detailed picture of one’s writing ability and performance is obtained in analytic scoring. It provides more diagnostic information about learners’ writing ability than holistic scoring. Moreover, it is more useful in rater training as inexperienced raters understand and apply the criteria more easily in separate scales (Weir, 1990). Besides, since multiple scores are given to a single script, it is more reliable than holistic scoring (Liz Hamp-Lyons, 1991; B. Huot, 1996). However, it is not practical and a good deal of information will be lost when a composite score is required (Jean Chandler, 2003).

Trait-based approach differs from other scoring methods in that it is context-sensitive. It is designed to clearly define the specific topic and genre features of a task being assessed (Jean Chandler, 2003). Trait-based approach is further divided into primary-trait scoring and multiple-trait scoring. Hyland (1986) states that “primary trait scoring represents a sharpening and narrowing of criteria intended for holistic scoring as it involves rating a piece of writing by just one feature relevant to that task” Since a separate scoring guide is needed to be developed for each task, primary trait scoring is very time consuming. Also Hyland (2003a, p. 230) states that “Multiple-trait scoring is often regarded as an ideal compromise by teachers as it requires raters to provide separate scores for different writing features, as in analytic scoring, while ensuring that these are relevant to the specific assessment task”.

In most approaches to assessment, stakeholders try to make the task of rating as objective as possible. However, as far as human raters, even trained raters, are involved, the task of judging learners' performance will be more or less subjective. While Inter-rater reliability has been recognized as a significant methodological factor, problems which could result in failed trials (Lee, 2003), or type II error (that is, to falsely accept a null hypothesis when the literature indicates that it should have been rejected) which makes high rater reliability a difficult task to achieve.

Weigle (2002) states that raters are affected as much by their experiences and cultural contexts as by the quality of the students' writings. Raters' background experience may affect their judgments. Raters from different disciplines have been observed to apply different criteria to nonnative English writing samples (Carl James, 1998; C. Knoblauch & Brannon, 1981; Leki, 1990; Leki, 1994). Also, raters familiar with L1 rhetorical conventions tend to be more accepting of L2 essays showing L1 traces in comparison with other raters (Burkland & Grimm, 2010; Radecki & Swales, 1988).

The number of levels in a scale that raters can accurately distinguish is limited. The more levels exist in a scale, the more difficult it becomes for the raters to decide accurately. Penny et al. (2000, p. 147) state that "it seems possible, and, moreover, it seems likely that the length of a scale may affect measurement error, serving to increase the error component of variance when the scale length surpasses the ability of raters to discriminate between levels of proficiency".

Raters have been found to assign higher scores to hand-written essays in comparison with those prepared by word-processors (Hahn, 1981; Leki, 1991; Marshall & Powers, 1969; Powers, Fowles, Farnum, & Ramsey, 1994; Russell & Tao, 2004; Sloan & McGinnis, 1982; Soloff, 1973; Sperling & Freedman, 1987). There are many other factors such as students' gender (Bolger & Kellaghan, 1990; Manke & Loyd, 1990; Natriello & McDill, 1986), their ethnic background (Keith, Reimers, Fehrmann, Pottebaum, & Aubey, 1986), socioeconomic status (Jussim & Eccles, 1992), and behavior (Manke & Loyd, 1990) which have been found to affect raters' rating task.

Presuming that one takes care of all the above-mentioned problems in the task of writing assessment, the question of whether to provide students with corrective feedback in writing courses still remains unresolved. In spite of the controversies present in this regard, many teachers still believe that following such a practice in their writing classes is one of their responsibilities, and insist on continuing to do so. The present study is an attempt to find evidence in support of either continuing or abandoning such a practice. In doing so, the researcher attempts to examine the effect of providing students with model essays as a form of feedback.

1.2 Statement of the Problem

During the last three decades, many studies have been carried out to show the effectiveness or ineffectiveness of the common practice of most writing teachers in providing feedback to their students. However, most of these studies were in one way or another flawed. That is why drawing any conclusion about the effectiveness of such a practice has been, if not impossible, very difficult. On the one hand, Truscott (1996, 1999) argues that providing students with corrective feedback is not only ineffective but also harmful. On the other hand those who argue against Truscott's claim do present results which are obtained from studies most of which are problematic in their

methodology. Many studies showing positive results (Chandler, 2000; Ferris, Chaney, Komura, Roberts, & McKee, 2000; Ferris, 1997; Lalande, 1982) were carried out without having a control group included, which makes drawing any conclusion about the comparison of correction/no correction approaches implausible.

Three studies mostly cited in different reviews are Ashwell (2000), Fathman and Whally (1990) and Ferris and Roberts (2001). These studies have shown positive results for the practice of providing students with CF and they did include a control group. However, having a control group is not sufficient. What is needed is a control group which is compatible with the treatment group in all aspects including proficiency level, writing conditions, and instructional context (Gu nette, 2007). It is a fact that the above mentioned studies, did have a control group, but their experimental and control groups differed in proficiency levels.

Also, most of these studies were of a short duration. For example, Fathman and Whalley's (1990) participants were required to write one essay and had 30 minutes to correct it. The effect of feedback on learners' accuracy in such a short time does not seem to be conclusive at all. Moreover, most of the studies examining the effect of CF do so by requiring learners to work on the same piece of work, i.e., working on different drafts of the same essay. Very few studies have examined the corrective feedback on a new piece of writing.

The present study, like all other studies carried out in this field, seeks to find some evidence on the effectiveness or ineffectiveness of providing students with corrective feedback. However, what makes this study stand out is the use of model essays as a form of feedback and exploring their effectiveness in teaching writing to the students of English as a foreign language. Apart from this, the carried out studies were not in the context of preparation for IELTS focusing on both IELTS Academic writing Task 1 and Task2. Additionally this study uses a comprehensive approach applying quantitative, qualitative and text Analysis based on the verbalized LREs derived from the interview transcriptions. Eventually, the mediating factors were assessed and a delayed post-test was administered substantiating the consistency of the outcomes. Most Iranian IELTS candidates have poor performance in IELTS writing task 1 and 2 as the conventional syntactic method could not yield desirable results (refer to Iran IELTS Official website) and it is evident that most of these candidates need to improve the instrumental use of language to find their way to one of the prime universities around the globe thereby the necessity to apply a more practical and analytic writing instruction to improve their writing proficiency becomes more paramount. In light of positive effects of applying written corrective feedbacks especially model essays as written corrective tool, this area of research needs more empirical studies. This particularly helps to introduce an effective writing instruction approach in the context of preparation for IELTS or other internationally recognized instrumental tests of English language. However, issues concerning how noticing is related to composing and subsequent feedback processing, and what impact such noticing has on L2 writing improvement, need to be addressed. Moreover the present study attempts to avoid the pitfalls of most studies carried out in this regard so far. Therefore, the methodological problems were addressed by including a control group; judging the effectiveness of the practice based on both quantitative and qualitative analysis and including a post test of a new piece of writing . To this aim, the homogeneity of treatment and control groups was taken into account and the time frame of the treatment was extended over

a course of 10 two-hour sessions followed with a delayed post- test to indicate the consistency of the results.

1.3 The Purpose of the Study

The present study as its primary objective attempts to explore the effectiveness of providing students with corrective feedback in teaching writing courses. In doing so, it examines the effectiveness of the use of model essays written by native speakers of English as a form of feedback. More specifically, this study seeks to find out which aspect of IELTS writing scoring rubric including the task response/achievement, coherence and cohesion, lexical resources, and grammatical range and accuracy is more affected by model essay approach through the use of an analytical rubric, the interview, text analysis and the operationalization of their noticing behavior. As its subsequent objective it examines the mediating effect of all four sub-scales of IELTS writing scoring rubric. Eventually, the research was carried out in Iran as one of the general objectives of this study was the application of model essays as written feedback tool in Iranian context.

1.4 Research Questions

The present study seeks to find an answer to each of the following questions:

1. Is there a significant difference in gain scores of IELTS mock writing test between experimental (receiving model essays) and control group in Task 1 and 2?
2. Is there a significant difference in 'task response /achievement' sub-scale of IELTS writing scoring rubric between experimental (receiving model essays) and control group in Task 1 and 2?
3. Is there a significant difference in 'coherence and cohesion' sub-scale of IELTS writing scoring rubric between experimental (receiving model essays) and control group in task 1 and 2?
4. Is there a significant difference in 'grammatical range and accuracy' sub-scale of IELTS writing scoring rubric between experimental (receiving model essays as a form of written corrective feedback) and control group in task 1 and 2?
5. Is there a significant difference in 'lexical resources' sub-scale of IELTS writing scoring rubric between experimental (receiving model essays as a form of written corrective feedback) and control group in task 1 and 2?
6. Does IELTS writing rubric sub-scales have any mediating effect between the model essay approach and holistic gain score of IELTS writing mock test?
7. What is the frequency of LREs (Language related episodes) and their noticing behavior in different sub-scales of IELTS writing rubrics for both academic writing Task 1 and 2 concerning low and high English proficiency levels?

1.5 Research Hypotheses

The following null hypotheses are based on the above research questions:

1. There is no significant difference in learners' gain scores of IELTS mock writing test between experimental (receiving model essays) and the control group for both Task 1 and 2.
2. There is no significant difference in learners' gain score as defined by their 'task response/achievement' sub-scale of the IELTS writing scoring rubric between experimental (receiving model essays) and control group.
3. There is no significant difference in learners' gain score as defined by their 'cohesion and coherence' sub-scale of the IELTS writing scoring rubric between experimental (receiving model essays) and control group.
4. There is no significant difference in learners' gain score as defined by their 'lexical resources' sub-scale of the IELTS writing scoring rubric between experimental (receiving model essays) and control group.
5. There is no significant difference in learners' gain score as defined by their 'grammatical range and accuracy' sub-scale of the IELTS writing scoring rubric between experimental (receiving model essays) and control group.
6. The IELTS writing rubric sub-scales does not significantly mediate the effect of model essay approach on holistic gain score of IELTS writing Mock test.

1.6 Significance of the Study

The first and foremost significance that the result of this study may have, is to help students to improve the writing proficiency in the context of preparation for IELTS or other internationally recognized English exams. It is noted that the main objective of many pupils is the instrumental use of language such as passing IELTS, TOEFL, FCE or any valid language testing system. In this case, improving the band score by 0.5 or 1 mark in IELTS or passing international English tests as language requirement of various universities could be a great achievement. As the academic writing has become more of a challenge for many students around the globe, the necessity to come up with an innovative and effective methodology improving the writing ability in the context of IELTS preparation has become more paramount. This particularly applies to IELTS academic writing module which is one of the needed requirements in most prime universities. The findings of the current study may be applicable in teaching writing to the students of English as a foreign language using model essays as written corrective feedback tool in writing instruction. This analytic approach could potentially improve the writing proficiency which plays an undeniable role within academic excellence and scientific productions in different fields. It is widely believed that using written corrective feedback in writing instruction will have tremendous effect on the development of analytical, more practical and less time consuming approaches. The efficacy of written feedback has become a crucial issue in writing instruction due to the ineffectiveness of most conventional methods in EAP or other academic contexts. This is due partly to the longer time frame and unsatisfactory outcomes derived from the results of the academic writing module in various international English testing systems. What pinpoints the efficacy of such an innovative methodology is its contribution to the development of useful, to the point materials and exercises in a preparatory context for various academic writing courses. It is clear that a systematic unified curriculum on the basis of an intensive and analytical approach yielding

positive outcomes in a shorter time frame would be considered as a great success. Interestingly enough, positive results substantiating the efficacy of model essay approach could potentially cast more light on the subtle and messy area of written corrective feedback and its interpolation in writing instruction. Additionally, only few studies have attempted to directly investigate whether L2 students who receive written corrective feedback or model essays are able to improve the accuracy of their writing. The carried out researches did not compare model essay approach group with the participants who do not receive model essays as a form of written feedback tool. Moreover, the efficacy of using model essays as feedback tool has never been investigated through a comprehensive study. However, the current study is an experimental approach with a time frame of 10 sessions endorsed by in depth qualitative supportive information. The qualitative section of the current study includes the operationalization of noticing by using think aloud protocols and counting produced LREs in all four major IELTS writing sub-scales in both academic Task 1 and 2. Subsequently, an interview based on IELTS writing rubric was carried out to gain more in depth details as well as conducting a full mediation model analyzing the mediating variables in both IELTS academic writing Task 1 and 2. The findings of this study may contribute to the enrichment of model essay approach in various aspects such as preparing course material, curriculum development and tailor made intensive and semi-intensive writing programs.

1.7 Limitations of the Study

Due to practicality issues, the inter- and intra-rater reliability indices for the essay components identified by IELTS writing rubric for the pretest and posttest writing samples were not calculated. Instead their agreement for the overall scores given to the writing samples in the pre and posttest was presumed to be representative of their performance for the four elements.

As mentioned in the previous section, the components of writing examined in the present study were restricted to those identified by IELTS. It was not possible to examine all the elements identified by other different rubrics.

One factor that may limit the generalizability of the results is the fact that all participants were of almost high proficiency. It was impractical to try to examine the existence of such an effect across different proficiency levels. As such, it is not possible to generalize the results to other situations. It is the users' responsibility to examine the context of the study and see if it matches the contexts to which they wish to apply the findings.

1.8 Definition of Key Terms

This study employs several key terms which are conceptually and operationally defined as follows:

1.8.1 Model Essays

Model essays are pieces of sample essays written by native speakers to give students a model of how to approach a topic. Usually, in most textbooks for teaching writing skill, students are provided with such essays for the first few units to gain insight into the points explained in the text (Cameron, 1999). The model essays as written

corrective feedback tools for L2 writing instruction aims to reflect how Iranian EFL pupils notice their language problems by analytically comparing their own writing with a model essay written by the native speaker or an expert of the academic panel in the context of preparation for the IELTS writing test. This terminology is partly replicated from Qi and Lapkin (2001) study of the reformulation method. Model essays in the context of this study are indirect and explicit written feedback.

1.8.2 Corrective Feedback

Corrective Feedback refers to teachers' or peers' reaction or treatment to one's errors by either correcting them directly or simply indicating that an error has occurred, which requires the learners to correct it. Corrective feedback (CF) comes in a variety of forms. It can be direct or indirect, selective or comprehensive, and implicit or explicit (Ellis et al., 2008). It is to be noted that model essays used in this study are written corrective feedback as indirect CF and the type of the essays are descriptive and argumentative in both types of presenting an argument, balanced argument or cause and effect ones tested in IELTS academic writing task 2. It is note worthy that both the model essays and reformulated texts are written corrective feedback in writing instruction context.

1.8.2.1 Direct vs. Indirect CF

'Direct' feedback is defined as the provision of the correct linguistic form or structure by the teachers above the students' linguistic error (D. Ferris, 2003). It may include the crossing out of an unnecessary word or phrase, or the insertion of the missing word. There are also other forms of direct feedback such as written meta-linguistic explanations at the end of students' essays, or even spoken meta-linguistic explanations as in individual conferences between teachers and students (Bitchener & Knoch, 2009). On the other hand, indirect corrective feedback is the one in which, in some way or another, it is indicated that an error has occurred without any explicit attention drawn to it (Ferris, 2003, 2011). It may involve one of the following forms: underlining or circling the error; recording in the margin the number of errors in a given line; or using a code to show where the error has occurred and what type of error it is. The teacher does not correct the error, but it is the students' task to do the job. In the context of the current research the use of model essay approach in treatment group is considered as written indirect feedback, however the methodology of writing instruction in control group was heavily based on conventional syntactic way of writing instruction and the use of reformulated texts of the original drafts of the students.

1.8.2.2 Comprehensive vs. Selective CF

In comprehensive approach (the most common practice) teachers correct all (or at least a range of) the errors in learners' written work. This approach is sometimes called the 'unfocused' approach. On the other hand, in selective or focused approach, specific errors are selected to be corrected or dealt with, and other errors are ignored. A highly focused CF will focus on a single type of error (Erlam, Ellis, & Batstone, 2013). Teachers can elect to correct all of the students' errors, in which case the CF is unfocused. Alternatively they can select specific error types for correction. For example, in the above examples the teacher could have chosen to correct just article errors. The distinction between unfocused and focused CF applies to all of the

previously discussed options. Processing corrections is likely to be more difficult in unfocused CF as the learner is required to attend to a variety of errors and thus is unlikely to be able to reflect much on each error. In this respect, focused CF may prove more effective as the learner is able to examine multiple corrections of a single error and thus obtain the rich evidence they need to both understand why what they wrote was erroneous and to acquire the correct form. If learning is dependent on attention to form, then it is reasonable to assume that the more intensive the attention, the more likely the correction is to lead to learning. Focused metalinguistic CF may be especially helpful in this respect as it promotes not just attention but also understanding of the nature of the error. However, unfocused CF has the advantage of addressing a range of errors, so while it might not be as effective in assisting learners to acquire specific features as focused CF in the short term, it may prove superior in the long run. Model essay approach applies the inherent qualities of selective feedback as the type of topic and learning context directs the approach applied. However, the more syntactic and conventional techniques are considered as comprehensive feedback similar to what was applied within the non treatment group.

1.8.2.3 Implicit vs. Explicit CF

In providing students with corrective feedback, options for teachers range from very explicit feedback such as marking an error at its exact location in the text and labeling it with a code or verbal cue, such as 'VT,' or 'wrong verb tense,' to placing a checkmark in the margin of the paper to let the writer know that there is an error somewhere in that line, but anyhow it is left for the students to find, diagnose and correct the error (James, 1998). Within the implicit written feedback types reformulated text are the most effective ones and model essays are the explicit types applied in the treatment group. Sheen(2010), pointed out that the explicitness of the feedback was a better predictor of the written accuracy. However, within the comparison of implicit feedback, written reformulations were more effective than the other investigated types.

1.8.3 Scoring Procedure

The process of assigning a score to a piece of writing produced by a student is known as scoring procedure. Scoring can be either norm-referenced or criterion-referenced. In the former, each student's performance is compared with those of other students. In the latter, each student's performance is assessed against a predetermined criterion called scoring rubrics. The scoring rubrics can be holistic, analytic, or trait-based. The scoring procedure used in this study is of criterion-reference and both a holistic and analytic type using IELTS scoring guide.

1.8.3.1 Holistic Scoring

A method based on the view that "there are inherent qualities of written text which are greater than the sum of the text's countable elements" (Hamp-Lyons, 1990). In holistic scoring a single score is assigned to a script based on the overall impression of the script. Each writing is read quickly and then judged against a rating scale, or scoring rubric that outlines the scoring criteria (Weigle, 2002).

1.8.3.2 Analytic Scoring

In analytic scoring, instead of a given single score, scripts are rated on several aspects of writing. The scripts may be rated on aspects such as vocabulary, grammar, content, organization, cohesion, register, or mechanics. As a result, analytic scoring provides us with more detailed information about a test taker's performance in different aspects of writing (Weigle, 2002). The fact that for each aspect of writing, a score must be reported helps us ensure that features are not collapsed into one; therefore, analytic scoring provides us with more information than holistic scoring (Hyland, 2003a).

1.8.4 Noticing, understanding and awareness in SLA

Schmidt (1995) defines noticing as “conscious registration of the occurrence of some event” and understanding as “recognition of a general principle, rule or pattern”. In other words, noticing deals with surface level language phenomena, while understanding is related to the learning at a more abstract level.

1.8.5 Think Aloud Protocol Approach

A verbal report type of data collection procedure which provides a rich amount of data reflecting on the subject's mental processes, (Schmidt, 1990) . He maintains that concurrent verbal reports such as think aloud protocols are trust worthy evidence as to whether something has been consciously perceived or noticed.

1.8.6 Language Related Episodes

LREs include —all interaction in which learners draw attention to form, that is, those that focus on form in the context of meaningful interaction as well as those that are set apart from such communication and simply revolve around questions of form itself (Williams, 1999, p. 595). LREs have been used as a unit of analysis in classroom-based studies of interaction and have been shown to occur frequently in classroom contexts (Ellis, 2009b; Ellis, Basturkmen, & Loewen, 2001a, 2001b; Loewen, 2003, 2004; Swain & Lapkin, 2001; Williams, 1999, 2001). Furthermore, LREs may promote the noticing of L2 forms and subsequent learning, so their role in interactions is important.

REFERENCES

- Airasian, P. W. (1996). *Assessment in the classroom*: McGraw-Hill New York.
- Aljaafreh, A., & Lantolf, J. P. (1994). Negative feedback as regulation and second language learning in the zone of proximal development. *The modern language journal*, 78(4), 465-483.
- Arnold, V. (1990). Do Students Get Higher Scores on Their Word-Processed Papers? A Study of Bias in Scoring Hand-Written vs. Word-Processed Papers.
- Ashwell, T. (2000). Patterns of Teacher Response to Student Writing in a Multiple-Draft Composition Classroom: Is Content Feedback Followed by Form Feedback the Best Method? *Journal of Second Language Writing*, 9(3), 227-257.
- Astika, G. G. (1993). Analytical assessments of foreign students' writing. *RELC journal*, 24(1), 61-70.
- Bachman, L. F. (1990). *Fundamental considerations in language testing*: Oxford University Press.
- Bachman, L. F. (1996). *Language testing in practice: Designing and developing useful language tests* (Vol. 1): oxford university press.
- Backman, L., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*: Oxford: Oxford University Press.
- Bamberger, M. (2000). *Integrating quantitative and qualitative research in development projects*: World Bank Publications.
- Bitchener, J. (2008). Evidence in support of written corrective feedback. *Journal of Second Language Writing*, 17(2), 102-118.
- Bitchener, J., & Knoch, U. (2008). The value of written corrective feedback for migrant and international students. *Language Teaching Research*, 12(3), 409-431.
- Bitchener, J., & Knoch, U. (2009). The relative effectiveness of different types of direct written corrective feedback. *System*, 37(2), 322-329.
- Bitchener, J., & Knoch, U. (2010). The contribution of written corrective feedback to language development: A ten month investigation. *Applied Linguistics*, 31(2), 193-214.
- Bolger, N., & Kellaghan, T. (1990). Method of measurement and gender differences in scholastic achievement. *Journal of Educational Measurement*, 27(2), 165-174.
- Breland, H. M. (1983). The Direct Assessment of Writing Skill: A Measurement Review. College Board Report No. 83-6.
- Breland, H. M., Bridgeman, B., Fowles, M. E., & Board, G. R. E. (1999). *Writing assessment in admission to higher education: Review and framework*: College Entrance Examination Board New York.
- Breland, H. M., & Jones, R. J. (1988). Remote scoring of essays.
- Briere, E. J. (1966). Quantity before quality in second language composition. *Language Learning*, 16(3-4), 141-151.
- Briggs, D. (1980). A study of the influence of handwriting upon grades using examination scripts. *Educational Review*, 32(2), 186-193.
- Brossell, G. (1986). Current research and unanswered questions in writing assessment. *Writing assessment: Issues and strategies*, 168-182.
- Brown, G. T., Glasswell, K., & Harland, D. (2004). Accuracy in the scoring of writing: Studies of reliability and validity using a New Zealand writing assessment system. *Assessing Writing*, 9(2), 105-121.

- Brown, J. D. (1991). Do English and ESL faculties rate writing samples differently? *Tesol Quarterly*, 25(4), 587-603.
- Bull, R., & Stevens, J. (1979). The effects of attractiveness of writer and penmanship on essay grades. *Journal of Occupational Psychology*, 52(1), 53-59.
- Burkland, J., & Grimm, N. (2010). Motivating through responding. *Journal of Teaching Writing*, 5(2), 237-248.
- Chalhoub-Deville, M., & Turner, C. E. (2000). What to look for in ESL admission tests: Cambridge certificate exams, IELTS, and TOEFL. *System*, 28(4), 523-539.
- Chandler, J. (2000). *The efficacy of error correction for improvement in the accuracy of L2 student writing*. Paper presented at the AAAL, Vancouver, BC.
- Chandler, J. (2003). The efficacy of various kinds of error feedback for improvement in the accuracy and fluency of L2 student writing. *Journal of Second Language Writing*, 12(3), 267-296.
- Chandler, J. (2004). A response to Truscott. *Journal of Second Language Writing*, 13(4), 345-348.
- Chapman, R. S. (1997). Language development in children and adolescents with Down syndrome. *Mental Retardation and Developmental Disabilities Research Reviews*, 3(4), 307-312.
- Chase, C. I. (1968). The impact of some obvious variables on essay test scores. *Journal of Educational Measurement*, 5(4), 315-318.
- Claffey, S. (2000). Teaching ESL Composition, Purpose, Process, and Practice: D. Ferris, and J.S. Hedgcock; Lawrence Erlbaum, London, 1998, 329 pp., £22.50 pb, ISBN 0-8058-2450-2. *System*, 28(2), 323-324.
- Clapham, C. (2000). Assessment for academic purposes: where next? *System*, 28(4), 511-521.
- Coffman, W. E. (1971). On the Reliability of Ratings of Essay Examinations in English. *Research in the Teaching of English*.
- Cohen, A. D., & Cavalcanti, M. C. (1990). Feedback on compositions: Teacher and student verbal reports. *Second language writing: Research insights for the classroom*, 155-177.
- Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20(1).
- Coley, M. (1999). The English language entry requirements of Australian universities for students of non-English speaking background. *Higher Education Research & Development*, 18(1), 7-17.
- Connor, U., & Carrell, P. (1993). The interpretation of tasks by writers and readers in holistically rated direct assessment of writing. *Reading in the composition classroom: Second language perspectives*, 141-160.
- Council, B. (2006). IDP: IELTS Australia and University of Cambridge ESOL Examinations 2005 IELTS Handbook.
- Creswell, J. (2003). *Qualitative, quantitative & Mixed Methods Design*: Thousand Oaks, CA: Sage.
- Creswell, J. (2007). *Qualitative inquiry and research method: Choosing among five approaches*. Sage, Thousand Oaks.
- Creswell, J. W., & Clark, V. L. P. (2007). *Designing and conducting mixed methods research*: Wiley Online Library.
- Cronbach, L., Linn, R., Brennan, R., & Haertel, E. (1995). Generalizability analysis for educational assessments. *Evaluation Comment*, 1-29.

- Crystal, D. (1997). The Functions of Language. *The Cambridge encyclopedia of language*, 10.
- Cumming, A. (1989). Writing Expertise and Second-Language Proficiency*. *Language learning*, 39(1), 81-135.
- Cumming, A. (1990). Expertise in evaluating second language compositions. *Language Testing*, 7(1), 31-51.
- Davies, A. (1999). *Dictionary of language testing* (Vol. 7): Cambridge University Press.
- DeMauro, G. E. (1992). *An investigation of the appropriateness of the TOEFL test as a matching variable to equate TWE topics*: Educational Testing Service.
- Dunne, M., Pryor, J., & Yates, P. (2005). *Becoming a researcher: A research companion for the social sciences*: McGraw-Hill Education.
- Eames, K., & Loewenthal, K. (1990). Effects of handwriting and examiner's expertise on assessment of essays. *The Journal of Social Psychology*, 130(6), 831-833.
- Ediger, A. (2001). Teaching children literacy skills in a second language. *Teaching English as a second or foreign language*, 3, 153-169.
- Elbow, P., & Yancey, K. B. (1994). On the nature of holistic scoring: An inquiry composed on email. *Assessing Writing*, 1(1), 91-107.
- Ellis, N. C. (2002). Frequency effects in language processing. *Studies in second language acquisition*, 24(02), 143-188.
- Ellis, N. C. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in second language acquisition*, 27(02), 305-352.
- Ellis, R. (1994). A theory of instructed second language acquisition. *Implicit and explicit learning of languages*, 79-114.
- Ellis, R. (2009a). Corrective feedback and teacher development. *L2 Journal*, 1(1).
- Ellis, R. (2009b). A typology of written corrective feedback types. *ELT journal*, 63(2), 97-107.
- Ellis, R., Basturkmen, H., & Loewen, S. (2001a). Learner uptake in communicative ESL lessons. *Language learning*, 51(2), 281-318.
- Ellis, R., Basturkmen, H., & Loewen, S. (2001b). Preemptive focus on form in the ESL classroom. *Tesol Quarterly*, 35(3), 407-432.
- Ellis, R., Sheen, Y., Murakami, M., & Takashima, H. (2008). The effects of focused and unfocused written corrective feedback in an English as a foreign language context. *System*, 36(3), 353-371.
- Enginarlar, H. (1993). Student response to teacher feedback in EFL writing. *System*, 21(2), 193-204.
- Erazmus, E. T. (1960). Second language composition teaching at the intermediate level. *Language learning*, 10(1-2), 25-31.
- Erlam, R., Ellis, R., & Batstone, R. (2013). Oral corrective feedback on L2 writing: Two approaches compared. *System*.
- Evans, N. W., James Hartshorn, K., & Strong-Krause, D. (2011). The efficacy of dynamic written corrective feedback for university-matriculated ESL learners. *System*, 39(2), 229-239.
- Fathman, A., & Whalley, E. (1990). Teacher response to student writing: Focus on form versus content. *Second language writing: Research insights for the classroom*, 178-190.
- Fazio, L. L. (2001). The effect of corrections and commentaries on the journal writing accuracy of minority-and majority-language students. *Journal of Second Language Writing*, 10(4), 235-249.

- Ferris, D. (1999). The case for grammar correction in L2 writing classes: A response to Truscott (1996). *Journal of Second Language Writing*, 8(1), 1-11.
- Ferris, D. (2003). *Response to student writing: Implications for second language students*: Routledge.
- Ferris, D. (2011). *Treatment of Error in Second Language Student Writing, ~ autofilled~*: University of Michigan Press.
- Ferris, D., Chaney, S., Komura, K., Roberts, B., & McKee, S. (2000). *Perspectives, problems, and practices in treating written error*. Paper presented at the Colloquium presented at International TESOL Convention, Vancouver, BC.
- Ferris, D., & Helt, M. (2000). *Was Truscott right? New evidence on the effects of error correction in L2 writing classes*. Paper presented at the AAAL Conference, Vancouver, BC.
- Ferris, D., & Roberts, B. (2001). Error feedback in L2 writing classes: How explicit does it need to be? *Journal of Second Language Writing*, 10(3), 161-184.
- Ferris, D. R. (1997). The influence of teacher commentary on student revision. *Tesol Quarterly*, 31(2), 315-339.
- Ferris, D. R. (2004). The “Grammar Correction” Debate in L2 Writing: Where are we, and where do we go from here?(and what do we do in the meantime...?). *Journal of second language writing*, 13(1), 49-62.
- Ferris, D. R. (2006). Does error feedback help student writers? New evidence on the short-and long-term effects of written error correction. *Feedback in second language writing: Contexts and issues*, 81-104.
- Gall, M., Borg, W., & Gall, J. (1996). *Educational research: an introduction*, 6th edn (White Plains, NY, Longman).
- Gearhart, M., Herman, J., Baker, E., & Whittaker, A. (1992). *Writing portfolios at the elementary level: A study of methods for writing assessment* (CSE Tech. Rep. No. 337). Los Angeles: University of California. *Center for Research on Evaluation, Standards, and Student Testing*.
- Godshalk, F. I., Swineford, F., & Coffman, W. E. (1966). *The measurement of writing ability*: College Entrance Examination Board.
- Gosling, G. W. (1966). *Marking English Compositions: Research Into the Marking of*: Australian Council for Educational Research.
- Graham, S., Weintraub, N., & Berninger, V. W. (1998). The relationship between handwriting style and speed and legibility. *The Journal of Educational Research*, 91(5), 290-297.
- Green, A. (2005). EAP study recommendations and score gains on the IELTS Academic Writing test. *Assessing Writing*, 10(1), 44-60.
- Green, A. (2006). Washback to the learner: Learner and teacher perspectives on IELTS preparation course expectations and outcomes. *Assessing Writing*, 11(2), 113-134.
- Gronlund, N., & Linn, R. (1990). *Constructing objective test items: multiple-choice forms*. Measurement and Evaluation in Teaching: New York: MacMillan.
- Guénette, D. (2007). Is feedback pedagogically correct?: Research design issues in studies of feedback on writing. *Journal of Second Language Writing*, 16(1), 40-53.
- Hahn, J. (1981). Students' reactions to teachers' written comments. *National Writing Project Network Newsletter*, 4(1), 7-10.
- Hamp-Lyons, L. (1990). Second language writing: Assessment issues. *Second language writing: Research insights for the classroom*, 69-87.

- Hamp-Lyons, L. (1991). *Assessing second language writing in academic contexts*: ERIC.
- Hamp-Lyons, L. (1991). Pre-text: Task-related influences on the writer. *Assessing second language writing in academic contexts*, 87-107.
- Hanaoka, O. (2007). Output, noticing, and learning: An investigation into the role of spontaneous attention to form in a four-stage writing task. *Language Teaching Research*, 11(4), 459-479.
- Harklau, L. (2002). The role of writing in classroom second language acquisition. *Journal of second language writing*, 11(4), 329-350.
- Hayes, A. F., & Preacher, K. J. (2012). Conditional process modeling: Using structural equation modeling to examine contingent causal processes. *Structural equation modeling: A second course*.
- Hinkel, E. (1994). Native and nonnative speakers' pragmatic interpretations of English texts. *Tesol Quarterly*, 28(2), 353-376.
- Horowitz, & Daniel. (1986a). Process, not product: Less than meets the eye. *TESOL quarterly*, 20(1), 141-144.
- Horowitz, & Daniel. (1986b). What professors actually require: Academic tasks for the ESL classroom. *TESOL quarterly*, 20(3), 445-462.
- Huck, S., & Bounds, W. (1972). Essay Grades: An Interaction Between Graders' Handwriting Clarity and the Neatness of Examination Papers. *American Educational Research Journal*.
- Hughes, A. (1989). *Testing for language teachers* (Vol. 1): Cambridge University Press Cambridge.
- Hughes, D. C., Keeling, B., & Tuck, B. F. (1983). Effects of achievement expectations and handwriting quality on scoring essays. *Journal of Educational Measurement*, 20(1), 65-70.
- Huot, B. (1990a). The literature of direct writing assessment: Major concerns and prevailing trends. *Review of Educational Research*, 60(2), 237-263.
- Huot, B. (1990b). Reliability, validity, and holistic scoring: What we know and what we need to know. *College Composition and Communication*, 201-213.
- Huot, B. (1996). Toward a new theory of writing assessment. *College composition and communication*, 47(4), 549-566.
- Huot, B. A. (1988). *The validity of holistic scoring: A comparison of the talk-aloud protocols of expert and novice holistic raters*: UMI.
- Hyland, K. (2001). Humble servants of the discipline? Self-mention in research articles. *English for Specific Purposes*, 20(3), 207-226.
- Hyland, K. (2003a). Genre-based pedagogies: A social response to process. *Journal of Second Language Writing*, 12(1), 17-29.
- Hyland, K. (2003b). *Second language writing*: Ernst Klett Sprachen.
- Jacobs, H. L., Zinkgraf, S. A., Wormuth, D. R., Hartfiel, V. F., & Hughey, J. B. (1981). *Testing ESL composition: A practical approach*: Newbury House Rowley, MA.
- James, C. (1998). *Errors in language learning and use*: Longman London New York.
- James, C. (1998). *Errors in language learning and use: exploring error analysis*: Longman.
- Johnson, L, R., & Bergman, T. (1996). Collaborating To Create a Portfolio Assessment in a Small-Scale Evaluation Context.
- Johnson, B., & Christensen, L. (2000). *Educational research: Quantitative and qualitative approaches*: Allyn & Bacon.

- Jussim, L., & Eccles, J. S. (1992). Teacher expectations: II. Construction and reflection of student achievement. *Journal of personality and social psychology*, 63(6), 947.
- Keech, C., & McNelly, M. E. (1982). Comparison and analysis of rater responses to the anchor papers in the writing prompt variation study. *Properties of writing tasks: A study of alternative procedures for holistic writing assessment*. Berkeley: University of California, Graduate School of Education, Bay Area Writing Project.
- Keith, T. Z., Reimers, T. M., Fehrmann, P. G., Pottebaum, S. M., & Aubey, L. W. (1986). Parental involvement, homework, and TV time: Direct and indirect effects on high school achievement. *Journal of Educational Psychology*, 78(5), 373.
- Kepner, C. G. (1991). An Experiment in the Relationship of Types of Written Feedback to the Development of Second-Language Writing Skills. *The modern language journal*, 75(3), 305-313.
- Klein, Joseph, & Taub, D. (2005). The effect of variations in handwriting and print on evaluation of student essays. *Assessing Writing*, 10(2), 134-148.
- Klein, D. F., Thase, M. E., Endicott, J., Adler, L., Glick, I., Kalali, A., . . . Bystritsky, A. (2002). Improving clinical trials: American society of clinical psychopharmacology recommendations. *Archives of general psychiatry*, 59(3), 272-278.
- Knoblauch, C., & Brannon, L. (1981). Teacher Commentary on Student Writing: The State of the Art. *Freshman English News*, 10(2), 1-4.
- Knoblauch, C. H., & Brannon, L. (1984). *Rhetorical Traditions and the Teaching of Writing*: ERIC.
- KOBAK, D. (2007). Continual efforts to achieve results. *EUROPEAN RAILWAY REVIEW*(1).
- Kobak, K. A. (2004). A comparison of face-to-face and videoconference administration of the Hamilton Depression Rating Scale. *Journal of telemedicine and telecare*, 10(4), 231-235.
- Kobak, K. A., Engelhardt, N., & Lipsitz, J. D. (2006). Enriched rater training using Internet based technologies: a comparison to traditional rater training in a multi-site depression trial. *Journal of psychiatric research*, 40(3), 192-199.
- Kobak, K. A., Engelhardt, N., Williams, J. B., & Lipsitz, J. D. (2004). Rater training in multicenter clinical trials: issues and recommendations. *Journal of Clinical Psychopharmacology*, 24(2), 113-117.
- Kobak, K. A., Williams, J. B., & Engelhardt, N. (2008). A comparison of face-to-face and remote assessment of inter-rater reliability on the Hamilton Depression Rating Scale via videoconferencing. *Psychiatry research*, 158(1), 99-103.
- Kobayashi, H., & Rinnert, C. (1996). Factors affecting composition evaluation in an EFL context: Cultural rhetorical pattern and readers' background. *Language learning*, 46(3), 397-433.
- Lalande, J. F. (1982). Reducing composition errors: An experiment. *The modern language journal*, 66(2), 140-149.
- Lantolf, J. P. (2000). *Sociocultural theory and second language learning*: Oxford University Press.
- Lapkin, S., Swain, M., & Smith, M. (2002). Reformulation and the learning of French pronominal verbs in a Canadian French immersion context. *The Modern Language Journal*, 86(4), 485-507.

- Lee, G., & Schallert, D. L. (2008). Meeting in the margins: Effects of the teacher–student relationship on revision processes of EFL college students taking a composition course. *Journal of Second Language Writing*, 17(3), 165-182.
- Lee, I. (1997). ESL learners' performance in error correction in writing: Some implications for teaching. *System*, 25(4), 465-477.
- Lee, I. (2003). L2 writing teachers' perspectives, practices and problems regarding error feedback. *Assessing Writing*, 8(3), 216-237.
- Leki, I. (1986). *ESL student preferences in written error correction*. Paper presented at the Southeast Regional TESOL Conference, Atlanta, Ga., October.
- Leki, I. (1990). Coaching from the margins: Issues in written response. *Second language writing: Research insights for the classroom*, 57-68.
- Leki, I. (1991). The Preferences of ESL Students for Error Correction in College-Level Writing Classes. *Foreign language annals*, 24(3), 203-218.
- Leki, I. (1994). Coaching from the Margins: Issues in Written Response. "Second Language Writing". Edited by B. Kroll: Cambridge University Press.
- Linn, R., & Gronlund, N. (2000). *Validity. su: Measurement and assessment in teaching*. Upper Saddle River, NJ: Merrill.
- Lipsitz, J., Kobak, K., Feiger, A., Sikich, D., Moroz, G., & Engelhard, A. (2004). The rater applied performance scale: development and reliability. *Psychiatry research*, 127(1), 147-155.
- Lloyd-Jones, R. (1977). Primary trait scoring. *Evaluating writing: Describing, measuring, judging*, 33-66.
- Loewen, S. (2003). Variation in the frequency and characteristics of incidental focus on form. *Language Teaching Research*, 7(3), 315-345.
- Loewen, S. (2004). Uptake in Incidental Focus on Form in Meaning-Focused ESL Lessons. *Language learning*, 54(1), 153-188.
- Long, M. H., & Sato, C. (1983). Classroom foreigner talk discourse: Forms and functions of teachers' questions. *Classroom oriented research in second language acquisition*, 268-285.
- Longford, N. T. (1994). *A case for adjusting subjectively rated scores in the Advanced Placement tests*: Educational Testing Service.
- Mackey, A. (2004). *Cognition and second language instruction*: Cambridge Univ Press.
- Manchón, R. (2011). *Learning-to-write and Writing-to-learn in an Additional Language* (Vol. 31): John Benjamins Publishing.
- Manke, M., & Loyd, B. (1990). *An investigation of non achievement-related factors influencing teachers' grading practices*. Paper presented at the annual meeting of the National Council on Measurement in Education, Boston.
- Mantello, M. (1997). Error correction in the L2 classroom. *Canadian modern language review*, 54(1), 127-131.
- Markham, L. R. (1976). Influences of handwriting quality on teacher evaluation of written work. *American Educational Research Journal*, 13(4), 277-283.
- Marshall, J. C., & Powers, J. M. (1969). Writing neatness, composition errors, and essay grades. *Journal of Educational Measurement*, 6(2), 97-101.
- McColly, W., & Remstad, R. (1965). Composition rating scales for general merit: An experimental evaluation. *The Journal of Educational Research*, 55-56.
- McMillan, J. H. (1997). *Classroom Assessment. Principles and Practices for Effective Instruction*: ERIC.
- McNamara, T. F., & Candlin, C. N. (1996). *Measuring second language performance*: Longman London.

- McNamara, T. F., & McNamara, T. (1996). *Measuring second language performance*: Longman London.
- Mendelsohn, D., & Cumming, A. (1987). Professor's ratings of language use and rhetorical organizations in ESL compositions. *TESL Canada Journal*, 5(1), 09-26.
- Messick, S. (1989). Meaning and values in test validation: The science and ethics of assessment. *Educational researcher*, 18(2), 5-11.
- Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational researcher*, 23(2), 13-23.
- Moore, T., & Morton, J. (2005). Dimensions of difference: a comparison of university writing and IELTS writing. *Journal of English for Academic Purposes*, 4(1), 43-66.
- Moss, P. A. (1994). Can there be validity without reliability? *Educational researcher*, 23(2), 5-12.
- Müller, M. J., & Szegedi, A. (2002). Effects of interrater reliability of psychopathologic assessment on power and sample size calculations in clinical trials. *Journal of clinical psychopharmacology*, 22(3), 318-325.
- Natriello, G., & McDill, E. L. (1986). Performance standards, student effort on homework, and academic achievement. *Sociology of education*, 18-31.
- Onwuegbuzie, A. J., & Leech, N. L. (2005). Taking the "Q" Out of Research: Teaching Research Methodology Courses Without the Divide Between Quantitative and Qualitative Paradigms. *Quality & Quantity*, 39, 261-272.
- Palmer, A., & Bachman, L. (1996). *Language testing in practice*: Oxford: Oxford University Press.
- Peacock, M. (1988). Handwriting versus wordprocessed print: an investigation into teachers' grading of English Language and Literature essay work at 16+. *Journal of Computer Assisted Learning*, 4(3), 162-172.
- Penny, J., Johnson, R. L., & Gordon, B. (2000). The effect of rating augmentation on inter-rater reliability: An empirical study of a holistic rubric. *Assessing Writing*, 7(2), 143-164.
- Peterson, E., & Lou, W. W. (1991). *The Impact of Length on Handwritten and Wordprocessed Papers*.
- Pincas, A. (1962). Structural linguistics and systematic composition teaching to students of English as a foreign language. *Language learning*, 12(3), 185-194.
- Polio, C., & Fleck, C. (1998). "If I only had more time:" ESL learners' changes in linguistic accuracy on essay revisions. *Journal of Second Language Writing*, 7(1), 43-68.
- Porte, G. (1996). When writing fails: How academic context and past learning experiences shape revision. *System*, 24(1), 107-116.
- Powers, D. E., Fowles, M. E., Farnum, M., & Ramsey, P. (1994). They Think Less of My Handwritten Essay If Others Word Process Theirs? *Journal of Educational Measurement*, 31(3), 220-233.
- Qi, D. S., & Lapkin, S. (2001). Exploring the role of noticing in a three-stage second language writing task. *Journal of Second Language Writing*, 10(4), 277-303.
- Radecki, P. M., & Swales, J. M. (1988). ESL student reaction to written comments on their written work. *System*, 16(3), 355-365.
- Raimes, A. (1983). Anguish as a second language? Remedies for composition teachers. *Learning to write: First language/second language*, 258-272.
- Reid, J. (1984). The radical outliner and the radical brainstormer: A perspective on composing processes. *TESOL quarterly*, 18(3), 529-534.

- Rivers, W. M. (1968). CONTRASTIVE LINGUISTICS IN TEXTBOOK AND CLASSROOM WILGA M. RIVERS. *Monograph Series on Languages and Linguistics*(21), 151.
- Robb, T., Ross, S., & Shortreed, I. (1986). Salience of feedback on error and its effect on EFL writing quality. *Tesol Quarterly*, 20(1), 83-96.
- Russell, M., & Plati, T. (2000). Mode of Administration Effects on MCAS Composition Performance for Grades Four, Eight, and Ten. A Report of Findings Submitted to the Massachusetts Department of Education. NBETPP Statements World Wide Web Bulletin.
- Russell, M., & Plati, T. (2002). Does it matter with what I write?: Comparing performance on paper, computer and portable writing devices. *Current Issues in Education*, 5(4), 24.
- Russell, M., & Tao, W. (2004). The influence of computer-print on rater scores. *Practical Assessment, Research and Evaluation*, 9(1).
- Santos, T. (1988). Professors' Reactions to the Academic Writing of Nonnative-Speaking Students. *Tesol Quarterly*, 22(1), 69-90.
- Scarcella, R. C., & Oxford, R. L. (1992). *The tapestry of language learning: The individual in the communicative classroom*: Heinle & Heinle Boston.
- Schmidt, R. (1993). Awareness and second language acquisition. *Annual review of applied linguistics*, 13(1), 206-226.
- Schmidt, R. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. *Consciousness in second language learning*, 11, 237-326.
- Schmidt, R. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. *Attention and awareness in foreign language learning*, 1-63.
- Schmidt, R., & Frota, S. (1986). Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese. *Talking to learn: Conversation in second language acquisition*, 237-326.
- Schmidt, R. W. (1990). The Role of Consciousness in Second Language Learning1. *Applied linguistics*, 11(2), 129-158.
- Semke, H. D. (1984). Effects of the red pen. *Foreign Language Annals*, 17(3), 195-202.
- Sheen, Y. (2007). The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *Tesol Quarterly*, 41(2), 255-283.
- Sheen, Y. (2010). Introduction: The role of oral and written corrective feedback in SLA. *Studies in Second Language Acquisition*.
- Sheppard, K. (1992). Two feedback types: Do they make a difference? *RELC journal*, 23(1), 103-110.
- Silva, T. (1990). Second language composition instruction: Developments, issues, and directions in ESL. *Second language writing: Research insights for the classroom*, 11-23.
- Sloan, C. A., & McGinnis, I. (1982). The effect of handwriting on teachers' grading of high school essays. *Journal of the Association for the Study of Perception*.
- Slomp, D. H., & Fuite, J. (2004). Following Phaedrus: Alternate choices in surmounting the reliability/validity dilemma. *Assessing Writing*, 9(3), 190-207.
- Soloff, S. (1973). Effect of non-content factors on the grading of essays. *Graduate Research in Education and Related Disciplines*, 6(2), 44-54.

- Sperling, M., & Freedman, S. W. (1987). A Good Girl Writes Like a Good Girl Written Response to Student Writing. *Written communication*, 4(4), 343-369.
- Sprouse, J. L., & Webb, J. E. (1994). The Pygmalion Effect and Its Influence on the Grading and Gender Assignment on Spelling and Essay Assessments.
- Stemler, S. E. (2004). A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. *Practical Assessment, Research & Evaluation*, 9(4), 66-78.
- Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied linguistics*, 16(3), 371-391.
- Swain, M., & Lapkin, S. (2001). Focus on form through collaborative dialogue: Exploring task effects. *Researching pedagogic tasks: Second language learning, teaching and testing*, 99-118.
- Sweedler-Brown, C. O. (1985). The Influence of Training and Experience on Holistic Essay Evaluations. *English Journal*, 74(5), 49-55.
- Sweedler-Brown, C. O. (1991). Computers and Assessment: The Effect of Typing versus Handwriting on the Holistic Scoring of Essays. *Research & Teaching in Developmental Education*, 8(1), 5-14.
- Sweedler-Brown, C. O. (1993). ESL essay evaluation: The influence of sentence-level and rhetorical features. *Journal of Second Language Writing*, 2(1), 3-17.
- Syndicate, U. o. C. L. E. (1999). *English Subjects: Report on the June 1999 Examination*: University of Cambridge.
- Taylor, S., & Asmundson, G. J. (2008). Internal and external validity in clinical research. *Handbook of Research Methods in Abnormal and Clinical Psychology*, 23-34.
- Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language learning*, 46(2), 327-369.
- Truscott, J. (1999). The case for "The case against grammar correction in L2 writing classes": A response to Ferris. *Journal of Second Language Writing*, 8(2), 111-122.
- Truscott, J. (2004). Evidence and conjecture on the effects of correction: A response to Chandler. *Journal of Second Language Writing*, 13(4), 337-343.
- Truscott, J. (2007). The effect of error correction on learners' ability to write accurately. *Journal of Second Language Writing*, 16(4), 255-272.
- Vaughan, C. (1991). Holistic assessment: What goes on in the rater's mind. *Assessing second language writing in academic contexts*, 111-125.
- Vygotsky, L. S. (1978). *Mind and society: The development of higher mental processes*: Cambridge, MA: Harvard University Press.
- Weir, C. J. (1990). Communicative language testing with special reference to English as a foreign language. *Exeter Linguistic Studies*, 11, 1-241.
- White, E. M. (1985). *Teaching and Assessing Writing: Recent Advances in Understanding, Evaluating, and Improving Student Performance. The Jossey-Bass Higher Education Series*: ERIC.
- Williams, J. (1999). Learner-Generated Attention to Form. *Language Learning*, 49(4), 583-625.
- Williams, J. (2001). Learner-Generated Attention to Form. *Language Learning*, 51(s1), 303-346.
- Williams, J. (2012). The potential role (s) of writing in second language development. *Journal of Second Language Writing*, 21(4), 321-331.

- Williams, J. B. (1988). A structured interview guide for the Hamilton Depression Rating Scale. *Archives of general psychiatry*, 45(8), 742-747.
- Zamel, V. (1983). The composing processes of advanced ESL students: Six case studies. *Tesol Quarterly*, 17(2), 165-188.



© COPYRIGHT UPM