In The Name of God

University of Isfahan
Faculty of Foreign Languages
Department of English Language

A PhD Thesis
Presented in Partial Fulfillment of the Requirements for
The Degree of Doctor of Philosophy in TEFL

A Critical Reflection on the High-stake Tests' Effects: Toward a Local Model for Teacher Professionalism in Iranian Preparation Courses

Supervisor:
Dr. Saeed Ketabi

Advisor:
Dr. Mansoor Tavakoli

By:
Sima Sadeghi

April 2012

کلیه حقوق مادی مترتب بر نتایج مطالعات، ابتکارات

و نوآوری های ناشی از تحقیق موضوع این پایان نامه

متعلق به دانشگاه اصفهان است.

دانشگاه اصفهان
دانشکده زبانهای خارجی
گروه انگلیسی

پایان نامه ی دکتری رشته ی زبان انگلیسی گرایش آموزش خانم سیما صادقی تحت عنوان

بررسی انتقادی تاثیرات آزمونهای مهم جهانی: ارائه مدلی بومی برای حرفه ای نمودن مدرسان در دوره های آمادگی پیش از آزمون

در تاریخ ۹۱/۱/۲۳ توسط هیأت داوران زیر بررسی و با درجه عالی (۱۹/۷۰) به تصویب نهایی رسید.

۱- استاد راهنمای پایان نامه دکتر سعید کتابی با مرتبه ی علمی استادیار امضا
۲- استاد مشاور پایان نامه دکتر منصور توکلی با مرتبه ی علمی استادیار امضا
۳- استاد داور داخل گروه دکتر داریوش نژاد انصاری با مرتبه ی علمی استادیار امضا
۴- استاد داور داخل گروه دکتر عزیزالله دباغی با مرتبه ی علمی استادیار امضا
۵- استاد داور خارج از گروه دکتر منصور کوشا با مرتبه ی علمی دانشیار امضا

امضای مدیر گروه

گروه انگلیسی
دانشکده زبانهای خارجی

Acknowledgements

I am grateful to those great people who inspired me and encouraged me to pursue this goal. First, I would like to express my heartfelt gratitude to my supervisor, Dr. Saeed Ketabi, who has been an incredible mentor and teacher. I always appreciate his effort toward ensuring that my PhD was not only a degree to be proud of but a continual process of inquiry. Thank you for unwavering support, ongoing recommendations and graceful inspiration you provided for me during my PhD program. I am also deeply indebted to my advisor and my consummate mentor, Dr Mansoor Tavakoli, for his sharp insights, tireless efforts, and encouragement. I am forever appreciative for the considerable part you played in this study.

I would like to acknowledge the generous cooperation of the Language Center at the University of Isfahan. I thank each and everyone who had a role in my accomplishment.

I wish to particularly acknowledge Iran Ministry of Science, Education and Technology which provided me with financial support for my PhD sabbatical leave in the Netherlands.

Finally, I would like to extend my sincere thank to my family for giving me their unconditional love, care and patience.

Dedication

This effort is dedicated to my late Quran teacher

Abstract

Given the increasing public demands for preparation courses, the purpose of this study was to explore how high-stakes testing influenced the nature and process of preparation courses as well as teachers' instructional decisions and professionalism. It also raised challenging questions regarding the teacher's reactions to tests' pressures and the factors involved in variations of responses. To achieve this end, concepts from Structuration Theory and Washback Hypothesis were employed to capture the complex and dynamic interaction between tests and teachers. The study was conducted at the Isfahan University Language Center (IULC) in Iran. It dealt with the performance of four instructors selected through convenient sampling at preparation courses of high-stake tests, two in the IELTS and two in TOEFL PBT.

This interpretive study was conducted in three phases. A mixed method was used for obtaining both quantitative and qualitative data. Phase one, an ethnographic case study, involved using observation and field notes to gather data about how high-stakes testing affected teachers' curriculum and methodology treatment. Part A of COLT and UCOS were used as observation schemes to collect in-depth and in-situ data. The second phase aimed at eliciting the effects of the tests on teachers' feelings, attitude and perceptions. A questionnaire survey was used to collect relevant data. To capture an accurate meaning of the teachers' professional decisions and practice, field notes and an interview were also conducted. Grounded Theory method was used for inductively analyzing data from interview and field notes. The constant comparison method and coding system, open, axial and selective, aided in uncovering the themes grounded in the data. Four themes, personality, professional, academic and strategic compromise, with several subthemes emerged revealing the factors related to teachers which could be influential for bringing about positive and negative washback. Cumulatively, the themes indicated high-stakes testing placed teachers in the dilemma position of pedagogical and ethical challenges to set a balance between tests' mandates and teachers' authority.

One theme, Strategic Compromise, related specifically to teachers' professional responses to the mandates of tests. Strategic compromise seemed more like a continuum accounting for the varying reactions of the teachers to the compliance of the same test, which went beyond the simple dichotomy of "blind adherence" or "ignorance".

Keywords: Ethnographic Case Study, Grounded Theory, High-stakes Testing, Strategic Compromise Repertoire, Structuration Theory, Teacher Professionalism

TABLE OF CONTENTS

Title **Page**

LIST OF TABLES.....VI

LIST OF FIGURES…..IX

LIST OF ABBREVIATIONS….....x

Chapter I: INTRODUCTION

1.1. Overview…….……..…….....1

1.2. Background of the Study…….…….....3

1.2.1. High-stake Tests ……..…….4

1.2.2. Preparation Courses ……..…6

1.2.3. Teacher Professional Performance……8

1.3. Statement of the problem ……....10

1.4. Purpose of the Study…….……..…….....13

1.5. Research Questions..…… ……..…….....14

1.6. Theoretical Framework ……15

1.7. Assumptions…….……..17

1.8. Significance of the Study …….…….....17

1.9. Definition of Terms……..……..20

1.10. Organization of the Dissertation…….……..22

Chapter II: REVIEW OF THE RELATED LITERATURE

2.1. Introduction……..……..25

2.2. The Origin of Washback ……..……..27

2.3. Definitions of Washback ……..……..33

2.4. Directions of Washback…….....35

2.5. Part II: Theoretical Models of Washback.....40

2.5.1. Hughes's Backwash Model ……...41

2.5.2. Alderson and Wall's Model ……..43

2.5.3. Bailey's Model ……..……..42

2.5.4. Burrows' Model ……..……..…45

2.6. Part III: Exploring Scopes of Washback.....47

2.6.1. Washback on Teaching.....47

2.6.2. Washback on Learning.....52

2.6.3. Washback on Materials.....55

2.6.4. Washback on Feeling.....58

2.6.5. Washback studies in Iranian Contexts.....60

2.7. Washback and Mediating Factors 61

2.8. Pedagogical Implication.....66

2.9. Chapter Summary.....68

Chapter III: METHODOLOGY

3.1. Introduction…….…..70

3.2. Qualitative/Quantitative Approach.…..…….....71

3.3. Mixed Method…..…….....75

3.4. Rational for Research Methodology……..77

3.5. Classroom Research …….……79

3.5. 1. Context of the Study…….……...82

3.5.2. Participants……..……83

3.5.3. Sampling Procedure……..……86

3.5.4. Ethical Consideration…...……87

3.5.5. Researcher Role…..…...……88

3.6. Overall Research Design……89

3.7. Data Collection Procedure…….……91

3.7.1. Ethnographic Observation……...92

3.7.2. Field note……95

3.7.3. Questionnaire Survey…….96

3.7.4. Interview ……99

3.8. Data Analysis Procedures…….…….101

3.8.1. Quantitative Data Analysis……..102

3.8.2. Qualitative Data Analysis……102

3.9. Trustworthiness…….…….……..108

3.9.1. Credibility…….……..….108

3.9.2. Dependability…….…….……..109

3.9.3. Transferability…….…….……..….110

3.9.4. Neutrality.…….…….…….……..…...110

3.10. Chapter Summary.....111

Chapter IV: RESULTS

4.1. Introduction…….……112

4.2. Ethnographic Observation: COLT…….....113

4.2.1. Participant Organization…… ……....114

4.2.2. Teaching Activity, Content & Materials……..……..115

4.3. Ethnographic Observation: UCOS.……..……..…...119

4.4. Washback Evidence from COLT and UCOS……..……...122

4.5. Phase II: Questionnaire Survey…….……..……125

4.5.1. TQ Part One…….……..……126

4.5.2. TQ Part Two…….……..……..……126

4.5.2. TQ Part Three…..……..……..…….129

4.6. Phase III: Grounded Theory……..……..……..131

4.6.1. Stage one: Open Coding……..……133

4.6.2. Stage Two: Axial Coding……...134

4.6.3. Stage Three: Selective Coding…….....134

4.7. Synthesizing Themes to Develop a Grounded Theory….....135

4.7.1. Personality Factors……..……136

4.7.2. Academic Factors ……...138

4.7.3. Professional Factors…….....140

4.7.4. Strategic Compromise…….....142

4.8. Reliability/credibility of Themes…….....147
4.9. Chapter Summary…….....148

Chapter V: CONCLUSION and IMPLICATIONS

5.1. Introduction……..149
5.2. Summary of the Study…….....150
5.3. Discussion…….....154
5.4. Limitations of the Study….....157
5.5. Suggestions for Future Research.....159
5.6. Implications of the Study.....160
5.7. Conclusion.....162

References.....165

- **Appendices**.....179
- Appendix 1: Modular Structure of IELTS Test….....179
- Appendix 2: IELTS Band Descriptors….....180
- Appendix 3: Modular Structure of TOEFL…….....182
- Appendix 4: BAK Model of Washback…….....183
- Appendix 5: Consent Letter (Director) ……..184
- Appendix 6: Consent Letter (Instructor).....185
- Appendix 7: Part A of the Communicative Orientation of Language Teaching (COLT)…….....186
- Appendix 8: University of Cambridge Observation Scheme (UCOS) …….....187
- Appendix 9: Teachers' Questionnaire.....188
- Appendix 10: Teachers' Semi-structured Interview Protocol.....191
- Appendix 11: Coding Structure.....193

List of Tables

Table 2.1. Messick's validity model……..32

Table 2.2. Trichotomy of washback Model (Hughes, 1993)……40

Table 2.3. Summary of empirical Studies on washback of high-stake tests (TOEFL & IELTS)……59

Table 2.4. Pan's model of intervening factors……65

Table 3.1: Characteristics of qualitative and quantitative research……...75

Table 3.2. Demographic characteristics of the teachers…….85

Table 3.3. Data collection design……....91

Table 3.4. Classroom observation scheme……..94

Table 4.1. Time segmented participant organization, COLT……114

Table 4.2. Segmented classroom activity, COLT……..……116

Table 4.3. Exam-related activity and text, UCOS results …….121

Table 4.4. Teachers' attitude toward high-stake tests……..……..127

Table 4.5. Teachers' attitude toward preparation courses.…..……..128

Table 4.6. Teachers' attitude toward effective teaching…….……128

Table 4.7. Teachers' attitude toward washback on feeling…….129

Table 4.8. Teachers' attitude toward washback on teaching ……..130

Table 4.9. Open/provisional codes…….133

Table 4.10. Teacher-related Factors affecting washback……..136

List of Figures

Figure 2.1. Green's theory on washback and construct……39

Figure 2.2. Bailey's basic model of washback (1996)……..44

Figure 2.3. Burrows' model of washback (1998)…….46

Figure 2.4. Shih's model of intervening factors……64

Figure 4.1. Participant organization, COLT……115

Figure 4.2. Teacher activities, COLT ……117

Figure 4.3. Student activities, COLT ……118

Figure 4.4. Washback and styles of preparation courses……....124

Figure 4.5. Ethical preparation courses……...124

Figure 4.6. Mechanism of Strategic Compromise..……..……...144

Figure 4.7. Professional performance based on SCR……..…….144

Figure 5.1. Schematic research procedure……..……..……152

Figure 5.2. Grounded Theory process……..…….153

Figure 5.3. Proposed washback model for accountability……156

List of Abbreviations

COLT: Communicative Orientation of Language Teaching

CR: Classroom Research

ETS: English Testing System

GP: General Proficiency

IELTS: International English Language Testing System

IULC: Isfahan University Language Center

MDI: Measurement-driven Instruction

QUAN vs. QUAL: Quantitative vs. Qualitative

SCR: Strategic Compromise Repertoire

ST: Structuration Theory

TOEFL: Test of English as a Foreign Language

TQ: Teacher Questionnaire

UCOS: University of Cambridge Observational Scheme

Chapter One
Introduction

1.1. Overview

Language testing has long been considered as a binding theme between two influential domains of applied linguistics: teaching and learning. In education, tests are used for several purposes: entrance, diagnostic, admission, placement, achievement, proficiency, etc. Thus, tests have the power to select, reward, motivate and at the same time they can exclude, punish and de-motivate (Hayes, 2003, p.1). One of the most important uses of language tests is as a gate-keeper: whether to allow individuals to enter into a program or not. Considering the mobility of knowledge/skill and the increasing interest in studying abroad and migration, there has been a growing demand for proficiency tests which are mostly used as gate-keepers, especially for admission into higher-education.

Two of these popular tests which are used internationally to provide language proficiency benchmarks for the admission into universities are the International English Language Testing System (IELTS) and Test of English as a Foreign Language (TOEFL). These tests are considered to be quite high-stake. High-stake tests refer to those tests whose results are used to make important decision regarding test-takers. It is argued that these tests, with their associated consequences, exert influence, either desirable or undesirable, over teaching and learning, a phenomenon which is frequently referred to as "washback", which receives an increasing attention from test developers, researchers, educators, material developers, administrators, teachers etc. The term washback or 'backwash' as it is sometimes referred to, can be generally defined as the effect of testing on teaching and learning, both positive and negative, and is therefore a form of impact (Alderson and Wall, 1993; Biggs, 1995).

High-stakes testing and accountability provoke numerous reactions. Opponents of the focus on testing and accountability believe it has been deleterious to teaching and learning (Popham, 2001). Some charge that teachers teach to the test, spend excessive class time on test preparation exercises, and place too much emphasis on standardized testing (Barksdale-Ladd & Thomas, 2000). In contrast, other research points to the beneficial outcomes of testing, such as improved professional development and increased student-centered pedagogies and achievement (Au, 2007; Cizek, 2001).

However, it should be noted that some studies indicate that washback is an extremely complex and elusive phenomenon (Alderson & Wall 2003; Cheng 2005). While the relationship between exams and washback is sometimes thought of as a simple one in which exams generate a direct, automatic and blanket washback, it seems that other factors beside the exam itself play their part in determining washback. Washback is a phenomenon that does not exist automatically in its own right but is rather one that can be brought into existence