LANCASTER UNIVERSITY

DEPARTMENT OF LINGUISTICS AND ENGLISH LANGUAGE

# The Reported Activities and Beliefs of the Students Preparing for the Specialised English Test (SPE)

**Mahmoud Moradi Abbasabady**

BA in TEFL (University for Teacher Education, Tehran, Iran)

MA in TEFL (Tarbiat Modares University, Tehran, Iran)



**This thesis is submitted in partial fulfilment of the requirements for the degree of**

**Doctor of Philosophy**

**2009**



ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

**Abstract**

The purpose of this study is to explore the 'washback effect' of a high stakes test known as the Specialised English Test (SPE) which is used to admit English-major applicants into English departments at universities in Iran. A belief commonly held about the SPE Test is that students admitted to university since its introduction in 2002 have been more proficient and successful than pre-2002 students. However, no research has been done on how the test might be affecting the learning of the students.

I examined the reported learning activities of the students to see what they reported they did in order to prepare for the test and why. As the SPE test preparation is done partly independently (by out-of-school resources), I also considered the students' reported learning beliefs to examine how the test was interacting with those criteria. Thus, reported activities and learning beliefs were the main themes of this study.

I collected the data from two contrasting groups- students who were going to take the SPE Test and students who were going to take the GE Test. I used three instruments to collect the data: questionnaire, letters, and interviews. 1038 students responded to the questionnaire, 91 students wrote the letters, and 18 students were interviewed.

The results showed that the SPE Test did have washback in the areas where washback was intended. The study also identified factors other than the test which influenced the learners. As regards learning beliefs, the results showed that the test preparation activities of the students were both affecting them as well as were affected by them.

The study suggested that in addition to test innovations, the beliefs of the learners should also be taken into account in order to promote positive washback. The

study also recommended the use of letters as an instrument in washback studies provided they are written by interested students and the results are followed up by other instruments such as interviews.

## DECLARATION

I hereby confirm that this thesis is originally my work and it has not previously been submitted for a higher degree or publication elsewhere. I alone am responsible for any possible mistakes.

Mahmoud Moradi Abbasabady

## ACKNOWLEDGEMENTS

I would like to express my deepest appreciation to my supervisors, Professor Charles Alderson and Dr Jayanti Banerjee for their critical comments without which this work would not have been possible.

Special thanks go to my sponsor, the Ministry of Science, Research, and Technology, for the financial support, as well as the representatives of the Ministry in London who have been immensely helpful and supportive.

I would like to thank all the members of the Language Testing Research Group in Lancaster including the teachers and the students who gave me feedback and helped me with this study.

My sincere thanks go to all the students who acted as participants in this study and all the staff members in the Education Ministry offices and the schools who cooperated with me in data collection.

Finally, I would like to dedicate this work to my wife and our families for their patience and support, and to my father who has been struggling with his illness since the beginning of this study.

# Table of Contents

# List of Appendices

# List of Tables

# List of Figures

# Chapter 1. Introduction

This study investigates how learners are influenced by the Specialised English Test (SPE), which has been used to admit students to BA programmes in English in Iranian universities since 2002. A belief commonly held about the SPE Test (Appendix 1) is that students admitted through this test have been more successful in their studies at university than pre-2002 students who were admitted through the General English Test (GE) (Appendix 2). However, no empirical research has been carried out on what is actually happening under the influence of this test. Therefore, the purpose of this study is to investigate the washback effect (Alderson and Wall, 1993) of the SPE Test.

In the following sections, I will discuss the rationale for the introduction of the SPE Test, how this study began, what changes occurred in the admission system which resulted in the introduction of the SPE Test, what English learning opportunities there are for the students from an early age until university, predicted washback, research questions that have driven this investigation, and finally the contributions of the study.

## 1.1. Rationale for the introduction of the SPE Test

One of the high stakes tests in Iran is the entrance examination to universities under the Ministry of Science, Research and Technology (MSRT). It is a norm-referenced test that is designed and administered by the National Organisation for Educational Testing (NOET) depending on what field of study the applicants will be undertaking at university. The examination consists of a General Section and a Specialised Section (Appendix 3, Sections 7.1 and 7.2). The General Section, which includes GE and three non-English subjects, is the same for all the groups of test takers, but the specialised section is different for each group of applicants pursuing a

different field of study at university.

Before 2002, the ability of English-major applicants was measured only by the General Section and they did not have to take the Specialised Section. The GE Test (in the General Section) was based on the high school textbooks. However, NOET found this system of admission inefficient. According to two authorities in the Organisation, this test proved to be too easy for the applicants because many applicants got most or even 100% of the items correct and, as a consequence, the students admitted through this test did not fulfill the expectations of the English departments. For example, based on NOET's 2002 report, in 2001- the last year in which GE Test was used for the admission of English-major applicants- the scores which had the highest frequencies were at 80%-90% (p. 148) (see Table 1.3 for the national average). This system did not necessarily lead to the admission of more proficient applicants as it did not have a good discrimination power. Therefore, the authorities decided to introduce some changes which resulted in the introduction of the SPE Test. The SPE Test has now more items than the GE Test and is based on materials from outside the high school textbooks (see 1.4). However, based on a brief conversation I had with English teachers, SPE materials were not taught in school. Following these changes I began to ask some questions which led to the present study.

## 1.2. Beginning to ask washback questions

As teachers of undergraduate students, my colleagues and I observed that students of English admitted to the university after 2002 were more proficient than the previous cohorts. They seemed to have fewer problems in English courses of general nature such as grammar, reading, and writing as well as specialised courses of Applied Linguistics and literature. Following these observations, I began to ask some questions concerning the ostensible improvement which were to be the beginning of a washback

study: Is the improvement due to the SPE test? How is the SPE Test affecting the learners? Does the fact that SPE materials are not taught at school encourage independent learning? Do the learners know what they should do in order to learn or is it the SPE Test which tells them what they should do? In other words, do the students know good language learning practices or is it the test which makes them do appropriate language learning practices? In order to find out how I would be able to answer these questions, I referred to the washback literature and the beliefs literature. As far as washback is concerned, Watanabe (2004) states that in order to establish washback, first it is essential to specify the changes in the contents of the test prior to the research.

## 1.3. Major Changes

As mentioned in 1.1, some changes were introduced in 2002 in the admission system of the students into English departments (see also the Test Specification/Descriptions in Appendix 3). With the introduction of the SPE Test, English was given more weight, i.e. the number of English items increased, two new sections were added, and the weight of 'non-English subjects' were kept constant. Another important change was to base SPE items on materials from outside the school textbooks. The third major change involved an increase in the difficulty of the SPE Test in order to measure the language proficiency of the applicants more accurately and to admit students with better English backgrounds. I will discuss each of these changes below so that I can then predict their possible washback (see 1.5). I will also mention the questions which arose as I was examining the changes.

### 1.3.1. Increase in the weight of English

#### 1.3.1.1. Increase in the number of items in the sections retained from the GE Test

At present, the SPE Test has 6 sections and the GE Test, which is still being

used, has 4 sections. The four sections of the GE Test were retained. These common sections include Vocabulary, Grammar, Reading, and Cloze. Although these sections were retained, they were not retained in the same way. The number of items in each of those sections was increased, but the increase was not equal in all the sections. Grammar items were increased less than those of other sections (see Table 1.1). As I was examining this change, the questions which arose for me included:

Will SPE students spend more time studying the common sections than GE students?

Will each group of the SPE and GE test takers spend time on their own test sections based on the number of items?

#### 1.3.1.2. Addition of two new sections

The two new sections which were added to the SPE Test were Language Functions and Sentence Structure. Table 1.1 shows the increase in the number of items in the retained sections as well as the number of items of the two new sections.

**Table 1-1 Increase in the number of SPE items**

| | Skills/Test Sections | GE | SPE |
|---|---|---|---|
| **Retained/common sections** | Vocabulary | 10 | 20 |
| | Reading | 5 | 15 |
| | Cloze | 5 | 15 |
| | Grammar | 5 | 10 |
| **Newly added sections** | Language Functions | 0 | 5 |
| | Sentence Structure | 0 | 5 |
| **Total** | | **25** | **70** |

As Table 1.1 shows, the SPE Test has 70 items and the GE Test has 25 items. The table shows that the increase in the number of items of the retained sections as well as the addition of two new sections not only resulted in the increase in the weight of each section but also in the weight of English/SPE Test as a whole (Appendix 3).

Question raised:

How much time will SPE students spend studying the new sections?

### 1.3.1.3. Keeping the weight of the 'non-English subjects' constant

In addition to the increase in the number of items (including the addition of the two new sections), the weight of non-English subjects did not increase. Under the SPE Test, knowledge of English, as opposed to the Non-English subjects, is now a stronger determinant for admission to university than before. Table 1.3 shows the weighting of English and Non-English subjects in the old and new admission systems.

**Table 1-2 Weighting of English and Non-English subjects for English and non-English-major applicants in the old and current admission systems**

| | | English | | Non-English | | Total weighting of subjects | |
|---|---|---|---|---|---|---|---|
| | | G | S | G | S | English | Non-English |
| SPE Group (English Major) | Old admission system | 100 | | 225 | | **100** | **225** |
| | New/current admission system | 50 | 280 | 225 | | **330** | **225** |
| GE Group (Non-English Majors) | Current admission system | 50 | | 225 | 443.2 | **50** | **668** |

G= General Section, S= Specialised Section

As Table 1.2 shows, in the current admission systems for the SPE group the weight of English is more than Non-English subjects (330 versus 225 respectively) which includes three general subjects, while for the GE group the weight of Non-English subjects is more than English (668 versus 50). This is because for the GE group Non-English subjects include three general subjects as well as all the specialised subjects, while for the SPE group Non-English subjects only include the three general courses.

Question raised:

Will SPE students study English more than Non-English courses, and conversely, will GE students study Non-English courses more than English courses?

### 1.3.2. Curriculum broader than school curriculum

The SPE Test is based on a broader curriculum than the GE Test. The main book the NOET authored for the SPE test preparation is *Bridging the Gap* (see

description of the book in Appendix 3). In addition to this main book, the authorities recommended a number of other books, which they claimed were represented by *Bridging the Gap* (Appendix 3). As I had a brief conversation with English teachers, while the GE Test is based on only the school textbook, the SPE Test is based on both the high school textbook and extra materials and the extra materials are not taught in school.

Questions raised:

Will SPE students rely on both school and out-of-school resources (including extra materials and preparation classes) and will GE students rely on school alone (including school textbooks and school classes)?

Considering the fact that the SPE students have to rely partly on their own criteria and decisions (e.g. for the choice of learning materials and methods of learning) rather than their teachers' in this partly independent learning situation, how does the test affect their criteria or beliefs?

Do their beliefs affect their test preparation activities? What beliefs do they have about their learning?

Will SPE students' beliefs about learning be different from those of the GE students?

### 1.3.3. Increase in the difficulty of the SPE Test

In a brief conversation with two NOET authorities, they claimed that the SPE Test was more difficult and a more accurate measure of language proficiency than the GE Test for the purpose of admitting students with a better English background. However, the only empirical evidence available for the difficulty of the test was the national averages of the students admitted to English departments from 2001-2004 (Table 1.3), but no information was available on whether the differences between the

averages were statistically significant. Concerning measurement accuracy also there was no empirical evidence on whether the two tests were measuring different constructs. The national averages are reported in Table 1.3.

**Table 1-3 Mean scores on GE and SPE of English students admitted to university from 2001-2004**

| Before SPE was introduced | | After SPE was introduced | | | | | |
|---|---|---|---|---|---|---|---|
| 2001 | | 2002 | | 2003 | | 2004 | |
| GE | SPE | GE | SPE | GE | SPE | GE | SPE |
| 72.9 | - | 85 | 51.9 | 74 | 65.5 | 69.9 | 60.4 |

Based on NOET's reports (2002; 2003; 2004; 2005)

Table 1.3 shows that the national averages on the SPE Test in the years 2002, 2003 and 2004 were lower than the averages on the GE Test, which give an indication of the difficulty of the SPE. They were also lower than the GE in 2001 when the SPE had not been introduced yet.

Questions raised:

Will SPE students perceive the SPE Test more difficult than GE students will perceive the GE Test?

How will the perceived difficulty affect the learners?

Will SPE students attend English language institutes to develop their English background?

Will SPE students engage in activities that are not tested by the SPE Test such as oral activities to develop their general proficiency?

Based on the washback literature, after specifying the changes in the test, the question to be asked at the next stage was 'what would washback look like?' This question was guided by the intended washback effects of the test constructors, which in this study was the test specifications/description (Appendix 3), a theory of washback, namely Alderson and Wall's (1993) Washback Hypotheses (2.4.1) and a literature review (Chapter 2). However, before I made any predictions, I considered the educational context in which the students learned English.

## 1.4. English learning before university exam

There are several opportunities for the students to learn English before they take the university entrance exam. They may learn English in school including state and private schools, or outside school through family members, in pre-school institutions, private language institutes, other private classes usually run by individual teachers, or they may learn English on their own. I will discuss these opportunities below.

### 1.4.1. English learning at school

Students learn English for a total of six years at school. They learn English for 2 years in junior secondary school (known as Guidance School), 3 years in high school, and 1 year in pre-university school. However, schools are of three major types which may offer English instruction with varying degrees and qualities: 'state public schools', 'non-profit schools', and 'state smart schools'. The latter type is commonly believed to enjoy the highest standards in teaching. The different standards in these kinds of schools may create different demands for extra private classes outside school. Furthermore, it is not clear whether SPE and GE students need these classes differently.

### 1.4.2. English learning outside school

There are many opportunities for the students to learn English outside school as well. Depending on the educational background of the family, students may be taught English by their family members at home. Families may send them to pre-school institutions such as nurseries or kindergartens where English may be taught as part of the regular programmes.

There are private classes conducted by individual teachers as well as classes in private institutions which students can attend to receive extra tuition to help with their

English courses at school, learn English for their interests, prepare for the SPE or GE tests, etc. Among the private institutions are the English language institutes where students can start studying English (usually English of a general nature with emphasis on conversation) even from a very young age. Attendance at the English institutes may play a very crucial role in the success of the English major applicants.

Depending on where the students study English, the quality and amount of English instruction might be different. However, these opportunities are open to both SPE and GE students and the demand for English classes might be dictated differently by the two tests. Now, based on the changes in the SPE Test and my knowledge of the educational context, I present my predictions of possible washback in the following section.

## 1.5. Predicted/expected washback

### 1.5.1. Expectation concerning the increase in the weight of English

#### 1.5.1.1. Expectation concerning the increase in the number of items

Based on the fact that the number of the items increased, I expect that SPE students will study each of the common test sections more than GE students (1.3.1.1).

#### 1.5.1.2. Expectation concerning the new sections

Given the addition of two new sections, the SPE students will study the new sections of Language Functions and Sentence Structure, but GE students will not (1.3.1.2).

I expect further that based on the number of items in each test (Table 1.1), each group of the SPE and GE test takers will spend time on their own test sections as follows:

**Table 1-4 Expected amount of time to be spent on the SPE and GE test sections ranked from most to least time based on the number of items**

| Amount of time | GE Test | SPE Test |
|---|---|---|
| **Most time** | Vocabulary | Vocabulary |
| | Reading / Cloze / Grammar | Reading / Cloze |
| | | Grammar |
| **Least time** | | Language Functions / Sentence Structure |

/ = equal amount of time spent

#### 1.5.1.3. Expectation concerning keeping the weight of Non-English subjects constant

As mentioned in 1.3.1.3, the weight of Non-English subjects did not increase against the weight of English in the SPE Test. Therefore, I expect while SPE students will study English more than Non-English courses, GE students will study Non-English courses more than English.

### 1.5.2. Expectation concerning the broader curriculum

As the SPE Test is based on a curriculum broader than the school curriculum, I expect GE students will rely on school alone (including school textbooks and school classes), while SPE students will rely on both school and out of school resources (including extra materials and preparation classes). I also expect that there will be a relationship between the learners' test preparation activities and their learning beliefs.

### 1.5.3. Expectations concerning the increase in the difficulty

SPE students will perceive the SPE Test more difficult than GE students will perceive the GE Test.

SPE students will attend English language institutes to develop their English background.

SPE students will engage in activities that are not tested by the SPE Test such as oral activities to develop their general proficiency.

## 1.6. Research questions

At this stage, I was in a position to form my research questions. Considering

the changes in the SPE Test which included an increase in the weight of English, a broader curriculum, and an increase in the difficulty of the SPE, I developed an overarching research question (Sunderland, 2010) as follows:

1. What activities do the SPE students report doing in order to prepare for the SPE Test?

As I was interested to know if the students were doing the activities because of the test or other factors, I posed my second research question as follows:

2. Why do the students report doing these activities in order to prepare for the SPE Test?

As the difficulty level of the test had increased (Table 1.3), I was interested to know how difficult the students perceived the test to be:

3. Do SPE students perceive the SPE Test to be more difficult than the GE Test?

As one of the test authorities had said that they introduced the changes in order to admit more proficient applicants, I asked the following question:

4. Do SPE students have better English backgrounds than GE students?

In 1.3.2, I explained that SPE materials were based on a broader curriculum and were not taught in school, which meant that SPE preparation occurred in a partly independent learning situation. Therefore, I was interested in how the SPE Test was interacting with the students' beliefs, i.e. whether it was affecting the students' beliefs, was being affected by the beliefs or both. To this end, I asked the following research questions:

5. What beliefs do the SPE students report holding about learning English?

6. Are the SPE students' reported activities consistent with their reported beliefs?

## 1.7. Contributions of the study

The questions which I posed in the discussion of the changes in the SPE Test as well as the research questions mean that there is no empirical evidence to answer these questions. Therefore, one of the contributions of this study could be to find answers for practical problems concerning the effect of the SPE Test which could be of use for the authorities in NOET, learners, teachers, parents, teacher training programs, textbook writers and publishers, and educational circles.

This research is also a response to researchers' call for the study of test washback in general and the effect of tests on learning in particular (e.g. Alderson and Wall, 1993; Bailey, 1996; 1999; Cheng et al 2004; Gosa, 2004; Tsagari, 2006; Wall, 2000; 2005).

An important contribution of this study is that it addresses the learners' independence/ autonomy, i.e. their learning criteria or beliefs, which have not been investigated in washback studies so far.

In addition, the results of this study could serve as the basis for further research studies into the effects of tests in the fields of language teaching and testing and general education, particularly with respect to the interaction of tests with learning independence and beliefs. This will hopefully shed more light on how washback works.

In the next two chapters, I will review the relevant literature on the two foci of this study, namely washback and learning beliefs, to see what they have to offer about the gaps in the literature and the research methodology.

# Chapter 2. Literature Review: Part 1, Washback

Chapter 1 has contextualised this study within the area of washback research. This chapter as well as the next chapter will present a literature review in two parts. Chapter 2 concerns washback and Chapter 3 concerns learning beliefs. In general, due to space limitations, I will only deal with issues which are central to this study. In the first part of the literature review I will deal with the concept and definition of the term, *washback*, concepts of positive and negative washback, the Washback Hypothesis, washback models, washback and validity, and some of the most frequently quoted research studies done on washback from the learners' perspective. In the second part, I will discuss the position of learner beliefs in washback studies, the significance of learner beliefs, debate on the stability of beliefs, and some research studies conducted in this area. Finally, I will situate my own research in the study of washback.

## 2.1. The concept and definition of washback

It has been asserted in both the general education and the language education literature that tests have a great deal of impact on teaching and learning (Alderson and Wall, 1993; Wall, 2005). The following statements are some of the typical claims:

'It is generally accepted that public examinations influence the attitudes, behaviour, and motivation of teachers, learners and parents' (Pearson, 1988: 98).

'Tests are held to be powerful determiners of what happens in classrooms' (Alderson and Wall, 1993).

The degree of the influence, however, depends on the stakes of the test, i.e. the higher the stakes of the test, the greater the influence of the test (Alderson and Wall, 1993; Alderson and Hamp-Lyons, 1993; Shohamy et al, 1996; Stecher et al, 2004).

A high-stakes test could have various types of effects ranging from effects on individuals to effects on practices and policies i.e. the classroom, the school, the

education system and society as a whole (Wall, 2005). However, authors use separate terms to refer to the various areas of exam influence. Wall (1997, p. 78) considers 'washback' as a form of 'impact' and uses 'washback' to refer to 'the effects of tests on teaching and learning'. Similarly, Bachman and Palmer (1996: 29-35) discuss exam influence at 'micro-level' and 'macro-level'. By micro-level they mean the effect of examinations on individual students and teachers, and by 'macro-level' they mean the impact on the educational system and society. In addition to the terms, 'washback' and 'impact', other terms have also been used for exam influence: 'measurement-driven instruction' (Popham, 1987), 'curricular alignment' (Madus, 1988; Smith, 1991), 'systemic validity' (Fredericksen and Collins, 1989), and 'backwash', which is used in the same sense as 'washback' by Hughes (1989). The terms 'measurement-driven instruction' and 'curricular alignment' will be discussed in connection with 'positive-negative washback' and 'systemic validity' in connection with 'washback and validity' in the next two sections.

In this study, I will use the term, 'washback' as it is more common than 'backwash' in language testing. As my study is concerned with the effect of the test at the micro-level, I will adopt the following definitions of washback:

'Washback ... is the influence that ... a test will have on the teaching [and learning] that precedes it' (Alderson and Hamp-Lyons, 1996: 280).

Alderson and Wall (1993) define washback as what teachers and learners do because of the test that 'they would not necessarily otherwise do' (p. 117).

## 2.2. Positive and negative washback

Washback has been considered sometimes positive and sometimes negative (Buck, 1988; Heaton, 1990; Bachman and Palmer, 1996; Bailey, 1996; Messick, 1996; Shohamy et al, 1996; Davies et al, 1999). On the positive side, Morris (1972)

considers examinations necessary to ensure that the curriculum is put into effect, and others, using the term 'measurement-driven instruction', claim that tests should drive teaching and hence learning (Bracey, 1987; Frederiksen, 1984; Li, 1990; Popham, 1987; Smith, 1991). Likewise, Popham (1987) claimed that measurement-driven instruction is the most cost-effective way of improving the quality of public education. On the negative side, Madaus (1988) criticized measurement-driven instruction as nothing more than 'psychometric imperialism'. 'Curricular alignment' also, which focuses on the connection between testing and teaching syllabus, has been associated with negative washback (Andrews, 1994; Linn, 1983; Madaus, 1988). It involves narrowing of the curriculum by teaching test-taking skills to students and focusing on topics known to be on the test (Mousavi, 1999). This is believed to lead to test score 'pollution', where teaching to the test would only increase test scores without real improvement of the ability being tested (Haladyna et al. 1991). Shohamy (1993: 186) states that there is negative washback when the means by which the students learn i.e. 'instructional activities, teaching methods, classroom learning, curricula, and textbooks', are compromised.

The notions of positive and negative washback reviewed above look like general statements which may not be very useful for empirical studies. However, Wall's (1999; 2005) criterion for positive-negative washback is match or mismatch between the test and syllabus/textbook in her study of the Sri-Lankan O-Level English Examination. This suggests that in any washback study we should specify exactly what we mean by positive or negative washback.

In this study, the criterion would be match/mismatch between the learners' activities and the learning tips in their textbooks. However, for oral activities, which are neither addressed in the textbooks nor in the test, the learners' doing of the

activities would be considered positive and not doing such activities would be negative.

Given that a test could have positive or negative effects, I discuss in the next section whether the validity of the test should be judged by the extent of positive or negative effects.

## 2.3. Washback and test validity

Some writers have suggested that a test's validity should be evaluated by the degree to which it has had a beneficial influence on teaching. Morrow (1986, cited in Alderson and Wall, 1993) coined the term 'washback validity' presumably meaning that if a test has positive washback, it is valid, and conversely, if it has negative washback, it is invalid. Similarly, Heaton (1990: 16) states that 'If it is a good examination, it will have a useful effect on teaching; if bad, then it will have a damaging effect on teaching'. Frederiksen and Collins (1989) introduce the term 'systemic validity', which they define as follows:

> A systemically valid test is one that induces in the education system curricular and instructional changes that foster the development of the cognitive skills that the test is designed to measure. Evidence for systemic validity would be an improvement in those skills after the test has been in place within the educational system for a period of time (p. 27)

However, Alderson and Wall (1993: 116) state that establishing such a cause-effect relationship is simplistic. Washback is related to the use of the test and when it comes to the use of the test i.e. in a teaching-learning situation, there are many forces, besides the test, that might prevent washback from appearing although the test might be valid by design (ibid). The presence of the many forces makes washback complex and therefore not *directly* relatable to a test's validity (ibid). The authors propose that causes of the teaching-learning practices be explored in order to separate the effects of

the test from the effects of other forces: '... only after we have established causal relationships will we be in a position to explore whether we are justified in relating washback to a test's validity' (p. 117). Messick (1996) expresses a similar idea. Although he includes washback within the 'consequential validity' of a test, he states that, due to the complexity of washback, evidence is needed to relate it to validity: '...washback is a consequence of testing that bears on validity only if it can be evidentially shown to be an effect of the test and not of other forces operative on the educational scene' (p. 242). He suggests achieving validity by design rather than by washback: '...rather than seeking washback as a sign of test validity, seek validity by design as a likely basis for washback'. To produce positive washback, he suggests minimizing construct under-representation and construct-irrelevant variance. In this regard, Hughes (2003: 53-56) makes the following proposals:

1) Test the abilities whose development you want to encourage.

2) Sample widely and unpredictably.

3) Use direct testing.

4) Make testing criterion- referenced.

5) Base achievement tests on objectives.

6) Ensure [that the] test is known and understood by students and teachers.

7) Where necessary provide assistance to teachers.

## 2.4. Theoretical frameworks of washback

### 2.4.1. Washback Hypotheses

Alderson and Wall (1993) argued (in their seminal paper, 'Does washback exist?') that the concept of 'washback' was vague and therefore proposed 15 hypotheses, known as 'Washback Hypothesis' in order to help clarify our thinking on

washback. They began with the most general hypotheses and moved on to the more refined ones.

1. A test will influence teaching.
2. A test will influence learning.
3. A test will influence what teachers teach.
4. A test will influence how teachers teach.
5. A test will influence what learners learn.
6. A test will influence how learners learn.
7. A test will influence the rate and sequence of teaching.
8. A test will influence the rate and sequence of learning.
9. A test will influence the degree and depth of teaching.
10. A test will influence the degree and depth of learning.
11. A test will influence attitudes to the content, method, etc. of teaching and learning.
12. Tests that have important consequences will have washback.
13. Tests that do not have important consequences will have no washback.
14. Tests will have washback on all learners and teachers.
15. Tests will have washback effects for some learners and some teachers, but not for others.

Clarifying the concept of washback through the above hypotheses, the authors have identified areas of washback for empirical research, or in Alderson and Hamp-Lyons's (1996) term, 'laid out the territory', so that it becomes clearer what questions washback studies should address. As Alderson and Wall (1993: 127) proposed that researchers state their own version of the Washback Hypothesis to be used in their study, considering that the focus of my study is washback on learning, the study will

be guided by Hypotheses 2, 5, 6, 8, and 10 which concern learning, and Hypotheses 11, 14, and 15 which concern both learning and teaching.

Since Alderson and Wall's (1993) landmark paper, a number of washback models have been suggested, in particular, Hughes (1994) and Bailey's (1996) models of washback.

### 2.4.2. Hughes's model of washback

Hughes (1994, in Wall 2000) made a distinction between washback on the 'participants,' the 'processes' and the 'products' of an educational system. 'Participants' include teachers, learners, administrators, materials writers and publishers whose attitudes and perceptions may be affected by the test (cited in Bailey, 1999). The term 'process' refers to any actions taken by the participants including materials development, syllabus design, changes in teaching methods or contents, learning and/or test-taking strategies (ibid). 'Product' refers to what is learned and the quality of learning (ibid). Hughes justifies his categories as follows:

> The nature of a test may first affect the perceptions and attitudes of the participants towards their teaching and learning tasks. These perceptions and attitudes in turn may affect what the participants do in carrying out their work (process), including practicing the kind of items that are to be found in the test, which will affect the learning outcomes, the product of that work (cited in Bailey, 1999: 10).

Though at first sight it seems impossible to separate participants from process, Hughes's rationale for his categories is his emphasis on perceptions and attitudes of the participants and how they affect what they do.

### 2.4.3. Bailey's model of washback

Bailey (1996) combined Alderson and Wall's hypotheses and Hughes' model and proposed two categories: 'washback to the learners' and 'washback to the

programme'. For instance, in her 'basic model of washback' (see Figure 2-1), she depicted different groups of 'participants' engaging in various 'processes' which resulted in certain 'products' and stated that Hypotheses 2, 5, 6, 8, and 10 were related to 'washback to the learners' and Hypotheses 1, 3, 4, 7, 9 and 11 were related to 'washback to the programme' (She also depicted 'washforward', to use Van Lier's term, to represent possible influences from the participants on the test). Her rationale for the separation of learners from all other stakeholders was that influences on learners affected their learning directly while influences on other participants affected efforts to promote learning (Bailey 1999: 12).

Regardless of the rationale behind each of the categories in the frameworks above and which frameworks or a combination of them will be used by researchers as their guide, the shared merit of these theories is that they tried to clarify washback and identify washback areas for empirical research.

Clarifying the concept of washback, Alderson and Wall (1993) also urged researchers to do empirical studies to find evidence for the existence of washback rather than just asserting its existence. They recommended classroom observations which had not been used to date as well as methodology triangulation and data triangulation, a more ethnographic approach, a widening of hypothesis formulation and taking account of findings in at least two areas, namely motivation and innovation in educational settings. Alderson and Wall having set the agenda, various studies have been conducted on washback which I discuss in 2.5. Since the focus of this study is washback on learning, I will only discuss studies which have been done in this area and for other areas such as washback on teaching and washback on learning materials, the reader is referred to encyclopedia entries by Wall (1997), Cheng (2007), Bailey (1999) and Tsagari (2006).

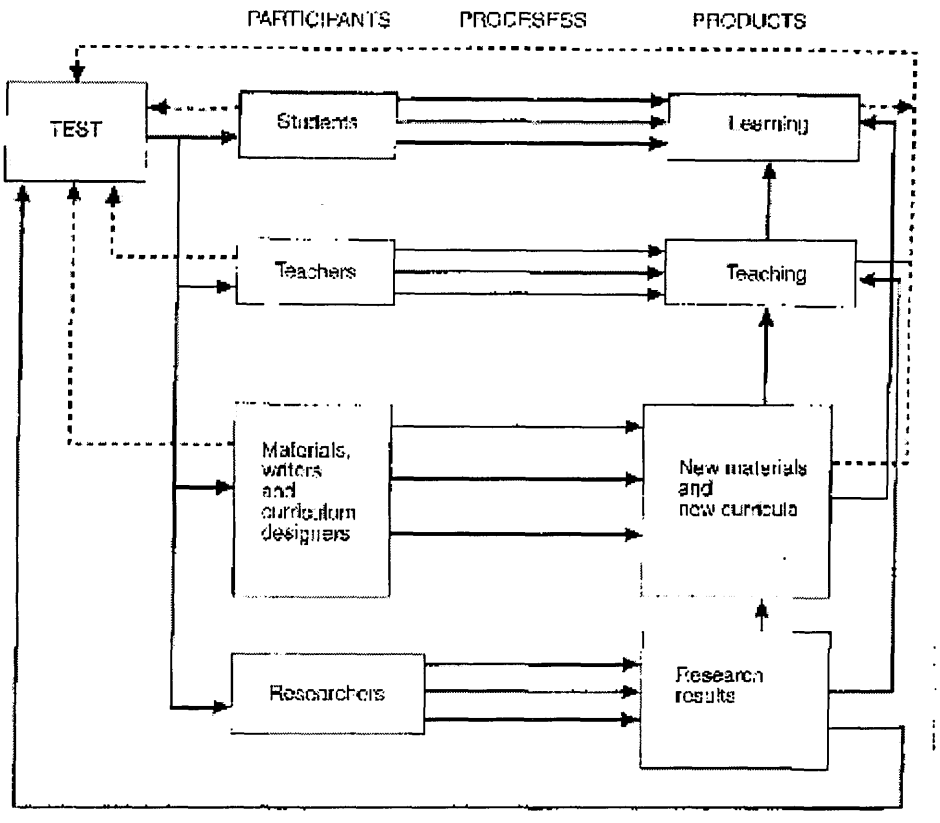


**Figure 2-1 Basic Model of Washback (Bailey, 1996, 1999)**

## 2.5. Washback studies on learning

Using SILL (Strategy Inventory of Language Learning) (Oxford, 1999), Watanabe (1992) investigated possible differences in learning strategy use of those students who entered college through entrance examinations and those who entered college through recommendations. He examined the students in their first and second year at college. The results showed that college entrance examinations had a positive effect in that the exam students used more learning strategies than the recommended groups. He speculated that in addition to the exam, language proficiency and motivation might have been the variables resulting in the wider range of strategy use. He further speculated that it might be due to the fact that the exam students had more chances to learn various strategies in supplementary preparatory classes. However, the

study showed that strategy use did not change over the two years. He thought that the reason might be because the second year students might not have been motivated by the college education to employ a wider range of strategies. He also stated that the reason might be that the washback effects of the examination drove students to learn EFL only in order to pass the examination. This argument was based on Saljo (1979), which suggested that students do not change their learning strategies until they change their conception about and purpose of learning.

The relevance of this study to my study is that I will investigate the learning methods of the students and will also address English background, motivation, attendance in preparation classes and methods of learning.

Using questionnaires, Cheng (1998) studied the impact of the new school-leaving English examination in Hong Kong by giving questionnaires to two groups of students. One group, which was the control group, was preparing for the old test in 1994 and one group, which was the main group, was preparing for the new test in 1995. The new test was integrated, task-based and process-oriented. She investigated students' perceptions of and attitudes towards teacher talk (to determine practice opportunities), teaching activities, learning activities, use of English inside and outside class, attitudes towards English lessons, motivation (i.e. reasons) to learn English, learning strategies, the effect of the test on anxiety, and the effect of aspects of public exams on students themselves and their learning processes. She argued that there was no fundamental impact on students' learning, for example with respect to their motivation to learn English or their learning strategies, and that classroom activities were similar to those designed in the new examination. She asserted that a change in the examination syllabus alone would not fulfil the intended goal and that what actually was taught and learned was far more complicated and involved more than the

examination and the school curriculum.

As far as my study is concerned, like Cheng's study, I will use questionnaires to address groups of learners who will be SPE test-takers and GE test-takers so that the effect of the SPE test can be separated from the GE test. I will also examine English learning outside school and teaching from the perspectives of the students.

Shohamy et al (1996) examined the impact of two national tests – Arabic as a second language test (ASL) and English as a foreign1anguage test (EFL) – to see whether the impact the tests had in 1993 still persisted. The researchers examined students' perspectives through questionnaires. The results showed that in each of the two tests, the impact found in the 1993 study differed from the impact of the 1996 study. While the washback effect of the ASL test significantly decreased over the years, the impact of the EFL test increased.

The impact of the ASL test decreased to the extent that it had no effect. Moreover, the students expressed negative attitudes that the test was unimportant and of poor quality. On the contrary, the increased washback of EFL included numerous oral teaching activities in the classroom, an increase in time allotment, generation of much new courseware, a high awareness of the test and a significant increase in the status of the subject-matter. However, the students expressed high anxiety and believed that the test results would affect their success in future studies. The researcher concluded that washback could change over time and depended on a number of factors including whether the stakes of the test and the status of the language were high or low.

The results of this study are relevant to my study in that I will highlight the changes in the students' attitudes at different points of time, although over a shorter span of time.

Wesdorp (1982) studied the effect of multiple-choice tests to investigate the validity of the objections to their use in Dutch schools. The results showed that there was no difference in the amount of writing practice and the quality of essays written in 1966 (when multiple-choice tests had not been introduced yet) and 1978. In terms of study habits, the results showed no difference between students who used multiple-choice reading comprehension tests and students who used other forms of tests. Neither group did much preparation for reading anyway. The results also showed that contrary to teachers' beliefs that multiple-choice tests had a negative emotional influence on the students, in general the students did not have negative attitudes towards them.

Wesdorp's study is relevant to my study in that the test under investigation in my study is a multiple-choice test. Therefore, it would be interesting to see whether the amount of time the students spend practising each test section is related to the test method. I will also examine possible differences between the teachers' (indirect) reports of their activities and students'.

Alderson and Hamp-Lyons (1996), in their exploratory interviews with groups of TOEFL learners, asked them how they would like their TOEFL preparation classes to be taught in comparison with what they had already experienced and what the best ways of preparing for TOEFL might be. The learners' expectations included having a placement test before a TOEFL preparation course, more opportunities for participation and questioning, diagnosis of their weaknesses, and combination of self-study and revision in class. However, the researchers stated that none of these were happening in any of the classes they observed. Concerning the best ways of preparing for TOEFL, the learners believed that 'having American friends', 'going to the movies', 'reading a lot', and 'using English outside class' were ideal methods of

preparation. However, the researchers stated that these ideas did not appear in the teachers' interviews at all.

The relevance of this study to my study is, as mentioned above, that I will see what criteria or beliefs the students have for appropriate ways of learning and will also consider possible differences between the teachers' views and the learners' views in the design of my data collection instruments and data analysis.

Watanabe (2001) investigated the relationship between Japanese university students' test preparation and their motivation through interviews. He found that the way the test affected motivation depended on the importance of the university/exam and the difficulty of the exam. Students tried harder for the sections of the exam which were related to the university of their first choice than the university of their second or third choice. The sections of the exam which were more difficult and thus more discriminating had a greater motivating power than those sections which were less difficult and less discriminating. However, the students did not bother to study for those sections which they perceived to be too difficult. The author concludes that washback is caused by a complex interplay between the test and the test-taker and that it is not the test alone that causes washback. He suggests that what counts is not the objective difficulty of the test but the learners' perception of the difficulty that may cause washback.

As Watanabe's study emphasises the role of motivation and difficulty of the test in producing washback, I will also collect data on students' motivation and test difficulty.

Ferman (2004) examined the washback effects of a national EFL oral matriculation test on Israeli high school students. Structured questionnaires were given to 120 students from three different types of schools and three ability levels.

Test washback was examined with regard to teaching-learning focus, time allotment, teaching strategies, learning strategies, promotion of language skills, parental involvement and anxiety. The test resulted in both positive and negative washback. The positive washback was an increased focus on the oral skills, which resulted in increased time allotment, accelerated pace of learning and employment of learning strategies towards test success. The negative washback was narrowing of the scope and content of learning, a high level of anxiety, fear of test results and memorisation. The test resulted in differential washback among learners. The author suggested that to ensure the desired washback, a variety of factors, particularly individual differences, should be taken into account and that a detailed examination of the educational context was necessary.

To consider individual differences in my study, I will sample students based on their English background and from different types of schools.

Andrews et al (2002) report a study on the effects of the changes made to the use of English (UE) oral examination on the performance of the students in their final school year in Hong Kong. The study had two phases: one discussed in Andrews and Fullilove (1997) and one in Wong (2001). The authors gathered data through a neutral testing instrument and administered it over a 3-year period to consider the effect of time as well. The results showed that the test did influence the students' performance, though not necessarily as predicted or intended, and that the washback effect was delayed. There was also evidence for the mediation of published materials, rote learning and memorization. The students learned which language features to use but not where or how to use them appropriately. The authors concluded that it was relatively easy to create changes in the content of teaching or learning and the time allocated, but that changes on teaching and learning were indirect and unpredictable

and depended on factors such as time, published materials, and individual differences among teachers and among learners.

Based on Andrews et al's study, there are two issues that I would address in my study: what types of materials the students will use and why, and whether they will use memorization as a learning method and why.

One aspect of Tsagari's (2006) study was to look at learners through the students' diaries. The study was set in a Greek private language school involving teenage students preparing for a high-stakes EFL test (First Certificate in English – Cambridge ESOL) over approximately 7 months. This study brought to light the influence of the high-stakes exams on a wide range of areas including students' feelings (stress, anxiety), motivation for learning English, perceptions, attitudes towards the content and methods of teaching and classroom assessment, reactions to an intensification of exam preparation, and the role of the exam textbook, teachers, parents, school and local society in the washback process.

The researcher showed that a wide variety of factors were involved in the washback process partly due to the type of data collection instrument where she did not impose any pre-determined categories for the students to write about. This has relevance in my study in that I will ask the students to recommend any appropriate method of test preparation to their friends.

Luxia (2005) used questionnaires and interviews to investigate the effect of the National Matriculation English Test (NMET), which was designed primarily to select candidates for tertiary education in China and to make changes in (i.e. to direct) English language teaching (ELT) in schools as well. The results showed that the NMET failed to bring about the intended washback. She examined the reasons for this failure and concluded that the two functions of the test were in conflict with each

other, making the test ineffective for changing teaching and learning. Since the primary role of the NMET was for university admission and was used for a large number of examinees, the test constructors had to work for high reliability, strong discrimination, and utmost fairness, all of which imposed constraints on the directing function of the test. The test constructors were restricted in terms of the choice of the test format (i.e. multiple-choice format) and the choice of the materials to be included in the test. The selection purpose of the test gave rise to another purpose which involved evaluation of teachers and schools based on the students' performance on the NMET. This misuse of the test scores, in turn, put pressure on the teachers and schools to strive for the immediate goal of high scores rather than the long-term goal of developing the ability to use English. As a result, the focus of teaching-learning was on test content, test format, abundant use of mock tests and in general on linguistic form rather than language use. The various communicative activities in the textbooks were also ignored. Among the factors which contributed to the failure of the NMET was the difference in views of the constructs between the test constructors and the test users. While the test users believed grammar and vocabulary were the main constructs, the test constructors believed that linguistic knowledge constituted only a small part of the NMET. Another reason why the learners were not exposed to real-life language use activities was the teachers' lack of experience in authentic language use in an English-speaking environment.

What this study has in common with my study is that the tests under investigation in both studies have been used for selection purposes. Another common feature of the two tests is that they have been used to select from a large population. Therefore, it would be interesting to see in my study what kinds of effects the SPE Test will have because of its primary function of selection.

Gosa (2004) studied the washback effects of the English component of the Romanian school-leaving exam (Bac) through students' diaries. She examined how the students experienced the teaching and learning of English inside and outside the classroom. The washback effects she identified were in terms of reference to the exam, content/topical information in the lessons, materials, teaching techniques, types of exercises/tasks employed, class atmosphere, attitudes to the lesson/contents/ methods, reference to private classes, learning strategies, and attitudes to doing homework and private study. The analyses of the diaries showed that although students expected washback in the classroom, they experienced very little, but their personal environment was affected to a great extent. The researcher concluded that students' expectations, feelings, attitudes, perceptions, beliefs, learning styles, motivation and anxiety should be taken into account when promoting positive washback as they were likely to interact with the effect of the test itself.

As this study addressed learning outside the classroom environment, I will also address students' self- study activities in my study.

Wall (1999) reports her study of the Sri-Lankan O-Level Examination in which she used innovation theory, i.e. Henrichsen's (1989) Hybrid Model of Diffusion /Implementation Process (see also Wall, 2005). Having done a baseline study of the 'antecedent' conditions i.e. before the introduction of the test, she tried to identify factors which hindered or facilitated the implementation 'process' of the innovation. To this end, she gathered data about the students as one of the 'users' of the innovation. The data gathered concerned the students' attitudes and their characteristics. Attitudes included those towards education, classroom teaching, language teaching, examinations, English, and attitudes towards new ideas (openness), and the data about the learners' characteristics included their levels of education,

abilities, personal life, economic situation, interests, and goals (p. 605).

However, due to some practical problems, it was not possible to talk to the students directly, and so she obtained 'second-hand information' by talking to the teachers (p. 648). Nevertheless, the author was aware that the teachers' views might not truly reflect what the students thought and felt and suggested that the learners be consulted directly in future studies (p. 723).

The results (from the teachers' perspectives) showed that many students considered education instrumental for a better future and were most motivated to study during the exam preparation period, while other students saw no point in education altogether; students differed in their interest in various topics; some students wanted to study more grammar, either for its use in language learning in general or for its importance in the examination; some students preferred teacher-centred tuition and spoon-feeding; some students needed more encouragement than others; students were different for social, economic and political reasons and due to their personal life in their abilities, attitudes towards studying, and learning activities. In the end, the author concluded that these different attitudes and characteristics had their own effects on the implementation process of the new test.

As Wall (1999) suggested direct access to the learners, in this study I will talk to the learners directly. Another feature of Wall's study was that she compared the situation before the introduction of the new test with the situation after the test was put to use. Similarly, I will have a contrast group preparing for the GE Test to compare it with the SPE test takers. Concerning learning beliefs, the author took account of the students' beliefs about learning grammar in general as well as learning grammar for the exam. In my study also, I will address both kinds of beliefs and will examine whether there is consistency between them. Another relevance of this study is that the

author used a qualitative method of data collection rather than a structured instrument with pre-determined categories. I thought it would be interesting to see in my study what other things the students have to say in addition to the categories addressed in the questionnaire.

**Table 2-1 Summary of individual washback studies on learners**

| Authors | Instruments | Main findings |
|---|---|---|
| Alderson and Hamp-Lyons (1996) | Interview | Learners had certain expectations about their TOEFL classes which were not happening in any of the classes observed. Students' criteria for the best ways of learning were not the same as the teachers'. |
| Andrews et al (2002) | Testing instrument | Washback effect was delayed, indirect and unpredictable and depended on factors such as published materials and individual differences among teachers and among learners. |
| Cheng (1998) | Questionnaire | The impact of test on learning was superficial in that learning activities were only similar to those designed in the new test. Washback was too complicated to be simply a matter of change in the exam and school curriculum. |
| Ferman (2004) | Questionnaire | To ensure the desired washback, a variety of factors particularly individual differences should be taken into account and a detailed examination of the educational context was necessary. |
| Gosa (2004) | Students' diaries | The personal environments of the students were affected more than their classroom environments. The researcher concluded that students' expectations, feelings, attitudes, perceptions, beliefs, learning styles, motivation and anxiety should be taken into account when promoting positive washback as they were likely to interact with the effect of the test. |
| Luxia (2005) | Questionnaire and interview | Multiple uses of the test, particularly the selection and direction functions of the test, were in conflict with each other making the test ineffective for changing teaching and learning. |
| Shohamy et al (1996) | Questionnaire | Washback could change over time and depended on a number of factors including whether the stakes of the test and the status of the language were high or low. |
| Tsagari (2006) | Students' diaries | This study brought to light the influence of high-stakes exams on a wide range of areas including students' feelings (stress, anxiety), motivation, perceptions, attitudes towards the content and methods of teaching and classroom assessment, reactions to an intensification of exam preparation, and the role of exam textbooks, teachers, parents, school and local society. |
| Wall (1999; 2005) | Interview | In general, positive attitudes (e.g. towards education, English, new ideas, etc), and positive characteristics (e.g. good ability in English, good economic situation, etc) facilitated the implementation of the new test and negative attitudes and characteristics hindered the implementation. |
| Watanabe (2001) | Interview | Effect of test on motivation depended on the learners' perception of the degree of the difficulty of the test sections. |
| Watanabe (1992) | Questionnaire (SILL) | Test affected the extent of learning strategy use but contextual factors such as proficiency, motivation, previous education, passage of time, college education, and purpose of learning interacted with washback. |
| Wesdorp (1982) | Testing instrument | There was no difference between the effect of multiple-choice tests and other forms of tests on the students' learning. Teachers' impression of the students' attitudes towards multiple-choice tests did not correspond with the actual attitudes of the students. |

## 2.6. Summary of the empirical studies combined

- The majority of the studies reviewed so far were based on self-report data rather than observed behavior i.e. mainly perceptions and attitudes were addressed.

- The data collection instruments included questionnaires, interviews, diaries, and a neutral testing instrument. Wall's report of the Sri-Lankan study was the only one which used a triangulation of observation and interviews.

- The effects of the tests were examined through the eyes of the learners on teaching, and learning in general, and in particular on content and methods of teaching, teacher talk, teaching-learning focus, time allotment, teaching strategies, classroom assessment, learning strategies, learning styles, content of learning, skills, materials, tasks, and use of English.

- Different contexts including classroom, out-of-class, and different types of school were examined.

- Feelings including motivation and anxiety were addressed. The effect of tests on motivation was examined along with the role of the importance of the test and the difficulty of the test.

- Individual differences such as language proficiency, purposes of the learners, etc were also examined in some studies.

- The multiplicity of washback factors such as the role of the exam textbook, published materials, parents, school, and local society was also considered.

- Some studies used baseline studies and compared the two situations before and after the introduction of the new test.

- Unintended or unpredicted washback was also studied.

- Use and misuse of the test and the consequent washback effects were examined in one study.

- One study was conducted longitudinally to examine the effect of time.

- There was one study which investigated the effects of the test on the product of learning or performance.

- While the majority of the studies examined the effects of the test on the students' test preparation, one study examined the effect of the test after it was taken by the students.

Clearly, the studies above have considered both the effect of the test as well as factors other than the nature and design of the test that exist in the context and may influence learning. They include affective and cognitive factors as well as the educational system and the society. However, despite the fact that we have gained a considerable knowledge on washback on learning through the above research studies, the literature urges washback studies on learning and students' perceptions of test preparation. According to Watanabe (2004), washback on learning is not well explored yet. Wall (2000) states, 'we know very little about students' perceptions of tests and much less about how new tests influence what students know and can do' (p. 506). Hamp-Lyons (1997:299) argues that 'it is not enough to evaluate tests from our own perspectives; neither is it enough to evaluate them by including teachers' perspectives'. She suggests 'many more studies are needed of students' views and their accounts of the effects on their lives of test preparation, test-taking and the scores they have received on tests' (p. 299). What is more noticeably missing from the literature is that the studies carried out so far have not examined how the test might interact with learners' independence / autonomy, i.e. learning beliefs / learning criteria. Although Gosa (2004) talks about beliefs, this is not her main focus. Therefore, the focus of this study will be on the washback effect of the SPE Test on the students' learning activities and its interaction with the students' learning beliefs.

In the next chapter, I will present a literature review on learning beliefs.

# Chapter 3. Literature Review: Part 2, Learning Beliefs

In Chapter 2, I examined the washback literature and concluded that no washback study addressed learner beliefs systematically. In this chapter, I present a literature review on learner beliefs to locate them in washback studies and to examine what beliefs are and how they are investigated.

## 3.1. Locating learner beliefs in washback studies

In theory, learning (and teaching) beliefs have already been addressed in the washback literature, although not very much empirically, systematically, or as a main focus. Washback Hypothesis 11 (Alderson and Wall, 1993) addresses beliefs with the inclusive term 'attitudes': 'A test will influence attitudes to the content, method, etc of (teaching and) learning'. Reber and Reber (2001: 63), in their definition of the term attitude, subsume 'beliefs' under 'attitudes'. They state that contemporary usage of the term includes several components, namely 'cognitive (consciously held belief or opinion), affective (emotional tone or feeling), evaluative (positive or negative), and conative (disposition for action)'. However, they believe that there is no consensus as to which component or components should be considered important and that this depends on the theoretical orientation of the researcher. For example, they continue, to a cognitive theorist, the underlying belief may be the most important and to a behaviourally oriented theorist the conative component may be fundamental. They define belief as '... an emotional acceptance of some proposition, statement or doctrine' (p. 86). Colman (2001; 2006) also sees attitudes as including those four components. However, he seems to be giving more prominence to the evaluative aspect: 'An enduring pattern of evaluative responses to a person, object, or issue' (p. 63). He defines belief as 'Any proposition that is accepted as true on the basis of inconclusive evidence. More generally, belief is conviction, faith, or confidence in

something or someone' (p. 85). Similarly, VandenBos (2007: 112) defines beliefs as 'acceptance of the truth, reality, or validity of something (e.g., a phenomenon, a person's veracity, a theory)'.

Furthermore, Wall and Alderson (1993) identified teacher beliefs as an important factor and Alderson (1998) suggested exploring teacher cognition and teacher thinking to understand why teachers do what they do. Therefore, it seems plausible to extend this suggestion to learners and explore their cognition.

Cotterall (1995: 195) links learner beliefs to autonomy. She states that autonomy involves readiness for changes in beliefs and behavior. Bailey (1996; 1999) discussed the relationship between washback and learner autonomy. She suggested a number of factors believed to promote beneficial washback: they included incorporation of language learning goals, authenticity, learner autonomy and self-assessment, and detailed score reporting. Concerning the philosophy of learner autonomy, she went on to say that 'students should have a large amount to say about what, how, and how fast they learn' (which, in this case, means that they have already been addressed by Washback Hypotheses 5, 6, and 8). Learner autonomy means that students develop their own internal values and take greater control over the content and methods of learning (Bailey, 1996; 1999; Holec, 1981, Chan et al, 2002). This ownership and self-regulation are likely to result in greater locus of control and deeper processing of the material at hand (Bailey, 1996; 1999).

However, Bailey (1996) says there is relatively little empirical research to substantiate the claim that learner autonomy contributes to positive washback (p. 272).

## 3.2. Significance of learner beliefs

Cotterall (1995) states that the reason for the importance of beliefs and attitudes is their profound influence on learning behaviour (see also Goh, 1997;

Horwitz, 1987; Victori, 1992; Wenden, 1991; 1999). According to Benson and Lor (1999), learning attitudes and behaviours are conditioned by beliefs about the nature of language and language learning (p. 459). If learners believe that the best way to learn a new language is to memorise its components, it is likely that they will hold positive attitudes towards learning vocabulary and grammar and using strategies that involve analysis, memorisation and practice (p. 459). If learners believe that the best way to learn a foreign language is to use it in its natural contexts, it is likely that they will be predisposed to communicating with native speakers and adopting a range of social and communication strategies (ibid). Considering that certain attitudes and behaviours may be more enabling than others (ibid), it is worth examining what attitudes and behaviours tests induce in the learners so that we can address the underlying beliefs in future tests and hence in the learners.

Having established a relationship between washback and learner beliefs, and the importance of beliefs in language learning, I will now explore types of learner beliefs or what they consist of. This will inform my study in terms of what types of beliefs should be addressed.

## 3.3. Types and components of learner beliefs

It has proved difficult to identify and classify learning beliefs in a systematic way (Benson and Lor, 1999: 460). However, despite the difficulty, there have been attempts to categorise them. Adapting Flavell's (1979, 1981) taxonomy for second language learning, Wenden (2001; 1999) states that there are three types of learner beliefs (or metacognitive knowledge, used interchangeably by Wenden, 1999; 2001; Yang, 1999; Kalaja, 1995): *person knowledge*, *task knowledge* and *strategic knowledge* (see also 3.4.2.1). *Person knowledge* refers to the laws of human learning or knowledge about how cognitive and affective factors may facilitate or inhibit

learning in general. These factors distinguish learners from one another and affect their success in language learning. This kind of knowledge includes age, language aptitude, personality (e.g. self-esteem and anxiety), motivation (e.g. purpose and interest), socio-cultural factors (e.g. attitudes toward the culture or language), cognitive style and learning style. Person knowledge also includes knowledge about how these factors apply in learners' own experience. For example, do learners consider their age to be an obstacle or a facilitator of their learning? *Task knowledge* refers to knowledge about the purpose and significance of the task (improving vocabulary, passing an exam, etc), the type or nature of the task (knowledge about the nature of language e.g. is language learning the same as learning biology? What is particular to the spoken language and the written language?), the demands of the task (what are necessary to complete the task? How should the task be done? Is it hard or easy?) and when deliberate learning is required (conscious effort is required when the task is new, when it requires conscious thinking e.g. writing, when it requires accuracy and when it is not been learnt appropriately). Finally *strategic knowledge* refers to knowledge about what strategies are, when to use them, how to use them and what the most effective or best strategies are, in other words, how best to approach language learning. Wenden also states that these kinds of knowledge are prerequisite to two key self-regulatory learning processes: task analysis and monitoring.

However, to my knowledge, only a few researchers have used the categories above, for example Wenden (1991; 2001), Victori and Lockhart (1995), Goh and Lin (1999), and Victori (1999). Instead, researchers have addressed issues of their own interest without particular concern for categorisation.

In the following sections, I will discuss approaches to the study of learner beliefs, review empirical studies in this area, and categorise them based on their

approach. Finally, I will discuss how my review of the literature on washback and learner beliefs informed my study.

## 3.4. Approaches to the study of learner beliefs

Barcelos (2003) identifies three approaches and categorises them based on the definition of beliefs which they offer, their methodology, the way they see the relationship between beliefs and actions, and finally their advantages and disadvantages. She says that these approaches can also be looked at on a continuum of studies based on grounded theory and studies based on a priori categories. She labels these approaches as normative, metacognitive, and contextual.

### 3.4.1. The Normative Approach

#### 3.4.1.1. Definition of beliefs

Most of the studies within the normative approach see beliefs as 'preconceived notions, myths or misconceptions' (Horwitz, 1988: 119). This implies that students' ideas are considered to be radically different from those of second language scholars. In other words, students' opinions about language learning are thought to be wrong and the scholars' opinions are thought to be right. This approach connects beliefs with autonomy and good language learning. Learners are compared with the ideal or autonomous language learner and their beliefs are considered as obstacles to autonomous language learning.

However, Riley (1997) criticizes stigmatizing students' ideas as wrong or as obstacles. He argues that the point is not to find '*the* truth' but students' subjective reality or '*their* truth' because it is the students' own beliefs that affect their learning more than anybody else's (p. 127). This argument recognizes the importance of context and the conditions in which learning takes place and means that not only should we focus on what beliefs the students have but also why they have these

beliefs.

#### 3.4.1.2. Relationship between beliefs and actions

Most of the studies within this approach have established a simple cause-effect relationship between beliefs and actions (Barcelos, 2003: 15). They have generally described and classified learner beliefs and have made assumptions about how they could affect learners' behavior (ibid). They have assumed that productive beliefs will lead to successful behaviour and erroneous beliefs to unsuccessful behavior (ibid). However the relationship between beliefs and behaviour is much more complex and depends on factors such as students' learning experiences in the past, teachers' approach to teaching, teaching materials, students' proficiency levels, motivation, contexts etc. For example in a test preparation situation, time limitation might force the students to use a bilingual dictionary rather than a monolingual one in order to save time, although they may believe a monolingual dictionary is more useful. In this case, their beliefs and behaviour are affected by their unique test preparation conditions.

In addition, the relationship between beliefs and actions is interactive and dynamic in the sense that beliefs both shape behaviour and are shaped by it. In the case of the above example, if the students really save time by using a bilingual dictionary, then the belief that they should use a bilingual dictionary is probably reinforced.

#### 3.4.1.3. Methodology

Likert-type questionnaires have been used for data collection and descriptive statistics for data analysis in this approach. The most widely used questionnaire is the Beliefs About Language Learning Inventory (BALLI) developed by Horwitz (1985; 1987). Other studies either adapted or modified it or developed their own

questionnaires. In some cases, other data collection techniques such as interviews were also used. However, the purpose of using these techniques was only to validate the questionnaires (for example, Cotterall, 1999, and Sakui and Gaies, 1999).

### 3.4.2. The Metacognitive Approach

#### 3.4.2.1. Definition of beliefs

The metacognitive approach is mainly advocated by Wenden (for example Wenden, 2001 and Wenden, 1999). She uses the term *metacognitive knowledge*, interchangeably with learner beliefs and states that research findings on the metacognitive knowledge of language learners are more frequently referred to as beliefs. She sees beliefs as a 'specialized portion of a learner's acquired knowledge base' (Flavell, 1979) and that part of long-term memory that contains what learners know about learning (Wenden, 2001). She sees beliefs as 'a sort of logic' that determines what learners do in order to learn. Although she considers beliefs as 'stable', she also says that it is possible that they may change over time. This knowledge may be acquired both unconsciously and consciously. It may be the outcome of observation and imitation (unconscious), of what teachers or parents tell learners about how to learn, or of learners' own reflection on the process (conscious). Learners' beliefs consist of 'a system of related ideas', some of which they accept without question and some of which they validate by their experience (Wenden, 1999: 436).

As mentioned in 3.3, Wenden (2001) draws on Flavell's (1979, 1981) categories and states that there are three types of metacognitive knowledge: person knowledge, task knowledge and strategic knowledge. Person knowledge refers to the knowledge learners have acquired about how cognitive and affective factors, such as age, language aptitude, personality, and motivation may affect learning in general and

in their own specific case. It also includes the knowledge learners have acquired about their proficiency in a given area. Task knowledge refers to knowledge about the purpose of the task, the type of the task and the demands of the task. Finally strategic knowledge refers to knowledge about what strategies are, when to use them, how to use them and what the most effective or best strategies are.

#### 3.4.2.2. Relationship between beliefs and actions

This approach, like the normative approach, sees metacognitive knowledge as essential in helping students to become autonomous learners. The three categories of metacognitive knowledge (person, task and strategic knowledge) mentioned above are prerequisite to two key self-regulatory learning processes: task analysis and monitoring (Wenden, 1999, 2001). However, as in the normative approach, still there is a cause-effect relationship posited between beliefs and actions (Barcelos, 2003: 19; Wenden, 1998: 522). The relationship is considered such that certain beliefs will lead to successful strategies or autonomous behaviour and certain others will lead to unsuccessful strategies ((Barcelos, 2003: 19).

#### 3.4.2.3. Methodology

The type of data collected within the metacognitive approach is verbal accounts gathered through semi-structured interviews and self-reports. This is because Wenden (2001) sees the three types of knowledge above as 'statable'. The basic assumption in this approach is that learners do think about their language learning process and are able to articulate some of their beliefs and bring them to consciousness. For example, Wenden (1987) showed that some of the 'stated' beliefs in her study were very different from those of the BALLI and that some others were not represented at all in the questionnaire. She concluded that a more comprehensive and representative set of beliefs needed to be developed.

One of the advantages that this approach offers is that the interview provides more in-depth information about metacognitive knowledge than questionnaires. It also gives the learners a chance to elaborate and reflect on their learning experience. However in this approach beliefs are not inferred from actions but only from what the learners say. In addition, although beliefs are connected to experience, learners are categorized into good and bad learners without taking into consideration the influence of context on the learners' beliefs (Barcelos, 2003: 19). For example it may not be justified to label as bad learner a bilingual dictionary user who uses this kind of dictionary because of time pressure or learns the meanings of words without spending enough time on using them.

In addition to the interview, a few studies may have used questionnaires, but none of them have used the BALLI.

### 3.4.3. The Contextual Approach

#### 3.4.3.1. Definition of beliefs

This approach sees beliefs from different perspectives and defines it according to various theoretical frameworks. Beliefs are characterized as social, dynamic, interactive and as embedded in students' contexts where each interaction in the context modifies the existing context and creates a new situation for the subsequent interaction. Beliefs are considered part of learners' experiences and interrelated with their environment. They are seen as 'situationally conditioned' (Sakui & Gaies, 1999) and 'relational and responsive to context' (Benson & Lor, 1999).

#### 3.4.3.2. Relationship between beliefs and actions

In the preceding sections we saw that a simple cause-effect relationship was established in the normative and metacognitive approach between beliefs and actions. Beliefs were labeled as right or wrong and positive beliefs were predictors of

successful behaviour and negative beliefs, predictors of unsuccessful behaviour. However, the contextual approach presents a less 'positivist' view of learners by seeing them as social beings interacting in their environment. As it is believed in this approach that knowledge or thinking is situated, it is important to investigate the context in which students interact. The studies in which classroom observation is used try to understand the relationship between beliefs and behaviour within the complexities of the specific contexts.

#### 3.4.3.3. Methodology

The studies within this heterogeneous approach have collected various types of data and have used various data analysis techniques. These studies have not aimed at making generalizations about beliefs but at a deeper understanding of beliefs in specific contexts. They use triangulation in order to take into account the students' own perspectives, i.e. to interpret students' beliefs in their contexts. The variety of methods used in this approach include ethnographic classroom observation (Allen, 1996; Barcelos, 2000), diaries and narratives (Miller & Ginsberg, 1995), metaphor analysis (Ellis, 2001), discourse analysis (Riley, 1994; Kalaja, 1995), case studies (Barcelos, 2003), biographies, phenomenography (White, 1999), and the discursive approach (Kalaja, 2003). These methods are grounded in students' own interpretative meanings and perspectives. Context which means learners' construction of their experiences is crucial to this type of analysis.

The contextual approach offers a number of advantages. It defines beliefs about SLA as dynamic, social, and interactive and takes into account the influence of environment and experience. It uses a variety of methodologies which come from qualitative and interpretative paradigms. However, the problem with this approach is that techniques such as classroom observation, metaphor, and discourse analysis are

very time consuming. Nevertheless, depending on factors such as the purpose of the study (to generate hypotheses or to conduct a case study), the number of subjects, the resources available, time constraints, the context in which learning takes place, the researcher may consider using any of these methods (Cohen, 1998).

## 3.5. Research studies on learner beliefs

In this section, I will review some of the studies which have been done on beliefs and will use the framework proposed by Barcelos (2003) to categorise them.

To identify cultural differences in learner beliefs, Horwitz (1999) reviewed several studies which had been conducted in various contexts. These studies had examined language learning beliefs under the categories of 'the difficulties of language learning', 'foreign language aptitude', 'the nature of language learning', 'learning and communication strategies', and 'motivations and expectations', which formed the 34 items of the BALLI (Beliefs About Language Learning Inventory) questionnaire outlined in Horwitz (1987). The review revealed that belief differences among participants of the same cultural background might account for as much variation as the cultural differences. The differences identified within the groups were more clearly attributable to differences in learning circumstances than cultural differences. The author argued that it was entirely possible that differences such as age, stage of life or differences in the language learning context such as specific classroom practices might contribute to within-group variation. A difference that was common was the difference in motivational issues such as 'wanting to get to know native speakers better' and 'expecting better job opportunities' with increased language proficiency. There was also a large amount of similarity across the beliefs which could not be ignored. This was attributed to the possibility of the existence of a world culture of language learning or some other shared characteristics such as the age

of the learners. The results suggested that language learning context should be considered as one of the sources of variation in the responses, for example the different instructional approaches or the particular nature of the language learned or taught. However, in this study conclusions about differences and similarities were based on only frequencies. Therefore, it is not clear whether the differences and similarities were also statistically significant.

Nevertheless, the significance of Horwitz's study to my study is that I will use the BALLI questionnaire and will take into account the learning circumstances of the students.

Davis (2003) addressed the similarities and differences between teachers' and students' conceptions of the nature and methods of language learning through a Likert scale questionnaire. Ten dimensions of language learning based on language learning theories were drawn from Lightbown and Spada (1993) and were included in the questionnaire. The subjects of the study were 18 full-time teachers and 97 students following Chinese-English translation courses. The results showed that teachers' beliefs corresponded largely with contemporary language learning theories for all the 10 statements, while the students' thinking corresponded with the theory for 6 of the items and differed substantially on four statements. Students sought a more structured and methodical approach than their teachers and had a positive view of being corrected when they made mistakes. Students' beliefs were stronger than their teachers' on statements which related to the language learning theory of behaviourism.

Davis's study is relevant to my study in that I will examine what positive beliefs and what negative beliefs the students have.

Yang's (1999) study addressed the relationship between language learning beliefs and learning strategy use. He used an English Learning Questionnaire which

was composed of the BALLI, SILL (Strategy Inventory for Language Learning by Oxford, 1990) and two open ended questions in order to collect data on the learners' beliefs, learning strategies and individual background. He used the two open-ended questions to elicit additional beliefs and strategy use. The subjects of the study were 505 Taiwanese undergraduate students. The results showed that the students had some conflicting beliefs which were reflected in their use of strategies. For example while they thought that it was necessary to practice speaking, they also worried about making mistakes. The author made several pedagogical suggestions: Teachers should encourage positive beliefs that lead to effective learning strategy use and minimize negative beliefs that inhibit learning. By providing knowledge about the nature and process of language learning, teachers can remove misconceptions from students. Teachers can raise students' awareness by methods such as group discussions and make them reflect on their learning by diary-keeping. Teaching methodology and strategy training programs should take into account students' beliefs.

The significance of this study to my study is that I will also study the relationship between learning beliefs and what the students do for test preparation. I will also give an open question to the students to talk about beliefs that might not have been addressed in the questionnaire.

Cotterall (1999) investigated the learning beliefs of a group of students taking an EAP course during a 12-week period. She reported on learning beliefs about factors which the literature suggests contribute to successful language learning, specifically autonomous language learning, such as the role of feedback, opportunities to practice and knowledge of learning strategies. The last part of the questionnaire contained a single item requiring subjects to write a letter to a friend giving advice on language learning. The results concerning the role of the teacher showed that the students

believed they should share responsibility with the teacher for their learning and that a key attribute of a language teacher was to show them how to learn. This finding showed students' willingness to assume responsibility, a characteristic of successful and autonomous language learning. Students also believed that making mistakes was a natural part of language learning, that different people learned languages in different ways and that language learning took a long time. Another autonomy-favouring belief was students' willingness to employ a range of key language learning strategies such as analyzing needs, setting goals and planning their learning even when they lacked knowledge of these strategies. Analysis of the responses to cognitive, social and metacognitive strategy items showed that the majority of the students reported knowing how to adopt six of the eight strategies. The author believed that these strategies represented important knowledge and behaviour for autonomous learning. However, the two metacognitive strategies which received the least response were monitoring and evaluating learning. This suggested that these two strategies might not have been well understood by the learners. The results showed that subjects' inability to use these strategies was related to a lack of confidence. This was identified in the items which investigated self-efficacy. While the majority of the learners showed confidence in general language learning ability, they showed less confidence in evaluating their work and measuring their progress.

This study is related to my study in that I will ask the students to write a letter to their friends making recommendations about the best ways of preparing for the SPE/GE Test hoping that they will also talk about their beliefs directly or indirectly.

Sakui and Gaies (1999) studied the beliefs of Japanese university learners of English using a questionnaire and interviews. The participants were 1296 students at public and private 2-and 4-year institutions of higher education. Some of the

participants were English majors, but the majority was not. The results showed that the learners responded differently to the 45 statements which reflected a variety of beliefs. They agreed strongly with certain statements, moderately agreed or disagreed with others, and strongly disagreed with yet other statements. Factor analysis yielded four factors: 1) Beliefs about a contemporary (communicative) orientation to learning English, 2) Beliefs about a traditional orientation to learning English, 3) Beliefs about the quality and sufficiency of classroom instruction for learning English, and 4) Beliefs about foreign language aptitude and difficulty. These factors included 25 of the 45 items. Many of them related to one another both statistically and logically. The authors suggested that language learners should be aware of and internalize a coherent set of beliefs about methodological options. They claimed that their data tentatively showed that the learners had such awareness and beliefs.

This study is related to my study in three ways. I will examine what contemporary beliefs and what traditional beliefs the students have, I will use both a questionnaire and interviews, and the participants of my study will be English major and non- English major applicants.

Using interviews and think-aloud methods, Victori (1999) investigated the relationship between beliefs and writing skills. The author asked two good writers and two bad writers to think aloud while writing an argumentative essay and then interviewed them to gain further insights into their beliefs. The data were analysed based on Flavell's (1979) general taxonomy of 'person knowledge', 'task knowledge' and 'strategy knowledge', which the author had adapted for the writing skill.

The study showed that the two better writers had a broader and complex view of their writing problems (person knowledge), the nature and requirements of the writing task (task knowledge) and their own approach to writing (strategy knowledge).

The protocols suggested that these writers appeared to have a flexible view of the composing task. For example, they believed that the contents of the different parts of the essay could vary depending on the type of task or topic knowledge. In contrast, the two poor writers had a more limited and inappropriate knowledge of the writing tasks which caused them to adopt inefficient strategies. The two kinds of writers were also different in the degree of effort they expended and in their commitment to writing. The poor writers' admitted laziness and lack of commitment caused them not to perform in the way they believed to be the best. They avoided correcting the errors they had identified, consulting a dictionary and planning the organization of the essay. Another finding concerning the poor writers was that in their interview accounts, they reported using strategies that were not observed in their protocols. In sum, the finding of the study substantiated the claim that in order to understand the differences between successful and unsuccessful students, we should take into consideration not only the processes and strategies they undertake but also the knowledge that they bring to the learning task, i.e. their person, task and strategy knowledge.

Victori's study is related to my study in four ways. First, I will examine the relationship between learning beliefs and learning activities. Second, although I will mainly use the categories of the BALLI questionnaire, I will incorporate adaptations. Third, I will compare two groups of learners, and fourth, I will use interviews.

White (1999) conducted a longitudinal study tracking the expectations, shifts in expectation and emergent beliefs of novice self-instructed language learners. A cycle of interviews, ranking exercises, questionnaires, scenarios and yoked subject procedures were used to collect the data. The study focused on the insider's perspective and self-instruction was investigated through the eyes of the learner. The means of data collection were not predetermined but were chosen, developed and

adopted according to the kinds of information that emerged. The results showed changes in the students' beliefs and expectations. In the later stages of learning, the students considered internal factors for success (effort, learner characteristics, knowing how to learn best) more important than external factors (interactions, course materials). The study suggested that not only learner beliefs, but also learner characteristics or predispositions influenced how learners conceptualised and experienced self instructed learning.

The significance of this study to my study is that I will have a flexible view of beliefs. Although I will use a questionnaire which assumes that beliefs are stable, in the end I will consider the context from which the beliefs emerge. The second relevance of White's study is that I will also use various instruments. Third, I will look at beliefs from the eyes of the learners, and fourth, I will study the learners in a self-study situation.

Kalaja (2003) (see also Huhta, Kalaja, and Pitkanen-Huhta, 2006), analysed expectations of success, as an aspect of learning beliefs, held by students who were about to take a high stakes school leaving test of English. To provide the data, the students were asked to keep a diary of their thoughts, feelings and experiences concerning the English test before and after each subtest and after receiving the official test results. This was followed later by a discussion in pairs or groups of three. Attention was paid to how the students talked about 1) themselves as test-takers, 2) the foreign language test, and 3) their performance on the test. The accounts of expectations about taking the test varied from one situation to another. They varied from getting the best mark, to full indifference ('I don't care at all how it goes') and from the third best mark, to a pass with the mark left unspecified ('as long as one passes the test'). However, the author found that there was unity in these accounts.

They seemed to be the ways of talking about taking the test that were culturally available to them. The author suggested that it was important for the teacher and the students to be aware of the ways in which accounts of test performance were given, or more specifically, of expectations and explanations of successes and failures. The author also suggested that students should be provided with opportunities to share their accounts with those of others so that they could become aware of alternative accounts, or repertoires, which in turn might lead them to reconsider their own and the reasons for resorting to them. The study concluded that the perceptions learners had of themselves, the difficulty level of the test and hence of their performance on the test might affect the degree to which they invested time and energy to prepare themselves for the test.

This study is related to my study particularly because it deals with preparation for a high-stakes exam. As in Kalaja's study, I will also examine the relationship between learning beliefs and test preparation activities particularly in terms of the amount of time students spend on the different skills and on the test preparation as a whole.

Kalaja et al (forthcoming) investigated the beliefs of teacher trainees about EFL learning through narratives. They showed how different research tools and tasks elicited different aspects of EFL learning. The authors suggested that this point should be considered in the selection of the methodological tools as learners might express themselves better by certain tools than by others, for example by drawing than by verbal mode. They showed that in the tasks they used to elicit responses from their participants, they could hear the participants' voices clearly, and that it was not a single voice but multi-voices which they heard. They urge teacher trainers to give students a number of opportunities (tasks, modes, etc) to reflect on aspects of their

learning.

The relevance of this study to my study is that I will give the learners different opportunities by using different instruments to elicit their beliefs.

**Table 3-1 Summary of the individual studies of beliefs**

| Authors | Instruments | Main findings |
|---|---|---|
| Cotterall (1999) | Questionnaire, Letters | The results showed that while the majority of the learners had some autonomy-favouring beliefs and showed confidence in general language learning ability, they showed less confidence in evaluating their work and measuring their progress. |
| Davis (2003) | Questionnaire (Lightbown & Spada, 1993) | There were differences between teachers' beliefs and students' beliefs. Teachers' beliefs corresponded largely with accepted language learning theory statements, while the students' thinking corresponded only partly with the theories. |
| Horwitz (1999) | Questionnaire (BALLI) | There was as much variation in beliefs within the cultures as there was between the cultures. The results suggested that language learning context should be considered as one of the sources of variation. There was also a large amount of similarity across the beliefs, which was attributed to the possibility of the existence of a world culture of language learning. |
| Kalaja (2003) | Students' diaries followed by discussions | Although the accounts of expectations about taking the test varied from one situation to another, there was unity in them. The study concluded that the perceptions learners had of themselves, the difficulty level of the test and hence of their performance on the test might affect the degree to which they invested time and energy to prepare themselves for the test. |
| Kalaja et al (forthcoming) | Narratives | The results showed that different research tools and tasks elicited different aspects of EFL learning and learners expressed themselves better by certain tools than by others. The authors urged teacher trainers to give students a number of opportunities (tasks, modes, etc) to reflect on aspects of their learning. |
| Sakui & Gaies (1999) | Questionnaire, interviews | Four kinds of beliefs were identified: 1) Beliefs about a communicative orientation to learning, 2) Beliefs about a traditional orientation to learning, 3) Beliefs about the quality and sufficiency of classroom instruction, and 4) Beliefs about foreign language aptitude and difficulty. The authors suggested that language learners should be aware of and internalise a coherent set of beliefs. |
| Victori (1999) | Interviews, think-aloud | The study showed that good writers had broader and more complex beliefs about writing and a flexible view of the composing task. The findings confirmed that in order to distinguish successful and unsuccessful students, we should take into consideration not only the processes and strategies but also the beliefs that learners bring to the learning task. |
| White (1999) | Interviews, ranking exercises, questionnaire, scenarios, yoked subject procedures | The results showed that the students' beliefs and expectations changed. The study suggested that not only learner beliefs, but also learner characteristics or predispositions influenced how learners conceptualised and experienced self instructed learning. |
| Yang (1999) | Questionnaires (BALLI, SILL) open-ended questions | The students had some conflicting beliefs which were reflected in their use of strategies. The author suggested that by providing knowledge about the nature and process of language learning, teachers can remove misconceptions from students. |

## 3.6. Summary of the beliefs studies combined

- Based on my literature review in 2.10, it is possible to group the studies described above within the three approaches. I categorise Horwitz (1999) and Davis (2003) as normative, Yang (1999), Cotterall (1999), Sakui and Gaies (1999) and Victori (1999) as metacognitive, and White (1999), Kalaja (2003), Kalaja et al (forthcoming), and Huhta, Kalaja, and Pitkanen-Huhta (2006) as contextual.

- Except for Kalaja (2003) (and Huhta, Kalaja, and Pitkanen-Huhta, 2006), none of the studies addressed test preparation situation.

- Each study addressed the types of beliefs the researcher was interested in.

- Depending on the approach of the studies, a wide variety of instruments was used.

- Some studies looked at the relationship between beliefs and actions as a cause-effect relationship (with beliefs causing behavior) while others considered a dynamic relationship between them looking at beliefs as situationally-conditioned and unstable or flexible.

## 3.7. How my literature review informs my study

Part 1 of my literature review established that more research was needed in the area of washback on learning and that hardly any washback studies addressed learning beliefs. Part 2 of my literature review showed that although Kalaja (2003) (and Huhta et al, 2006) addressed beliefs, they dealt with only one aspect of beliefs i.e. expectations. This study, therefore, will investigate the effect of the SPE Test on the students' learning activities and the relationship between the test and learning beliefs. The reason why I would like to investigate beliefs as a prerequisite for autonomy is the unique test preparation situation in which SPE students have to fend for themselves as their materials are not taught at school and they have to find themselves

a private preparation class, a private teacher, appropriate materials, etc (see also Chapter 1).

In this study, I will use the three instruments: questionnaire, letters, and interviews, each of which will address both washback and beliefs. I will explain in my Methodology Chapter that due to some limitations I will not be able to use classroom observation or conduct a longitudinal study.

As far as beliefs are concerned, I will not follow one approach but a combination of them. I will use one instrument from each approach. I will use a questionnaire from the normative approach, interviews from the metacognitive approach, and letters from the contextual approach. I will use the popular BALLI questionnaire and a few items from other questionnaires. The reason I will use BALII is because it is most frequently used and it deals with learning methods which are dealt with in the SPE students' textbooks as learning tips. Although my approach was combinatory for the research instruments, my view of beliefs will be contextual as I am interested in how the SPE Test will interact with beliefs. My position of the relationship between learning beliefs and the SPE Test looks as follows:

**Table 3-2 View of beliefs: Possible interaction between the SPE Test washback and learner beliefs**

| Activities likely to be influenced by the SPE Test | Interaction | Learner beliefs |
|---|---|---|
| What learners learn (contents)<br>How learners learn (method)<br>Rate and sequence of learning<br>Degree and depth of learning,<br>Etc. | →<br>← | Person Knowledge<br>↓ ↑ ↑<br>↓Task Knowledge↑<br>↓ ↑ ↑<br>Strategic knowledge |

In Table 3.2, the left hand column shows the areas affected by the SPE Test and the right column shows learning beliefs. I depict an interactional relationship rather than a linear one between beliefs and the washback areas by using two arrows pointing in two opposite directions. The rationale behind this interactional view is

expressed by Barcelos (2003: 19). She states that not only may beliefs drive actions, but also actions and reflections on experiences (i.e. the washback areas where the learners do their preparation activities) may lead to changes in beliefs or create other beliefs. I also established an interactional relationship between the kinds of beliefs as well. In this connection, Barcelos (2003) states that 'belief systems are not linear or structured, but complex and embedded within sets of beliefs forming a multilayered web of relationship' (p. 26).

Finally, my research questions are as follows:

1. What activities do the SPE students report doing in order to prepare for the SPE Test?

2. Why do the students report doing these activities in order to prepare for the SPE Test?

3. Do SPE students perceive the SPE Test to be more difficult than the GE Test?

4. Do SPE students have better English backgrounds than GE students?

5. What beliefs do the SPE students report holding about learning English?

6. Are the SPE students' reported activities consistent with their reported beliefs?

# Chapter 4. Pilot Study: Methodology, Results, and the Lessons Learned

In chapters 2 and 3, I reviewed the literature on washback and learner beliefs and discussed how they helped me arrive at my research questions. I also mentioned that in addition to the literature my knowledge of the SPE Exam context contributed to the formulation of the research questions (3.7).

In the literature review I also discussed the methodologies that have been used in washback studies so far. In this chapter, I will operationalise my research questions and discuss how I selected my methodologies from the available methodologies and what approach they were based on. More specifically, I will discuss issues of the selection of the instruments, validity, reliability, sampling, participants, data handling, and ethical considerations.

## 4.1. Design of the study in general

### 4.1.1. Approach

I adopted a combination of quantitative and qualitative approaches in my study. To determine the approach, first I asked myself what the purpose of the study was, or in other words, what it was that I was going to investigate (Cohen et al, 2000: 73). The purpose of my study was to investigate the effect of the SPE Test on the learning of the students. As I was familiar with the research context, including familiarity with the changes introduced into the SPE Test, to some extent I was able to predict the nature of the test washback. In other words, I started my research with some hypotheses which I was going to test. In addition, I was going to test these hypotheses across a large population. These characteristics meant that I was going to apply a positivist and quantitative approach. According to Ary et al (2006: 27), studying relationships, testing hypotheses, and using large samples are major

characteristics of quantitative research approaches.

However, I was also interested in the 'why and 'how' of the influence of the SPE Test from the learners' perspective, i.e. to gain a more comprehensive picture, through qualitative methods. Examining a phenomenon in rich detail is a major characteristic of qualitative research (ibid). Consequently, the approach I chose to use in my study consisted of both quantitative and qualitative methods at least as far as the study data was concerned. Best and Kahn (2006) not only do not see quantitative and qualitative approaches as mutually exclusive but also believe that in order to answer all of the questions, a research study may need to combine both approaches. They give survey research (as is the case in this study) as an example of a research method which often uses both methodologies: 'A single survey will often contain questions that provide for quantitative responses and also ask questions that result in qualitative data' (p. 271). The approach I took affected the choice of my instruments which I discuss in the following section.

### 4.1.2. Selection of the instruments

I chose to use three instruments in my study: questionnaires, letters, and interviews, the second of which had never been used in washback studies so far (see Table 2.1). Not only did the study approach influence the choice of my instruments but also I weighed the merits and demerits of the instruments that have been used in washback studies and came up with my own selection of suitable instruments as follows.

The questionnaire had certain advantages which made it possible for me to address large groups of respondents and test the hypotheses that I had (Cohen, 1998). It also allowed for ease of statistical data analysis (ibid). However it had a disadvantage as well, which was superficiality of responses (ibid). One solution was

to ask students to write letters to their friends and advise them on how to prepare for the test (Table 2.2) so that students would have a chance to write about issues not covered in the questionnaire. Another solution was to use semi-structured interviews in order to take further the questionnaire responses and the comments in the letters and use the interview as a new source of data as well. Both solutions made it possible to pursue topics of interest and identify new dimensions (ibid). However, the use of interview and letters would have resulted in a large amount of data for which I thought of using qualitative data analysis software such as ATLAS.ti (Muhr, 2003-2004). I considered one further disadvantage which the three instruments had in common, i.e. all would produce self-report data. As part of my research concerned learning activities/behaviour, I had to think of some other instrument i.e. observation. However, the problem I had with observation was that teachers were reluctant to allow me to observe their classes. Furthermore, what I was interested in was not only classroom activities but also out-of-class self-study activities. Clearly it was not practical to follow students around out of class to observe what they did. Therefore, I reluctantly decided not to conduct observation.

Having decided which instruments I would use, I took the next step which was operationalisation of the instruments in order to address my research questions. However, since I was able to use only the questionnaire in my pilot study, I only discuss operationalisation of the questionnaire in the following section, and will discuss the other instruments in Chapter 5.

### 4.1.3. Operationalisation of the questionnaire

According to my research questions, two main themes were to be included in the instruments: reported learning activities, and reported learning beliefs. I also needed to know whether the students were familiar with the test so that I would be

able to relate their activities to the test (Bailey, 1996: 276; Wall 2005: 52). I also needed some information about the students' background. The following table shows the categories/themes used in the questionnaire and the research questions they address.

**Table 4-1 Categories and the research questions addressed in the questionnaire**

| Background information | Difficulty | Test preparation activities | Test knowledge, Purposes, Motivation, Interest, Anxiety | Language learning beliefs |
|---|---|---|---|---|
| Gender, school, field, LP, SPE/GE grouping | | | | |
| RQ 4 | RQ 3 | RQ 1 | RQ 2 | RQ 5 & 6 |

As Table 4.1 shows, the items on reported learning activities were supposed to answer Research Question 1, items on reported beliefs Research Question 4, and items on background information were to answer Research Question 3.

There was not much problem in determining the content of the 'background information' as it would ask about some basic information such as gender as well as about variables I was interested in such as English background, field of study, and type of school. To determine the content of 'beliefs', I used BALLI (Horwitz, 1987, 1999). However, I did not use the questionnaire in its original form but made some adaptations (see the items in Appendix 4). Items 13 and 24 concerned 'Americans' as native speakers, which I changed to 'English speakers'. Similarly, I changed 'American friends' in Item 32 to 'English-speaking friends'. Item 26 concerned practising 'with cassettes or tapes', to which I also added 'CDs'. Item 28 dealt with translating from mother tongue to English. Since it is much more common among the students in my context to translate from English to mother tongue, I modified the direction of translation accordingly.

For 'knowledge about the test' I adapted Wall and Horak's (2006) Students' Post-Observation Questionnaire. I drew on Herington (1996) for items relating to

'perceptions and attitudes about the test', and on Cohen and Chi (2002) for items relating to 'reported learning activities'. However, in order to ensure that the questionnaire items also reflected the students' likely actual learning behaviours, I also ran focus groups with former SPE and GE students who were currently in university. In this regard, authors suggest that focus groups are useful for identifying questions and developing themes and topics for questionnaires (Morgan, 1997: 2; Krueger and Casey, 2000: 19; Ary et al, 2006: 481).

Morgan (1997: 25) states that there are three ways in which focus groups can contribute to the development of questionnaire items: 1) focus groups can help capture the domains that need to be measured in the survey, 2) they can help determine the dimensions that make up each of these domains, 3) they can provide item wordings that could be interpreted in the same way by the researcher and the respondents. The fact that the items cover the contents of the domains and their dimensions and the fact that they are likely to mean the same for both the respondents and the researcher would result in reducing invalidity (ibid: 26). Item wordings would also reduce unreliability by 'minimising differences in how the respondents interpret the questions' (ibid).

In order to design focus groups, I referred to the literature and found out about who could participate in the groups, how structured the groups would be (including the level of moderator involvement), the size of the groups, and the number of the groups (Morgan, 1997: 34).

As a rule of thumb, focus groups most often use homogeneous strangers as participants, a relatively structured interview with high moderator involvement, 6 to 10 participants, and 3 to 5 groups. However, these rules of thumb are not a standard. In reality, a project rarely matches all four of these criteria. For example

acquaintanceship may be unavoidable (ibid).

Based on the literature reviewed above, I selected the participants and decided on the structure, size, and number of the groups. I asked for volunteer students, as my assumption was that they would be more motivated to discuss the topics. They were homogeneous in that they were either English major students or non-English major students at university who had taken the SPE and the GE tests respectively the year before. I assigned the students to separate groups based on whether they were English-major students or non-English-major students. It was not possible to recruit 'strangers' as this required inviting students from different universities and agreeing on a venue for conducting the focus groups. This was not practical given the time constraints I had. The structure I adopted was neither fully structured nor fully unstructured, but semi-structured. As Morgan states, this kind of structure is used when there are a pre-existing agenda and research questions, and when focus groups are part of a larger project. For the same reason, I set the number of participants at six for each group and the number of groups at two for each group of English majors and non-English majors (i.e. 4 groups in total). I set the size of each group at the minimum of six because small groups work best when the participants have a high level of involvement with the topic. However, that the students would have a higher level of involvement with the topic was based on an assumption, the students being volunteers. I also decided to have a one-week interval between the first round of focus groups and the second round so that I could reflect on the first round and solve possible problems for the second round.

After planning the design of the groups, I decided to determine the procedures for conducting the interviews including the contents of the interview guide.

Based on these guidelines, I prepared an interview guide. The guide was

accompanied by stimulus materials, which were copies of the SPE and the GE tests.

After I conducted the focus groups, I used the results as input to my questionnaire. For example, if students in the focus group stated that the book authored by NOET was not enough for the SPE preparation, I formed this statement into an item on the scale of agreement such as 'The book authored by NOET was enough for the SPE preparation'.

In addition to the inputs from the focus groups and existing questionnaires, I also included open-ended items in the questionnaire. I included these items so that students would have a chance to express ideas which might not have been covered by the questionnaire.

Having incorporated the items in the questionnaire, I ended up with a questionnaire with 204 items. I included this number of items hoping to pilot various questions in order to determine the best ones for the main study. However, I split the questionnaire into three hoping that this would allow me to give a shorter questionnaire each time so that the students wouldn't feel tired answering them. I included background information, reported learning beliefs, and knowledge about the test in Questionnaire I, perceptions and attitudes about the test in Questionnaire II, and reported activities in Questionnaire III.

### 4.1.4. Identifying the participants

Washback studies typically require researchers to contrast the teaching and learning situation before the introduction of the test with the teaching and learning situation after the introduction of the test (Wall and Alderson, 1993; Wall, 2000; 2007; Bailey, 1996). However, since I did not have access to baseline data (the teaching and learning situation before the test was introduced) I had to adopt a contrasting groups design i.e. students who will take the SPE and students who will

take the GE (Alderson and Hamp-Lyons, 1996). In the following section, I will explain how I selected my participants.

Since washback generally refers to the effect of an exam on the teaching and learning which precede it (Alderson and Hamp-Lyons, 1996: 280), the participants had to be the students who were preparing for the SPE Test as a component of the University Entrance Examination. These students were to be Pre-University students doing their last year at school. They were preparing for the SPE test in order to get admitted to university to do a four-year programme in English literature, translation, or Teaching English as a Foreign Language (TEFL). The SPE Group's examination consisted of Specialised English (which I call the SPE component) as well as General English (which I call the GE component). I will refer to the whole examination of this group as the SPE Exam (though it is a combination of the SPE and the GE components). So that I would be able to separate the washback effect of the SPE Component, I had to find another group who was not going to take this component. As mentioned at the beginning of this section, this type of data is usually provided in the baseline study. However, due to the lack of a baseline study, I selected a group of students to compensate for the lack of this data. This contrasting group was going to take the exam in which there was only one component of English which was General English and a Specialised Section which included subjects of interest in Farsi (law, education, medicine, engineering, or whatever the interest of the students might be). I will refer to this whole examination as the GE Exam. The following table summarises the information above, indicating the type and number of English components in the University Entrance Examinations of the SPE and GE groups.

**Table 4-2 English components in University Entrance Examinations taken by SPE and GE groups**

| Groups | University Exam | | No of English components |
|---|---|---|---|
| | Specialised Section | General Section | |
| SPE Group | Specialised English | General English (+ 3 non-English subtests) | 2 |
| GE Group | Specialised subjects of interest in Farsi | General English (+ 3 non-English subtests) | 1 |

### 4.1.5. Sampling of the participants

In this section, first I will briefly review the literature on the theory of sampling, and then I will discuss how I applied the theory. Since this study is both qualitative and quantitative, I will discuss sampling and its application in connection with both types of research.

According to Ary et al (2006), usually it is neither possible nor necessary to study all possible cases to understand the phenomenon under investigation. Sampling is indispensable in that it enables the researcher to observe only a small number of all possible cases. The small group is called a sample and the larger group is called a population. In quantitative research, the researcher draws conclusions from these observations to generalise to the larger population using inductive reasoning. If it were possible to observe all instances of a population, one could base their conclusions about the population on these observations. However, observing a small sample of population, the researcher can only infer that these observations will be true of the population. Making inferences about the population with reasonable confidence requires that the sample be representative of the population. If individuals in the sample are different in their characteristics from the target population, the sample would not be representative. An unrepresentative sample is termed a biased sample.

The first step in sampling is the identification of the target population. For example if the researcher is interested in learning about teachers in public schools in a

district, all teachers and the public schools of that district would be the target population. There are two major types of sampling procedures, probability sampling and non-probability sampling. In probability sampling, individuals in the sample are selected by chance, and each individual has an equal chance to be chosen in the sample. Non-probability sampling, however, is not based on chance but the knowledge and expertise and judgment of the researcher. It is used when probability sampling is not practical. It is driven by convenience and economy (ibid).

Sampling is important in qualitative research as well; however, the size of the sample is much smaller. This is because what the qualitative researcher looks for is depth and extent of information (ibid). Though random sampling may be ideal in quantitative research, it is purposive in qualitative research (Silverman, 2001: 250). Based on their knowledge and experience, researchers select a sample that has features they are interested in and that can provide the information they want. Guba and Lincoln (1981: 276) wrote, 'Sampling is almost never representative or random but purposive, intended to exploit competing views and fresh perspectives as fully as possible'. The size of a sample is determined by whether any new information can be provided by a new member of the sample. Of course it is also determined by practical considerations such as time, money, and access (ibid).

Based on the literature reviewed above, first I identified the target population and their characteristics so that I could base my sampling on them. The population consisted of Pre-University students in the whole country, males and females, in three types of regions with different levels of educational facilities, in five types of schools with different educational qualities, and in the three fields of study of mathematics, natural sciences, and humanities.

Having identified the target population and the various strata they came from, I

had to determine whether to assign equal or unequal proportions to each stratum, and if unequal, what proportions. This data was fortunately available in the National Organisation for Educational Testing (NOET). The data was University Entrance Examination statistics which is published every year by NOET. The statistics I had access to concerned the year 2004 which gave information about the number of students who *participated* as foreign language applicants and the number of the students who were finally *admitted* to university. As both kinds of information were useful, I took their average. The averages indicated how many students to select for each group and subgroups. However, the information the statistics did not provide was how many of the applicants or the admitted students were from which 'type of school', so I decided to collect the data based on the number and the population of each type of school and on the extent to which I would manage to arrange with the school authorities to administer the questionnaires.

Table 3.3 shows to what extent the study sample was going to approximate the NOET statistics. The statistics showed how many students each group and the subgroups would consist of, for example how many *female* students from the field of *mathematics* in *Region* 1.

**Table 4-3 Sample of questionnaire respondents based on NOET statistics on 2004 University Entrance Exam showing the number of students in main groups and subgroups**

| **Gender** | | **Female** | | **Male** | | **Total** |
|---|---|---|---|---|---|---|
| **Groups** | | **SPE** | **GE** | **SPE** | **GE** | |
| **Region 1 (390)** | **Maths** | 71 | 71 | 22 | 22 | 186 |
| | **NS** | 62 | 62 | 20 | 20 | 164 |
| | **Hum** | 15 | 15 | 5 | 5 | 40 |
| **Region 2 (420)** | **Maths** | 77 | 77 | 24 | 24 | 202 |
| | **NS** | 67 | 67 | 21 | 21 | 176 |
| | **Hum** | 16 | 16 | 5 | 5 | 42 |
| **Region 3 (190)** | **Maths** | 35 | 35 | 11 | 11 | 92 |
| | **NS** | 30 | 30 | 10 | 10 | 80 |
| | **Hum** | 7 | 7 | 2 | 2 | 18 |
| **Total** | | **380** | **380** | **120** | **120** | **1000** |
| **Total gender** | | **760** | | **240** | | |
| **Total Maths= 480, Total NS= 420, Total Hum= 100** | | | | | | |
| **Total SPE= 500, Total GE= 500** | | | | | | |

**Note:** NS= Natural Sciences, Hum= Humanities

Due to access problem, however, I was able to collect data from students with the following proportions.

**Table 4-4 Percentage of respondents from each category who actually answered the pilot questionnaires**

| Categories | | SPE (%) | GE (%) | Total (%) |
|---|---|---|---|---|
| **Gender** | Male | 9.9 | 51.7 | 41.6 |
| | Female | 90.1 | 48.3 | 58.4 |
| **Region** | 1 | 39.6 | 12.4 | 19.0 |
| | 2 | 53.2 | 81.6 | 74.7 |
| | 3 | 7.2 | 6.0 | 6.3 |
| **Types of schools** | Smart | .9 | 19.0 | 14.6 |
| | Nemooneh | | 1.7 | 1.3 |
| | Shahed | | 3.2 | 2.4 |
| | Non-profit | 12.6 | 16.4 | 15.5 |
| | Public | 86.5 | 59.8 | 66.2 |
| **Fields of study** | Mathematics | 30.6 | 43.1 | 40.1 |
| | Natural Sciences | 36.0 | 26.4 | 28.8 |
| | Humanities | 33.3 | 29.9 | 30.7 |
| **Attendance in English institutes by terms** | Mean | 10.76 | 6.11 | 7.22 |
| | SD | 9.41 | 8.13 | 8.67 |
| **Total respondents in each group** | | 111 | 348 | 459 |

### 4.1.6. Validation of the instruments

According to Alderson and Banerjee (2001), there has been a gap in the literature on research instrument validation. The literature is mainly concerned with construction of the instruments rather than a set of procedures for establishing their validity and reliability. In other words, it is generally believed that if an instrument is carefully constructed, it is a good instrument. The authors address this gap in the literature and offer a set of procedures for the validation of instruments used in washback studies, i.e. a set of procedures to enable the researchers to check whether their instruments have captured the information required to answer their research questions and whether they have captured the information reliably (p. 152). In this section I explain some of the authors' suggestions which were applicable to my study and show how I applied them.

The authors believe that one of the aims of validation is to ensure whether the questions in the instrument mean the same to the respondents as they do to the researcher. They refer to this as piloting in which the researcher finds out where respondents may have problems answering the questions and why.

First I piloted the questionnaires with two SPE and two GE students. I asked them to answer the questions and tell me what they thought each question was asking and where there were ambiguities (Alderson, 1992). Then, I piloted the questionnaires with a whole class of SPE and a whole class of GE students who were doing their first semester at university and had taken SPE and GE tests (respectively) about 4 months ago. The pilot was carried out in my presence to enable the students to ask any questions they might have (Cheng, 1998). The students pointed out several cases of ambiguities which I revised.

They recommended giving the instruments to a few of the people for whom the instruments have been designed. For example, a discussion can be held with groups of students in order to explore similarities and differences in responses and indicate possible problems with the items. Similarly, the writers suggested making a profile of the expected answers and then comparing them with the actual responses. I believe what enabled me to put this suggestion into practice was the focus groups (Section 4.1.3). Not only did the focus groups help me to provide items for the questionnaire, but also they helped me to come up with possible answers.

The authors suggest expert judgment and correction of item wordings for the purpose of content validity. To put this suggestion into practice, in one of my presentations in the Language Testing Research Group (LTRG) in Lancaster University, I asked some staff members, research students and MA students of language testing to comment on the appropriacy of my questionnaire items. The group

criticised two of the items which I deleted from the questionnaire.

They also suggested exploring the extent to which individual items discriminated between respondents appropriately based on the respondents' characteristics (e.g. those who were going to take the test versus those who were not) and their response patterns (e.g. favourable versus unfavourable attitudes toward the test).

I also explored whether the items discriminated between the respondents by running significance tests- not necessarily for the purpose of validation but because of what this whole study was all about (separating the washback effect of the SPE Test from the GE Test). Tests of significance showed that SPE and GE groups were significantly different on most of the items (Appendix 4). Therefore, the statistical analyses demonstrated the validity of the questionnaire responses which also indicated the input from the focus groups was appropriate. The authors also state that differences between the groups would be an indication of construct validity (i.e. divergent validity) as well.

As another kind of evidence for questionnaire validity, the authors suggest the Test-Retest method of estimating reliability, i.e. checking whether respondents give the same responses on different administrations (see also Alderson, 1992; Alderson and Banerjee, 1996). However, they warn that as this procedure assumes that constructs are stable, it may be problematic for measuring attitudes which are subject to change over time. Therefore, they suggest that too much time should not intervene between the administrations.

I checked the reliability of my questionnaires by administering them to the same groups of SPE and GE students twice with one week between the two administrations. The results will be explained in 4.6.2.1 (see Appendix 5).

The authors also suggest grouping of the questions based on their contents and categories. Then the responses can be cross-checked within groups for consistency and across groups for expected differences. They believe that the results could also indicate content validity. I will discuss the results of the internal consistency estimate in 4.6.2.2 (see also Appendix 6).

## 4.2. Data collection

### 4.2.1. Administration of the questionnaires

On average, it took the students about 30-45 minutes to complete the questionnaires. I observed that the students took the questionnaires seriously. Many students came to me and volunteered to answer the questionnaires. This might be because the questionnaire was in direct relevance to their high stakes exam or it might also be because sometimes school principals or their assistants talked to the students about the importance and benefits of research in general before the administrations. The significant difference between the SPE and the GE groups on most of the items further confirms this conclusion (Appendix 4).

### 4.2.2. Ethical considerations

I got consent for data collection in several ways. First of all, I got a permission letter from the head education office as well as letters from the education office of each city in which the schools were located. Then, I presented the letters to the school principals or their assistants. I also got consent from the individual participants on the questionnaires. In order to give the students something in return for their time and cooperation, I offered to share the results of the study with them. I also offered that I would be willing to give them advice on exam preparation or on language learning in general and, therefore, I told them that they could keep in touch with me (Ary et al, 2006: 485).

Throughout the data collection process, I tried to respect the rules of the schools. For example, some classes had exams in which case I arranged for another day for those classes, or in one case a math teacher did not agree to give his class time to me because he said he would fall behind schedule as the final exams were drawing near.

## 4.3. Preparing the data for analysis

To prepare the data for analysis, I entered the data in SPSS (version 11) and cleaned the data. I explain each stage below.

### 4.3.1. Data entry

Entering the data involved naming the variables and determining what codes I would assign to the responses (Barker, et al, 2002: 226; Argyrous, 2005: 33). The codes I used were in the form of numbers. For example I used 1 and 2 for Male/Female or Yes/No responses, 1, 2, and 3 for the three fields of study, 1- 5 for five-point Likert scale of agreement and frequency. I also used numbers for responses to ranking items, with 1 indicating highest rank (see Appendix 7 for the full code book).

### 4.3.2. Cleaning the questionnaire data

I checked the data by running 'frequencies'. Data which have been entered by mistake tend to appear in the form of extra choices in the output. For example, if expected responses for an item are 1, 2, and 3, but if, say, 5 or 11 also appear in the output, they are likely to have been entered by mistake. In the debugging process, I eliminated these extra choices by deleting them and then replaced them with the students' original responses by retrieving their questionnaires.

## 4.4. Data analyses

I used Statistical Package for Social Sciences (SPSS), version 11, to analyse

the questionnaire data. The purpose of the analysis was to find out possible differences between the SPE and GE groups so that I would be able to separate out the washback effect of the SPE Test. To this end, I used the statistical techniques of 'frequencies' (Horwitz, 1999) and 'mean' as 'descriptive statistics' and 'Independent Samples T-Test' and 'Chi-Square' as 'inferential statistics'. As 'frequencies' and 'mean' did not indicate whether the differences between the groups were statistically significant, I used 'inferential statistics' to supplement the analysis.

I used 'inferential statistics' for all items but used 'descriptive statistics' differently for different types of items. For 5 and 6-point scale items, I reported both frequencies and means. The mean values were helpful in that they served as a summary of the information frequencies provided. However, for ranking items, I did not report frequencies as they turned out to be very confusing. For example, in a question in which students were asked to rank items from 1-6, the SPSS output gave 6 percentage figures for each of the items that was ranked by the students. This made it difficult to compare the ranked items. Therefore, instead of reporting frequencies, I reported the average of the frequencies (mean), which gave the same amount of information as frequencies did.

For those multiple-choice items which collected categorical data, I reported frequencies but not the mean, as it did not provide useful information. For example, when students were asked to choose, from types of textbooks, the one on which they spent most time, it would not make sense to report the average factor, which in fact did not exist. However, for those multiple-choice items which were based on a continuous scale, I reported both frequencies and the mean. For example, the item which asked the students how many terms they had attended English language institutes and items which asked factual questions about the SPE/GE tests were based

on a continuous scale.

As far as 5 and 6-point scale items are concerned, I reduced them to a 3-point scale. I combined the figures for 'agree and 'strongly agree' and presented them as 'agree', combined 'disagree and 'strongly disagree' and presented them as 'disagree', and presented 'no opinion' separately. Similarly, for items based on a frequency scale, I combined 'It's a good idea but I don't do it' with 'never', 'rarely' with 'sometimes', and 'often' with 'always'.

## 4.5. Lessons learned from the pilot study

The purpose of the following sections is to show whether my methodological decisions worked well within the pilot study or not. In particular, they will show whether the groups of participants I selected were appropriate, whether the research instrument worked well in terms of reliability, coverage, and focus, and whether there were gaps in the study which required different or additional instruments. Where necessary, I will present the results of the pilot study to achieve these aims.

### 4.5.1. Reconsidering sampling

I decided to remove 3 out of 5 types of schools from sampling for the main study. The reason for this decision was based on the fact that the actual sample of the pilot study showed that I would not be able to have a balanced number of SPE and GE students from these schools. The sample showed that there was only one SPE student in the Smart School and no SPE students in the Nemooneh and Shahed schools. In addition, as I enquired from the SPE student in the Smart School, his purpose of taking the SPE test was first to assess his English ability and secondarily to study English only in case he was not admitted to a major of his interest at university. The other reason for removing the three schools was that the number of these schools was very few. In some cities there was only one such school and in some cities there were

no such schools at all. Therefore, I decided to remove these three school types from sampling and only consider Non-Profit and Public schools in the main study.

### 4.5.2. Consistency of the questionnaire

#### 4.5.2.1. Consistency of the questionnaire over time

As I discussed in 4.1.6, one of the methods which I used to estimate the reliability of the questionnaire was the test-retest method to examine if the questionnaire produced consistent results in two different administrations. I did the test-retest reliability for the review of the coverage and the focus of the items and whether the items were being interpreted correctly or in the same way as they were intended. The data is a slightly different data set from university students who had already taken the SPE and the GE and were now engaged in their university study. The two samples consisted of 39 SPE students and 35 GE students. As my data were ordinal, I used the non-parametric Spearman Rho. The reliability of the instrument for the SPE Group was 0.62 and the reliability of the GE Group was 0.77 (the full results are presented in Appendix 5). As the figures show, the reliability of the SPE Group was lower than that of the GE. Based on Barker et al (2002: 70), the reliability estimate of the GE Group was 'good', and the reliability of the SPE Group was 'acceptable'. The lower reliability of the SPE Group might suggest that the attitudes and beliefs of the SPE Group underwent more changes between the two administrations than those of the GE Group. The reason might be because the SPE were having specialized English courses for about 15 hours a week, but the GE students were having their General English Course for 3 hours a week. The English lessons the SPE students were getting were actually important, while the General English Course was not very relevant to them because they were doing university subjects that were not related. Therefore, there were some changes in the students and

in the light of that, the questionnaire actually showed stability. This also suggests that the time interval between the two administrations, particularly for the SPE students, should probably have been shorter than one week.

#### 4.5.2.2. Internal consistency of the questionnaire

In section 4.1.3 which concerned the design of the questionnaire, I discussed the inclusion of certain items (see Appendix 6) in order to check the internal consistency of the instrument. Although I expected these items to receive similar responses, i.e. to be highly correlated, I was not sure if they would really correlate. However, I needed to run a correlational analysis (Spearman Rho) to check both the internal consistency of the questionnaire and my assumptions about whether the items that I grouped together produced the same patterns of results. The results (Appendix 6) showed that, unlike my expectations, the great majority of item pairs i.e. 460 in the SPE Group and 463 in the GE Group had low correlations (0 - 0.50), 12 pairs in the SPE Group and 9 pairs in the GE Group had 'marginal' correlations (0.51 – 0.60) and only 2 items in each group had 'acceptable' correlations (0.61 – 0.70). While authors suggest that we aim for correlations above 0.70 (Dornyei, 2003; Pallant, 2007; Field, 2005), there were no items having internal consistency of this size. The low correlations imply that the items had little in common and were measuring different constructs. However, although the items did not overlap, I had to shorten the questionnaire for practicality reasons, which I describe in the following section.

### 4.5.3. Length of the questionnaire

Dornyei (2003: 18) warns that we should resist the temptation to include everything in a questionnaire that seems interesting to us. He also warns that a long questionnaire can become unproductive. He suggests that the questionnaire should not be more than 4 pages or take more than 30 minutes unless the topic is important to

the respondents. Similarly, Gillham (2000:39) sees the optimal length of a questionnaire from 4 to 6 pages maximum.

The questionnaire I used in my study was very long and took the students a long time to complete and I had to administer it in three batches, namely Questionnaire 1, 2, and 3. However, my pilot study experience showed that the topic was important to the SPE students but unfortunately not so important to the GE students. Therefore, I could not have a long questionnaire just based on the interest of the SPE students. Finally, having both Dornyei and Gillham's suggestions in mind, I decided to shorten the questionnaire aiming for 5 pages.

#### 4.5.3.1. Shortening the questionnaire

I used some criteria to shorten the questionnaire for the main study including removing non-discriminating items and narrowing the focus of the study. I explain these in more detail below.

##### 4.5.3.1.1. Removing non-discriminating items

I eliminated questions that did not discriminate between the groups. As I mentioned in Section 4.1.4, in the absence of a baseline study, I used a contrasting groups-design in order to compare the SPE students with the GE students. Therefore, if an item showed that there was no difference between the groups, it was not fulfilling its purpose and had to be removed.

The results showed that 78 out of 207 items did not significantly differentiate the groups (Appendix 4). I removed most of these items from the questionnaires but kept a few of them which were related to the main focus of the study and my interest. This is explained further in the following sections.

#### 4.5.3.1.2. Shortening the questionnaire by narrowing the focus of the study

I decided to narrow the focus of the study by concentrating on reported activities based on the changes in the SPE Test (1.3) and based on the learning tips in the students' textbooks, i.e. Pre-University English and *Bridging the Gap* (Appendix 3). This enabled me to eliminate items by removing opinion items, removing similar items, and by using alternative items.

#### 4.5.3.1.3. Removing items of opinions about the SPE/GE tests

I decided to keep items of reported activities which were based on a frequency scale and remove items of various opinions about the test which were based on an agreement scale. The reason for this decision was because, as I will explain in 5.2.2, I was going to use other instruments, particularly interviews, which would give me a chance to ask about why the students were doing the activities they reported on the questionnaire. This meant that the majority of the items in Questionnaire 2 could be removed and the majority of the items in Questionnaire 3 could be kept.

For example, from the following items which addressed the issue of preparation class, I kept Item 9 which reported an activity and removed items 1 and 50 which were concerned with opinions:

9. Are you attending a preparation class?

1. Preparation classes are useful for the SPE Exam.

50. School classes have a more important role in preparing us for the SPE Exam than private preparation classes.

I removed other opinion items such as Item 49 below thinking that I could not investigate everything.

49. The SPE Exam is a fair Exam.

I decided that issues such as test fairness might deserve a study on their own.

However, exceptionally I kept some items which were not items of reported activities. I kept some items on motivation, test difficulty, and test preparation purposes as they were important considerations based on the washback literature (Chapter 2), irrespective of whether they were statistically significant or whether they were items of opinions or reported activities.

I had still other options to remove items from the remaining significant items and items of reported activities. These options included removing similar items and replacing items with alternative items.

#### 4.5.3.1.3.1. Removing similar items

I reduced the number of similar items (see Appendix 6) that addressed the same issue and were in different parts of the questionnaires but were phrased differently. In the meantime, I applied logic to decide which one to remove. For example from the two items below, I kept Item 9:

9. Are you attending a preparation class?

17. Attending preparation class

I kept Item 9 which was a yes-no question and removed Item 17 which was based on frequency because attendance at a preparation class was a matter of either-or not frequency on which Item 17 was based. In the following example I kept Item 83 and removed Item 15.

15. Practicing with exam papers of previous years

83. Answering sample questions

Although both items showed significant differences between the groups, the reason for preferring one over the other was that Item 83 was inclusive of Item 15.

Finally, in the following pair of examples, the reason for keeping Item 47 and removing Item 67 was that Item 47 was less likely to be misinterpreted by the students because of the word 'English'.

47. Reading English story books

67. Reading story books

#### 4.5.3.1.3.2. Replacing items with alternative items

In some cases, I decided that I could reduce the number of items by testing them in a different way. The following set of items is an example:

28. Studying grammar

41. Learning grammar

29. Studying vocabulary

40. Increasing my vocabulary

30. Studying reading

These items asked about the frequency with which students studied the sections of the SPE/GE Test. I removed these items (not only because some of them were similarly phrased but also) because I also asked about them through a ranking item i.e. Item 52 below.

52. Which of the following language skills are more essential to your success than other skills? Please rank 5 of them according to importance to your success (1=most essential, 5=least essential).

I modified Item 52 as well. As I was interested in reported activities, I asked about how much time the students spent on each skill instead of asking about the importance of the skills, which had made it an opinion item.

56. On which skill do you spend most time in the test preparation? Please rank

5 of them from most to least time you spend on them (1= most time, 5= least time).

One final point about Item 52 is that, although it showed insignificant differences between the groups, I kept it because it gave an indication about the constructs of the SPE Test, which is an important consideration in washback studies.

### 4.5.4. Deepening the focus of the study

#### 4.5.4.1. Using interviews

During the analyses of the questionnaire data, I was often faced with questions that called for richer data. Typically, I began my analyses with the students' test preparation activities and then looked to see if I could find any responses that could possibly account for those activities. However, at best what the questionnaire data helped me to do was only to establish some relationship. It was not able to identify the (reported) causes of those activities, i.e. it was not able to indicate whether what the students did was due to the effect of the test or any other factor. I felt a need to listen to students' accounts to see why they did what they did. Therefore, I needed some evidence to establish (reported) cause-effect relationships, which the interview was likely to provide (although what the students say might be different from actual causes).

Another gap in the pilot questionnaire was that it only addressed self-study activities but not classroom activities. To examine whether the effect of the SPE Test on self-study activities was also mediated by the classroom activities, I decided to ask the students in the interviews what their teachers did in the classroom. Because of the concerns about the length of the questionnaire, I did not want to address these activities in the questionnaire but in the interviews.

#### 4.5.4.2. An additional instrument

The interviews were still based on some predetermined categories, i.e. the

questionnaire items. Therefore, I decided to collect some grounded data with a new instrument so that I would not impose specific issues on students to talk about. The new instrument would hopefully also provide lenses different from those of the questionnaire and the interviews. I would also be able to try a new instrument which had never been used before in washback studies. To this end, I decided to ask students to write a letter in which they would give advice to a friend on appropriate ways of preparing for the SPE/GE Test (Cotterall, 1999; see also Wall and Horak, 2006). Cotterall (1999) used a questionnaire, a part of which contained an open-ended item requiring students to write a letter to a friend providing advice on language learning (p 499). However, Dornyei (2003:14) does not recommend open-ended items to be used in questionnaires. He believes that questionnaires inherently involve superficial responses and 'brief engagement with the topic', and therefore open-ended items are unlikely to produce rich data. He also quotes Sudman and Bradburn (1983) who say that if respondents are asked to give responses longer than a sentence, they will often refuse to answer the question or the entire questionnaire, and even if they give long responses, many such answers would be uncodable and inappropriate (p. 15). Therefore, based on these arguments, I decided to use letter writing as a separate instrument in the main study. Then, I decided to provide a list of the things the students recommended in their letters and follow them up in the interviews.

In this chapter I discussed the methodology of the pilot study and what lessons I learned from the pilot study in terms of sampling, appropriacy of the instrument, focus of the study, and manageability of the study. Some of the lessons were related to the narrowing of the focus of the study and some others were related to the deepening of the study. In the next chapter, I will discuss the methodology of the main study.

# Chapter 5. Methodology of the Main Study

In Chapter 4, I focused on different aspects of the pilot study including methodology, results, and lessons learned from the pilot study. I discussed how the pilot study informed the main study particularly in terms of the instruments and sampling. Concerning the instruments, I discussed how I shortened the questionnaire and addressed the gaps in the questionnaire which called for new instruments, i.e. letters and interviews to supplement the questionnaires. As for sampling, I discussed the reasons for removing three types of school from the original sample. In this chapter, therefore, I will focus on those aspects of the methodology which were different from those of the pilot study with respect to the instruments, sampling, data collection, and data analysis.

## 5.1. Aims and contents of the instruments

### 5.1.1. Questionnaire

The purpose of using the questionnaire was to test some hypotheses across a large population and to find general patterns of responses to probe in the interviews (4.1.1 and 4.1.2). In terms of the contents, the difference between the questionnaire in the pilot study (4.1.3) and the questionnaire in the main study was due to two major changes which I made (as a result of the pilot study - 4.5) in the size and in the themes of the questionnaire. As I explained in 4.5.2, I shortened the questionnaire by removing non-significant items and narrowing the focus of the study. As a result, the focus of the main study questionnaire was on two major themes of reported activities which concerned test preparation, and reported learning beliefs which concerned language learning in general. Other aspects of test preparation addressed in the questionnaire included issues of motivation, anxiety, test difficulty, purposes of test preparation, and knowledge of the test. Table 5.1 shows the themes of the

questionnaire and the research questions they address.

**Table 5-1 Themes and the research questions addressed in the questionnaire**

| Background information | Difficulty | Self-study test preparation activities | Test knowledge, Purposes, Motivation, Interest, Anxiety | Language learning beliefs |
|---|---|---|---|---|
| RQ 4 | RQ 3 | RQ1 | RQ2 | RQ 5 & 6 |

**Note:** The whole focus of 'background information' was not RQ4 but to check sampling as well.

### 5.1.2. Letter

I had two purposes in mind for using the letter as an instrument (4.5.3.2). One purpose was to collect grounded data, unlike the questionnaire data and partly unlike the interview data (4.5.3.1 and 5.1.3) which were based on pre-determined categories. The second purpose was to try the letter as a new instrument which had never been used in washback studies, and the third purpose was to use it as a tool which would provide different lenses. I asked the students to write a letter in which to give advice to a friend on appropriate ways of preparing for the SPE/GE Test and give reasons for their advice (Appendix 8). Therefore, the content of the letters was determined by the students' agenda, i.e. what they thought was important for test preparation.

### 5.1.3. Interviews

Kvale (1996: 95) believes that the first stage of an interview investigation is deciding on the 'what' and the 'why' of the investigation. He calls this 'thematising' which includes determining the content and the purpose of the study. Regarding the content of an interview study, he says determining the content involves developing a conceptual and theoretical understanding, description, and clarification of the phenomena or the themes to be investigated, and with respect to the purpose of the study, he says an interview study can be explorative or hypothesis testing. The exploratory interview is open and has little structure, while the hypothesis testing interview is more structured (p. 97). The interview can also be both exploratory and hypothesis testing, in which case it is a semi-structured interview. In a semi-structured

interview, there is a sequence of themes and questions to be covered and at the same time, the interviewer is open to changes in the sequence and forms of the questions so that they can follow up the answers and seek new information about the topic.

In this study, the aims of the interviews were both hypothesis testing and exploratory, which means that they were semi-structured (4.1.2 & 4.5.3.1). This, in turn, determined the content of the interviews. As far as the hypothesis testing aspect of the interview is concerned, its contents were the same a priori themes as those of the questionnaire, which involved asking the students whether they confirmed the responses on the questionnaire or not. This aspect of the interview provided different lenses for the results as well, which was another purpose of the interview. However, the interview addressed one more theme than the questionnaire which was 'classroom activities' (4.5.3.1). Table 5.2 shows the themes in the interview and the research questions they address.

**Table 5-2 Themes and the research questions addressed in the interviews**

| Background information | Difficulty | Test preparation activities | | Test knowledge, Purposes, Motivation, Interest, Anxiety, | Language learning beliefs |
|---|---|---|---|---|---|
| | | Self-study activities | Classroom activities | | |
| RQ 4 | RQ 3 | RQ 1 | | RQ 2 | RQ 5 & 6 |

The exploratory aspect of the interview, however, involved asking the students why they did the activities they reported on the questionnaire. As this part of the interview was open, the contents were determined by the students.

## 5.2. Validation of the instruments

### 5.2.1. The questionnaire

#### 5.2.1.1. Piloting

As I modified the questionnaire after the pilot study, I decided to re-pilot it. However, this time I only piloted the questionnaire with one SPE and one GE student. I asked them to answer the questions and point out where there were ambiguities

(Alderson, 1992). However, they stated that they did not have particular problems understanding the questions.

#### 5.2.1.2. Test- retest reliability

I reexamined the reliability of the questionnaire using the test-retest method (4.1.6) with one week interval between the two administrations. I gave the questionnaire to two samples which consisted of 58 SPE students and 59 GE students. I used Spearman Rho to calculate the reliability. As the greater number of SPE students was attending preparation classes, I expected their attitudes and beliefs to be less stable than those of the GE. However, contrary to my expectation, the reliability for the SPE Group was higher than that for the GE. The reliability of the instrument for the SPE Group was 0.78 and the reliability for the GE Group was 0.70 (the full results are presented in Appendix 9). The lower reliability of the GE Group suggests that their attitudes and beliefs underwent more changes between the two administrations than those of the SPE Group. However, the information the questionnaire provided about the respondents did not account for the possible (reported) causes of these changes because it was the SPE Group, the majority of which attended preparation classes, not the GE Group. The results contradicted those of the pilot study where the correlation was higher for the GE Group than that of the SPE Group. However, as attitudes are subject to change over time, the reliability of the questionnaire can be considered sufficiently high (Dornyei, 2000; Alderson and Banerjee, 2001).

### 5.2.2. Piloting the letter and the interview

I piloted the letter instrument with one SPE and one GE student before I collected my main data. I sent the letter writing instruction (see Appendix 8) via email to these students and asked them to email me their letters at their convenience. They

wrote about various issues. The SPE student recommended certain skills which should be learned and certain learning materials including the pre-university textbook, *Bridging the Gap*, and some extra materials. She also mentioned what factors contributed to success in the exam including interest, motivation, and previous English background. The GE student recommended the pre-university book (like the SPE student), as well as a preparation classes and a lot of practice tests. However, both students wrote only short paragraphs and what they did not write very much about were the reasons for their recommendations (Appendix 8). Therefore, I decided not to ask the students for the reasons.

To ensure reliability of interviews, Silverman (2001: 229) suggests pre-testing the interview schedules, tape recording all the interviews, and carefully transcribing these tapes (5.6.2). I pre-tested/piloted the interview schedules with 3 students to make sure the procedures worked including whether the students understood the questions in the same way as I intended and whether the questions and their sequencing worked. As I had recorded the interviews (5.4.3), I listened to the recording after each interview and thought about how I could have conducted it in a better way and how I could have probed the issues in a more effective way. I used the insights to make improvements on later interviews. However, the main problem I was faced with during piloting was that it was near exam time and the teachers were trying to catch up with the schedules. The students were only given permission to leave the class in the last half hour for the interviews. Therefore, while the initial plan was to examine reported activities, motivation, anxiety, test difficulty, purposes, and learning beliefs, I decided to focus on reported activities and a selected number of reported beliefs.

## 5.3. Sampling

The main study sample was different from the pilot study sample (Table 4.4) in terms of the types of schools which were considered in the sampling. As I explained in 4.5.1, the reason why I removed 3 out of 5 types of schools from sampling was that, the actual sample of the pilot study indicated that I would not be able to have access to a balanced number of SPE and GE students from the Smart, Nemooneh, and Shahed schools. Therefore, I only considered Non-Profit and the Public schools in the main study sampling. Other aspects of the sampling, however, were the same in the two studies. The following tables show the samples for each instrument.

**Table 5-3 Sampling of Pre-university students as questionnaire respondents in the main study**

| Gender | | Female | | Male | | Total |
|---|---|---|---|---|---|---|
| Groups | | SPE | GE | SPE | GE | |
| Region 1 (390) | Maths | 71 | 71 | 22 | 22 | 186 |
| | NS | 62 | 62 | 20 | 20 | 164 |
| | Hum | 15 | 15 | 5 | 5 | 40 |
| Region 2 (420) | Maths | 77 | 77 | 24 | 24 | 202 |
| | NS | 67 | 67 | 21 | 21 | 176 |
| | Hum | 16 | 16 | 5 | 5 | 42 |
| Region 3 (190) | Maths | 35 | 35 | 11 | 11 | 92 |
| | NS | 30 | 30 | 10 | 10 | 80 |
| | Hum | 7 | 7 | 2 | 2 | 18 |
| **Total** | | **380** | **380** | **120** | **120** | **1000** |
| **Total gender** | | **760** | | **240** | | |
| **Total Maths= 480, Total NS= 420, Total Hum= 100** | | | | | | |
| **Total SPE= 500, Total GE= 500** | | | | | | |

**Note:** NS= Natural Sciences, Hum= Humanities

As Table 5.3 shows, I considered Region, Gender, and Field of study to make the sample as representative as possible. As I was aiming for 1000 students, I was hoping that there would be students from a wide range of English background in the actual sample.

Another difference between the pilot study and the main study was that two more samples were going to be used in the main study i.e. a sample of letter writers and a sample of interviewees. Using NOET statistics and the pilot study sample as a guide, I considered the same factors in sampling the letter writers as I considered for

the questionnaire respondents, which included field of study, gender and region. Table 5.4 shows the sampling of the letter writers.

**Table 5-4 Sampling of letter writers**

| Region | Field | Public | | | | Total |
|---|---|---|---|---|---|---|
| | | Female | | Male | | |
| | | SPE | GE | SPE | GE | |
| Region 1 | Maths | +2,-2 | +2,-2 | +1,-1 | +1,-1 | 12 |
| | NS | +2,-2 | +2,-2 | +1,-1 | +1,-1 | 12 |
| | Hum | +1,-1 | +1,-1 | +1,-1 | +1,-1 | 8 |
| Region 2 | Maths | +3,-3 | +3,-3 | +1,-1 | +1,-1 | 16 |
| | NS | +3,-3 | +3,-3 | +1,-1 | +1,-1 | 16 |
| | Hum | +1,-1 | +1,-1 | +1,-1 | +1,-1 | 8 |
| Region 3 | Maths | +2,-2 | +2,-2 | +1,-1 | +1,-1 | 12 |
| | NS | +2,-2 | +2,-2 | +1,-1 | +1,-1 | 12 |
| | Hum | +1,-1 | +1,-1 | +1,-1 | +1,-1 | 8 |
| Total | | 34 | 34 | 18 | 18 | 104 |

**Note 1:** NS= Natural Sciences, Hum= Humanities

**Note 2:** + indicates strong student and – indicates weak student.

As Table 5.4 shows, I sampled twice as many female students as male students in the fields of Mathematics and Natural Sciences in Regions 1 and 3 and three times as many female students in Region 2. For Humanities, however, I sampled an equal number of male and female students in the three regions. Although I should have sampled more students of Mathematics than Natural Sciences and more female students than male students, I was concerned that this might make data analysis impractical.

As the table shows, I tried to have a balance between students of good (+ sign) and weak (- sign) English background (5.4.2). As the plus and minus signs indicate, for every strong student there was a weak one in the sample. Contrary to the sampling of the questionnaire respondents, I considered the English background in the sampling of the letter writers because they were limited and therefore I had to be selective.

To sample the interviewees, however, I did not consider gender and region for practical reasons (Tables 5.3 and 5.4). As the number of the interviewees was going to be much more limited than the questionnaire respondents and the letter writers, I only

considered 'field of study'. I also tried to have a balance between good students and weak students as in the sampling of letter writers.

**Table 5-5 Sampling of interviewees**

| Field | SPE | GE | Total |
|---|---|---|---|
| Maths | +1,-1 | +1,-1 | 4 |
| NS | +1,-1 | +1,-1 | 4 |
| Hum | +1,-1 | +1,-1 | 4 |
| **Total** | 6 | 6 | 12 |

Table 5.5 shows that for each field of study there were four students and for English background there was a balance between strong and weak students (5.4.3). The total number of the sampled interviewees was 12.

## 5.4. Data Collection

The data collection lasted about 2.5 months from mid-October 2006 until the end of January 2007. I collected questionnaire data and letter data at the same time and then carried out the interviews. In the following sections, I explain how I collected the data through each instrument.

### 5.4.1. Data collection through the questionnaire

There were some differences between the data collection processes in the main study and in the pilot study (4.2). As I mentioned in 4.2.1, in the pilot study it took the students 30-45 minutes to complete the three questionnaires. However, in the main study, there was one questionnaire which took the students about 20-30 minutes to complete.

### 5.4.2. Data collection through the letters

I asked for volunteers from the questionnaire respondents for writing the letters keeping in mind the factors which I considered for sampling. I checked the questionnaires of the students who returned them earlier than others for sampling information (e.g. English background) and if they matched the sampling criteria, I

asked them whether they were willing to write the letters. If they were willing, I wrote their name on the questionnaire which they had just returned to me after getting their permission (I did this to be able to check the consistency of the students' responses on the different instruments - 5.4.3 and 5.8). However, because usually good students (i.e. with good English background) volunteered, I also got help from the teacher, school principal or their assistants in identifying the weak students. Again there was no pressure on the students to get their agreement for cooperation. After the identification of the students, I gave them the letter writing instructions (see Appendix 10) which were typed in Farsi on A4 sheets of paper with enough space for them to write their advice on. Based on my pilot study experience (4.2.1), I had to ask the students to write the letters under time restriction. I asked the students to write the letters in class and return them to me rather than to the school later. At the top of the paper, I asked them to write their name, gender, region, field of study, and number of terms they had attended in English institutes for future reference. As I was going to select the interviewees from the letter writers, at the bottom of the sheet, I asked them to tick 'Yes' or 'No' depending on whether they were willing to be interviewed later. They were also asked, if they were willing, to provide their phone numbers and email addresses for appointments. I managed to collect a total of 91 letters including 42 letters from SPE students and 49 letters from GE students (Table 5.7). After collecting the letter data, I made a list of the students' suggestions in the letters to probe in the interviews. Initially the list was expanded and included all the issues raised in the letters, but after piloting the interview, I only listed the same activities addressed in the questionnaire in order to be able to check the students' consistency in their responses and possible changes in their views. The reason why I did not include beliefs in the list was that the beliefs addressed in the letters were different from those

addressed in the questionnaire.

### 5.4.3. Data collection through the interviews

I identified the interviewees based on the sampling information provided in the letters and the questionnaire (Table 5.8). Having identified the interviewees, I made appointments with them through the school principal or their assistant. In the interview sessions, first I asked the students to tick the list I had made from the suggestions in the students' letters and asked for permission to record the interviews which were conducted in Farsi. Then, I used the questionnaire items as the interview guide. I had two major types of questions to ask the students. One was to ask them whether they confirmed or disconfirmed their responses on the questionnaire and one was to ask why they did the previously reported activities (see Appendix 11). However, I allowed for the order of the questions to be flexible in case a fixed order might interrupt the flow of the interviewee's thoughts. Throughout the interviews I tried to establish and maintain rapport so that the students would be encouraged to talk. As I was able to make more appointments than I expected, I managed to interview 9 SPE and 9 GE students, which was more than the original sample. I made this decision so that the results would be more generalisable.

#### 5.4.3.1. Actual samples and the data collected

However ideal my sampling and data collection plan might be, I collected the data under some practical constraints. Therefore, the actual samples and the data might not truly reflect the original plan. The data I actually managed to collect are summarised in Tables 5.6, 5.7, and 5.8. The tables show to what extent the sample and the collected data represent the population.

**Table 5-6 Actual sample of questionnaire respondents / the collected questionnaire data**

| Categories | | SPE | GE | Total |
|---|---|---|---|---|
| **Gender** | Male | 18.6% | 49.8% | 38.6% |
| | Female | 81.4% | 50.2% | 61.4% |
| **Region** | 1 | 18.6% | 18.4% | 18.5% |
| | 2 | 78.7% | 50.1% | 60.3% |
| | 3 | 2.7% | 31.5% | 21.2% |
| **Types of schools** | Non-profit | 17.5% | 28.3% | 24.5% |
| | Public | 82.5% | 71.7% | 75.5% |
| **Fields of study** | Mathematics | 28.3% | 31.5% | 30.3% |
| | Natural Sciences | 43.4% | 39.4% | 40.8% |
| | Humanities | 28.3% | 29.1% | 28.8% |
| **Attendance in English institutes by terms** | Mean | 13.20 | 4.71 | 7.77 |
| | SD | 9.28 | 7.22 | 9.00 |
| **Total respondents (in percentage)** | | 35.7% | 64.3% | 100% |
| **Total respondents (in numbers)** | | 371 | 667 | 1038 |

**Table 5-7 Actual sample of letter writers / the collected letter data**

| Categories | | SPE | GE | Total |
|---|---|---|---|---|
| **Gender** | Male | 16.7% | 53.1% | 36.3% |
| | Female | 83.3% | 46.9% | 63.7% |
| **Region** | 1 | 16.7% | 18.4% | 17.6% |
| | 2 | 73.8% | 51.0% | 61.5% |
| | 3 | 9.5% | 30.6% | 20.9% |
| **Types of schools** | Non-profit | 19.0% | 34.7% | 27.5% |
| | Public | 81.0% | 65.3% | 72.5% |
| **Fields of study** | Mathematics | 38.1% | 36.7% | 37.4% |
| | Natural Sciences | 38.1% | 32.7% | 35.2% |
| | Humanities | 23.8% | 30.6% | 27.5% |
| **Attendance in English institutes by terms** | Mean | 15.00 | 7.55 | 10.93 |
| | SD | 10.45 | 7.13 | 9.50 |
| **Total respondents (in percentage)** | | 46.2% | 53.8% | 100% |
| **Total respondents (in numbers)** | | 42 | 49 | 91 |

**Table 5-8 Actual sample of interviewees / the collected interview data**

| Categories | | SPE | GE | Total |
|---|---|---|---|---|
| **Gender** | Male | 0% | 55.6% | 27.8% |
| | Female | 100% | 44.4% | 72.2% |
| **Region** | 1 | 44.4% | 0% | 22.2% |
| | 2 | 55.6% | 100% | 77.8% |
| | 3 | 0% | 0% | 0% |
| **Types of schools** | Non-profit | 0% | 33.3% | 16.7% |
| | Public | 100% | 66.7% | 83.3% |
| **Fields of study** | Mathematics | 33.3% | 44.4% | 38.9% |
| | Natural Sciences | 33.3% | 33.3% | 33.3% |
| | Humanities | 33.3% | 22.2% | 27.8% |
| **Attendance in English institutes by terms** | Mean | 17.78 | 6.67 | 12.22 |
| | SD | 9.05 | 9.01 | 10.46 |
| **Total respondents (in percentage)** | | 50% | 50% | 100% |
| **Total respondents (in numbers)** | | 9 | 9 | 18 |

### 5.4.4. Ethical considerations

I got consent for data collection in several ways. First of all, I got a permission letter from the head education office as well as letters from the education office of each city in which the schools were located. To collect data from the schools, I presented the letters to the school principals or their assistants. I got consent from individual questionnaire respondents on the questionnaires as well (see Appendix 12). As regards the letter writers who did not volunteer (mainly weak ones), their cooperation was requested and they had the option of refusing. I also obtained permission from the letter writers (volunteers and non-volunteers) to write down their name on the questionnaire. Interviews were voluntary; nevertheless, I got permission from the interviewees for recording their interviews. In general, I assured the students that I would protect their anonymity and did not include any personal or sensitive questions in any of the instruments (ibid: 440).

Throughout the data collection process, I tried to respect the rules of the schools. For example, in one case a math teacher did not agree to give his class time to me because he said he would fall behind schedule as the final exams were drawing near, or some classes had exams in which case I arranged for another day for those classes.

In order to give the students something in return for their time and cooperation, I offered to give them advice on exam preparation or on language learning in general (ibid: 484). A few of the students were in touch with me until they took their exam.

## 5.5. Preparing the data for analysis

### 5.5.1. Preparing the questionnaire data for analysis

The procedures for preparing the questionnaire data for analysis were the same as those of the pilot study, which I explained in 4.3.

### 5.5.2. Transcribing the letters and interviews

I transcribed the interviews from the audiotapes which I used during the interviews (Appendix 11). According to Silverman (2001), the degree to which transcription captures the details depends on what the researcher is trying to do in the analysis (p. 189). In this regard, Barker et al (2002: 222) state that 'the transcription method should be chosen for the task at hand'. What I was concerned about in the interview was *what* the student said not the nature of the interaction (Banerjee, 2003: 174). Therefore, I adopted simplified conventions for the transcription as it was not necessary to record every detail of the interviews. I used conventions only in cases where I thought the intended meanings by the interviewees would be distorted otherwise.

As far as the letter data are concerned, I did not need to transcribe them but only type what the students gave to me in the form of hand written scripts in Farsi (Appendix 10). I transcribed both the interviews and the letters in Farsi as they were going to be analysed in the same language.

## 5.6. Data Analysis

### 5.6.1. Analysis of questionnaire data

The procedures for the analysis of questionnaire data were the same as those of the pilot study which I explained in 4.4.

### 5.6.2. Coding the letter and the interview data

In this section, I will review the literature on coding focusing on two major approaches and will demonstrate how I applied the theory to develop an analytic framework for the letter and the interview data.

Marshall and Rossman (2006: 156) explain the stages of data analysis as follows. They suggest that revisiting the piles of data is very important before the

actual analysis begins. They state that the researcher needs to list the data that have been gathered, do minor editing if necessary, clean up the data that may seem unmanageable, and organise the data in terms of where and when they have been gathered and with whom. At this time, the data are ready to be entered into a software program for management or analysis (ibid). The next stage is immersion in the data which involves the researcher reading and rereading the data so that they become intimately familiar with them (ibid). Generating categories and themes is the next stage which requires awareness of and focused attention to the data and openness to implicit meanings (ibid). The codes may be in the form of abbreviations, colour dots, numbers, etc (ibid). In software programs, however, abbreviations are typically used (ibid). In this regard, Silverman (2001) suggests that the researcher begins with a small dataset or one case and generates a provisional analytic scheme or hypothesis. Then s/he goes on to test the hypothesis or scheme against other data (p. 238). If the scheme doesn't apply to the new cases, it is reformulated (ibid). The process of reformulation or modification against deviant cases goes on until the scheme can incorporate all the data (ibid).

Authors talk about two approaches for this stage, an a priori approach and a grounded approach. Concerning the a priori approach, Kvale (1996) believes that the researcher can come to the transcripts with the thematic questions she or he asked at the start of the inquiry (see also Seidman, 2006, and Marshall and Rossman, 2006). He states that the preliminary research questions, the related literature, as well as the pilot study can be used to develop several categories.

Regarding grounded approach, Seidman (2006) states that the researcher has an open attitude, lets the data 'speak for itself' and seeks what emerges from the data as important and interesting (p. 117). The researcher identifies the salient themes,

recurring ideas, and patterns of belief that link people and settings together uncovering patterns, themes, and categories (ibid). The emerging categories would serve as containers in which sections of texts can be placed (ibid).

Using the literature reviewed above as a guide, I applied it to develop an analytic framework. In the following section, I discuss the procedure for developing the framework.

After transcribing the data (5.5.2), I organised them according to Marshall and Rossman's suggestions. I drew up tables (Table 5.7 and Table 5.8) in which I included information about the characteristics of the data. I also took note of the time when I collected these data.

After organising the data, I read them in order to do some minor editing if necessary. Still using the paper copy of the transcripts, I reread the data in order to intimately familiarise myself with them while looking for tentative patterns. Then I entered the data into Atlas.ti software (Muhr, 2003- 2004) for categorisation and coding.

In order to categorise and code the data, initially I made use of the themes in the research questions and the questionnaire as the super-ordinate, a priori categories. They showed that 'reported beliefs', 'reported activities', 'purposes', and 'knowledge of the test' were the main themes. Then I went through the texts and coded the parts of the texts with these general labels. Then I broke these general labels into more fine-grained ones. For this purpose, I used the subcategories which were already available in the questionnaire. For example, one of the questionnaire items asked the students to rank their 'test preparation purposes' (super-ordinate category) in terms of their importance to them. The subcategories in the questionnaire included the options which reflected the various purposes the students were supposed to rank, for example

'passing the exam', 'interest', 'traveling abroad'. As I had asked the students about their 'test preparation purposes' in the interviews, I expected to find the same subcategories in the transcripts. Therefore, again I went through the texts and coded the parts which matched these subcategories. So far, the development of the framework was based on a priori or theory-driven themes.

However, there were times when students talked about issues in the interviews and the letters that were not addressed by the research questions or the instruments beforehand. These new categories were grounded categories which made the framework more complex. Therefore, I did not adopt an entirely a priori approach but had an open attitude and let the data 'speak for itself'. In other words, I looked to see what new themes and categories emerged from the data. For example a student said in the interview that she was preparing for the SPE Test for the 'development of her speaking ability'. Therefore, I added this new subcategory to the list of the various 'test preparation purposes' (see Appendix 13 for the codes).

### 5.6.3. Reliability of the coding

There is always the risk that researchers may do coding unreliably. Patton (2002) cautions the researcher against leading and imposing categories. Seidman (2006: 127) warns that the analyst may be tempted to force data into categories or use categories which might better reflect the researchers' world than the research participants' world. To avoid these problems, Silverman (2001: 229) suggests ensuring inter-rater reliability of the codes, arguing that categories should be standardised, so that any researcher would categorise in the same way. The process of examining inter-rater reliability involves giving the data to a number of analysts and asking them to code the data based on an agreed set of categories (ibid).

In this study, I checked the inter-rater reliability of my coding in three stages,

at the beginning, in the middle, and at the end of the development of the coding framework. In the first two stages, I gave one extract from the letters and one extract from the interviews and my tentative categories to a number of staff members and research students interested in language teaching and language testing from the Department of Linguistics in Lancaster University. These two stages were mainly spent on receiving feedback and coming to an agreement on a set of categories. In the last stage, however, I calculated the inter-rater reliability i.e. the percentage of agreement between the coders. First of all, I sent an email to three research students and requested their help. Then, I sent them two interview and two letter extracts which I coded before, a table including the codes in the form of numbers with meanings of the codes, instructions for coding, and some brief background information about my study (Appendix 13). The inter-rater reliability turned out to be 93% which was reasonably high (Appendix 14). However, as the figure shows, there was 7% unreliability which could be attributed to the coders' insufficient familiarity with my study. Although I gave the coders some background information about this study, it might not have been sufficient for adequate knowledge of the research context, themes and foci of the study, and as much familiarity with the data as the researcher.

## 5.7. Consistency of the students' responses across different instruments

As mentioned before, I used three instruments and a list from the students' letters for data collection. As the data were collected within 2.5 months, I was interested to see to what extent the students were consistent in their responses or whether there were any changes in their views. Students whose responses were examined were 9 SPE and 9 GE students who had responded to all the four instruments. However, there was a problem with comparability of the responses which

I had to resolve first. The responses were given to instruments which had different types of scales. While the questionnaire items were based on scales of 5, 6, and 7 points, the letter, the list, and the interview were based on two-point/yes-no scales. In the letters the students either recommended certain activities (yes) or did not recommend them (no); in the list which was derived from the letters the students either ticked (yes) or did not tick (no) certain activities as necessary; and finally in the interview before I asked the students for the reasons for their reported activities, I asked them whether they confirmed (yes) or disconfirmed (no) their responses on the previous instruments. The reason I only considered responses to the confirmation questions and not responses to the 'why' questions was that the latter were not comparable with the responses on the other instruments. I made other adjustments on the scales which are detailed in Appendix 15. Therefore, the consistency figures may reflect a rough estimation rather than the exact calculation. After the adjustments, I calculated the percentage of agreement of each individual student's responses on the different instruments. The average agreements for the SPE and GE groups were 0.72 and 0.77 respectively (Appendix 15). These figures show that in both groups there was some degree of inconsistency which suggests that the students' attitudes might have undergone some changes during the data collection period. The questionnaire and the letter data were collected on one day approximately during the first two third of the data collection period, and data from the list and the interviews were collected on one day near the end of the period. Therefore, some change of attitudes might have occurred in between. I present a few examples which show that some changes occurred in the students' attitudes.

Concerning practice tests GE 4 said 'it's a good idea but I don't do it' in the questionnaire but he ticked it in the 'list' and in the interview also he said he did it. As

regards the inconsistency between his response in the questionnaire on the one hand and the list and the interview on the other, he said this was because he had not quite started using practice tests when he was answering the questionnaire. Therefore, some of the different answers might be due to the point of time when the students were preparing for the test.

Regarding the use of extra materials, SPE 7 and 8's responses slightly changed from those on the questionnaire. While on the questionnaire they reported that they spent more time on the Pre-university book, in the interview they reported that both the Pre-university book and extra materials were necessary for the SPE Exam. However, they were not sure which they had to spend more time on and said the emphasis on either type of materials would depend on the teacher. GE 6, 7, and 8 reported on the questionnaire that they spent an equal amount of time on the Pre-university book and extra materials, while in the list and in the interview they reported spending more time on the Pre-university book. When I asked them about these inconsistencies, GE 6 and 8 said that what they meant by extra materials was actually supplementary materials and GE 7 said what she had in mind was school books of the previous years but that now she thought the Pre-university book itself included all the points of the previous years.

Concerning the weight of English versus non-English, SPE 3, who had reported in the questionnaire that she was focusing on English rather than on Non-English subjects, said the opposite in the interview. When I asked about the contradiction, she emphasised that 'at present' i.e. only temporarily, she was focusing on Non-English subjects, which was not what she generally believed, that more emphasis should be given to English.

Finally, when I asked GE 7 why he ranked Grammar 3 on the questionnaire

i.e. less important than reading, he replied that at that time he thought improving reading speed was more important than Grammar.

As the consistency figures showed, the SPE Group was less consistent than the GE Group. The lower consistency of the SPE Group suggests that their attitudes might have undergone more changes than those of the GE Group during the data collection period. This could have been due to the majority of the SPE students' attendance in preparation classes or their allocation of more time to studying English.

However, in addition to the time factor which may have contributed to the change of attitudes (and hence, inconsistencies), there were a number of issues with the instruments which could have given rise to the inconsistencies.

First, the instruments were different in approach. While the questionnaire, the list, and the interview were theory-driven, the letter was grounded i.e. in the questionnaire, the list, and the interview, the students were asked what specific issues to report on, while in the letter only a general topic was given and the students themselves decided what specific topics to write about. In addition, letters were written under time restrictions. Therefore, due to the memory factor, the students might not have remembered to write everything they wanted to or might not have had enough time to write everything.

Second, while the questionnaire and the interviews were concerned with what activities the students reported they 'did', the letters and the list were concerned with what activities the students 'recommended' or 'considered essential', and doing certain activities and recommending certain activities (or considering them as essential) may not necessarily be the same, i.e. students may recommend certain activities but may not do them themselves for various reasons.

# Chapter 6. Results from the Questionnaire Data

In Chapters 4 and 5, I discussed the research methodology of the study including the aims and contents of the instruments and their validation, sampling, data collection, and methods of data analysis. This chapter as well as the next two chapters will focus on the results from three sets of collected data i.e. questionnaire data, letters, and interviews. In this chapter, I will present the results of the questionnaire data. The results will be presented in the order of the research questions, and predicted versus unpredicted areas of washback. Predicted washback is based on the features of the SPE Test and unpredicted washback examines what else was happening under the influence of the test.

## 6.1. RQ1: What activities do the SPE students report doing in order to prepare for the SPE Test?

### 6.1.1. Predicted washback

#### 6.1.1.1. Test Sections

As regards the test sections, I asked the students how much time they spent on each section of their tests, i.e. I asked them to rank the amount of time from most to least. As the number of items in each section of the SPE Test was more than the number of items in each section of the GE Test, I expected that SPE students would say they spent more time on each section than GE students. Also based on the number of items in the GE and SPE tests (Appendix 3), I expected that both groups would rank 'vocabulary' first. However, since 'reading', 'cloze', and 'grammar' had the same number of items in the GE Test, it was not predictable which would come next. Similarly for the SPE Group, the order of 'reading' and 'cloze' was not predictable because of the same number of items except that they were expected to be placed somewhere below 'vocabulary'. However, I expected 'grammar' to follow 'reading'

and 'cloze' by 'language functions' and 'sentence structure' which again had the same number of items.

To analyse this item (See Figure 6-1), I calculated the mean of each test section. The section with the lowest mean would be the one on which most time was spent and the section with the highest mean would be the one on which the least time was spent. The different mean values allowed me to examine how SPE and GE groups ranked the test sections in order of the amount of time they spent on them. I compared the two groups using Independent-Samples T-Test. The results are presented in Table 6.1 preceded by the questionnaire item through which I collected the data.

*Please number from 1 to 6 the sections of the SPE/GE Test on which you spend most time 'during test preparation' where 1= most time and 6= least time.*

| Grammar | Vocabulary | Sentence Structure | Language Functions | Cloze Passage | Reading |
|---|---|---|---|---|---|
| | | | | | |

**Figure 6-1 Questionnaire Section 4 Item 28**

**Table 6-1 Ranking of the amount of time spent on SPE and GE test sections**

| Test Sections | Mean | | Significance | Resultant ranking | |
|---|---|---|---|---|---|
| | SPE | GE | | SPE | GE |
| **Grammar** | 2.66 | 2.95 | .005 | Vocabulary | Vocabulary |
| **Vocabulary** | 1.77 | 1.96 | .021 | Grammar | Reading |
| **Reading** | 2.70 | 2.92 | .034 | Reading | Grammar |
| **Cloze** | 3.50 | 3.77 | .003 | Cloze | Cloze |
| **Sentence Structure** | 4.18 | 4.37 | .200 | Language Functions | Sentence Structure |
| **Language Functions** | 4.12 | 4.38 | .104 | Sentence Structure | Language Functions |

The P-values in Table 6.1 show that the two groups were significantly different on the common test sections (grammar, vocabulary, reading, and cloze) but not on the new sections. The results also show that SPE mean values are lower than the GE mean values suggesting that SPE students spent more time on each section than GE students. This might be due to the increase in the number of SPE Test items (Appendix 3) which resulted in the increase in the weight of English versus non-English subjects. However, the lack of significant difference in the new sections might mean that the new sections were not functioning efficiently in that they probably did

not encourage additional or different activities on the part of the SPE students. Whether the significant or non-significant differences were due to the test will be examined later in this chapter (Table 6.19) as well as in the discussions of the results from the letters (Tables 7.1 and 7.8) and the interviews (Table 8.14).

Based on the mean values in Table 6.1 (lowest mean= most time, highest mean= least time) SPE and GE students ranked the different sections of their tests almost similarly. Both groups of students spent most time on 'vocabulary'. While the SPE Group ranked 'grammar' second and 'reading' third, the GE Group ranked 'reading' second and 'grammar' third. 'Cloze' was ranked fourth by both groups, followed by 'language functions' and 'sentence structure' (with no significant difference) in the lowest position. The lack of total agreement between the number of items and the rating of the amount of time spent on the test sections might suggest that factors other than the test might have affected the students' ranking (see Table 8.14).

#### 6.1.1.2. Oral skills versus the test sections

The results presented in Table 6.2 are from the item which examined the status of oral skills activities in comparison with the skills tested on the SPE Test. The reason for the inclusion of listening and speaking in the item was that although oral skills are not tested in the SPE Test, they are recommended by the book, *Bridging the Gap*. Therefore, it was interesting to see whether this recommendation could make any difference between SPE and GE students and whether the textbook had a stronger influence than the test.

I expected that both groups would report spending less time on oral skills than on vocabulary, grammar, and reading, but expected that SPE students would report doing more oral activities than GE students. The item and the results are presented below.

*Please number from one to five the following language skills on which you spend most time where 1= most time and 5 = least time.*

| Vocabulary | Grammar | Reading | Listening | Speaking |
|---|---|---|---|---|
| | | | | |

**Figure 6-2 Questionnaire Section 4 Item 29**

**Table 6-2 Ranking of the SPE and GE test sections by SPE and GE groups**

| Test Sections | Mean | | Significance | Resultant ranking of the skills | |
|---|---|---|---|---|---|
| | SPE | GE | | SPE | GE |
| **Vocabulary** | 2.20 | 2.43 | .018 | **Vocabulary** | **Vocabulary** |
| **Grammar** | 2.84 | 3.17 | .001 | **Reading** | **Reading** |
| **Reading** | 2.75 | 2.59 | .125 | **Grammar** | **Grammar** |
| **Listening** | 3.57 | 3.46 | .352 | **Speaking** | **Speaking** |
| **Speaking** | 3.28 | 3.36 | .493 | **Listening** | **Listening** |

As I expected, Table 6.2 shows that listening and speaking have higher mean values than vocabulary, grammar, and reading, which means that the students spent less time on them. The lack of significant difference between the two groups (.352 and .493) suggests that probably the influence of the test was stronger than *Bridging the Gap*. In the discussion of the interviews also (Chapter 8), I will discuss why the students considered oral skills less important than vocabulary, grammar, and reading.

Concerning the ranking of the test sections, the SPE Group ranked 'reading' above 'grammar', which is different from the results presented in Table 6.1. This could be due to the wording of the question or because of some overlap with the item in Figure 6-1, the students might have considered it a repeat question and therefore not have taken it seriously. These contradictory results also suggest that further investigation is needed into the constructs of the SPE Test.

#### 6.1.1.3. Extra materials versus Pre-university textbook

The following item concerned the use of school materials and extra materials. I expected that both groups would use the Pre-university book but SPE students would use more extra materials than GE students (Appendix 3).

*On which of the following types of materials do you spend more time? Please tick the appropriate box.*

| Pre-university textbook | Extra materials | I spend equal amount of time on them |
|---|---|---|
| | | |

**Figure 6-3 Questionnaire Section 4 Item 30**

**Table 6-3 Percentage of students spending time on extra materials and Pre-university book**

| | Pre-university textbook | Extra materials |
|---|---|---|
| **SPE** | 11.6 | 39.6 |
| **GE** | 33.9 | 16.9 |
| **Significance** | .000 | .000 |

The results in Table 6.3 show that the two groups of students were significantly different in their use of school and extra materials. The descriptive statistics show that more SPE students spent time on extra materials than the Pre-university textbook, while more GE students spent time on the Pre-university textbook. The results confirmed my prediction that SPE students would use more extra materials than GE students. However, it is not clear whether this is due to the requirement of the SPE Test or other factors (see Table 6.19). Therefore, it was worth following up the issue by asking SPE students why they spent more time on extra materials than the Pre-university textbook and by asking GE students why they spent more time on the Pre-university book than on extra materials (8.2.4).

#### 6.1.1.4. The Use of *Bridging the Gap*

Since the book *Bridging the Gap* was introduced by NOET as supplementary material for the SPE students, it was worth examining how much time the students spent on this book than on other types of extra materials.

*Among extra materials, on which of the following do you spend more time? Please tick the appropriate box.*

| *Bridging the Gap* | Other extra materials | I spend equal amount of time on them |
|---|---|---|
| | | |

**Figure 6-4 Questionnaire Section 4 Item 31**

I expected that SPE students would use *Bridging the Gap* more than other types of extra materials (Appendix 3). The results are presented in Table 6-4.

**Table 6-4 Percentage of students spending time on *Bridging the Gap* and other extra materials**

| | *Bridging the Gap* | Other extra materials |
|---|---|---|
| **SPE** | 18.9 | 39.6 |
| **GE** | 0.4 | 2.1 |
| **Significance** | .945 | .000 |

The results show that there was no significant difference between SPE and GE students concerning the amount of time they spent on *Bridging the Gap*. This was probably because only a few GE students tried this item, presumably because it was not relevant to them. A possible reason for the little use of the book by the SPE students (although the book was the main book specifically designed for them) could have been because the students were only at the beginning of their test preparation period and had barely started using this book. .

However, the SPE and GE students were statistically different in the use of 'other extra materials'. Possible reasons for the use of extra materials including *Bridging the Gap* will be investigated in the interviews.

#### 6.1.1.5. Preparation classes

The following item asked whether the students were attending a preparation class. My expectation was that the great majority of SPE students would attend preparation classes but the great majority of GE students would not. The item and the results are presented in Figure 6-5 and Table 6-5.

*Are you attending a preparation class?*

| Yes | No |
|---|---|
| | |

**Figure 6-5 Questionnaire Section 4 Item 27**

**Table 6-5 Percentage of the students attending preparation classes**

| SPE | | GE | | Significance |
|---|---|---|---|---|
| **Yes** | **No** | **Yes** | **No** | |
| 60.6 | 38.5 | 29.7 | 69.4 | .000 |

According to Table 6.5, the two groups of students attended preparation

classes significantly differently. While the majority of SPE students (60.6%) attended preparation classes, the majority of GE students (69.4%) did not. This might be because SPE Test questions come from outside the Pre-university book but GE questions from inside the book (Table 6.19). However, the question to be probed further is what reasons the students had for attending or not attending preparation classes.

#### 6.1.1.6. English VS Non-English courses

One of the features of the SPE Exam was that more weight was given to English than to Non-English courses (Appendix 3). Therefore, my expectation was that SPE students would spend more time on English than on non-English subjects but that GE students would spend more time on non-English than on English. Through the following item (See Figure 6-6), I asked the students how much effort they put into each of these subjects.

*On which of the following subjects do you spend more time? Please tick the appropriate box.*

| SPE/GE | Non-English Courses | I put in equal amount of time on them |
|---|---|---|
| | | |

**Figure 6-6 Questionnaire Section 4 Item 32**

**Table 6-6 Percentage of SPE and GE students spending time on English and Non-English subjects**

| | SPE/GE | Non-English courses |
|---|---|---|
| **SPE** | 38.0 | 12.9 |
| **GE** | 2.2 | 70.2 |
| **Significance** | .010 | .000 |

The two groups of students studied English and non-English courses significantly differently. The reason why greater number of SPE students spent more time on English and greater number of GE students spent more time on non-English might be because of the different requirements of the two tests (Table 6.19). However, the reason why a considerable number of SPE students spent time on Non-English

courses might be because of the way these students planned their study activities, which means that later they might change their focus of the activities. Therefore, it would be interesting to collect data at different points of time in the test preparation period to track possible changes in the students' activities. The results may also suggest that test authorities should reconsider the weighting of the subtests.

#### 6.1.1.7. Summary of the results for predicted reported activities

In answer to Research Question 1, I examined the reported activities of the students which I had made predictions about based on the features of the SPE Test. I present a summary of the results for the predicted washback in Table 6.7 with some explanations.

**Table 6-7 Summary results for the predicted reported activities**

| Activities- weight | SPE | GE |
|---|---|---|
| Vocabulary | Longer | Shorter |
| Grammar | Longer | Shorter |
| Reading | Longer | Shorter |
| Cloze | Longer | Shorter |
| Sentence Structure | No difference | |
| Language Functions | No difference | |
| Weight of English | Longer | Shorter |
| Weight of Non-English | *Shorter* | *Longer* |
| Listening | No difference | |
| Speaking | No difference | |
| **Activities- school** | | |
| Pre-university book | *Shorter* | *Longer* |
| **Activities- out of school** | | |
| Extra materials | Longer | Shorter |
| *Bridging the Gap* | No difference | |
| Preparation class | Longer | Shorter |
| **Test difficulty** | | |
| Test Difficulty | Higher | Lower |
| **English background** | | |
| Total attendance in English institutes | Longer | Shorter |
| Self-assessment | Higher | Lower |

**Note 1:** 'longer' and 'shorter' compare the amount of time spent on the activity by the SPE and GE groups. **Note 2:** 'higher' and 'lower' refer to the level of test difficulty and level of English ability through self-assessment. **Note 3:** where GE's average is higher, it is indicated by italics.

The results showed that SPE students were statistically different in most of the areas where washback was predicted. They spent longer on the sections which the

SPE Test had in common with the GE Test and on out-of-school resources including extra materials and preparation classes. This was further confirmed by the fact that they spent longer on English than on non-English subjects. SPE students also considered the SPE Test to be more difficult than the GE Test and had a better English background than GE students. Whether these differences were due to the test or not will be examined in 6.2. However, there were also cases in which SPE students were not statistically different from GE students, including the amount of time they spent on the new sections, doing oral skills, and using *Bridging the Gap*. While it was not expected that SPE students would be different from GE students in doing oral skills because of the features of the SPE Test, it was speculated that the lack of differences in the use of *Bridging the Gap* and the time spent on the new sections might be because of factors other than the test. As far as reading activity is concerned, the results were contradictory (Tables 6.1 and 6.2).

### 6.1.2. Unpredicted washback

#### 6.1.2.1. Specific vocabulary learning activities

The following items asked about the frequency with which the students reported doing various activities for learning vocabulary. I used frequency, mean, and Independent-Samples T-Test to analyse these items. For easier processing of the results, I combined frequency categories.

*3. I memorise word meanings.*
*18. I use new words in sentences that I make.*
*26. I use English-to-English dictionary.*
**Figure 6-7 Questionnaire Items for Vocabulary Activities**

**Table 6-8 Frequency with which SPE and GE groups reported doing various vocabulary activities**

| | A good idea/Never | | Rarely-Sometimes | | Often-Always | | Mean | | Significance of the means |
|---|---|---|---|---|---|---|---|---|---|
| | SPE | GE | SPE | GE | SPE | GE | SPE | GE | |
| **3** | 4.9 | 6.3 | 22.9 | 22.6 | 71.2 | 70.2 | 3.98 | 3.85 | .053 |
| **18** | 15.1 | 39.3 | 39.6 | 37.3 | 43.9 | 21.7 | 3.19 | 2.29 | .000 |
| **26** | 10.5 | 37.6 | 25.6 | 35.7 | 63.1 | 26.1 | 3.65 | 2.43 | .000 |

The descriptive statistics in Table 6.8 show that both groups of students memorised meanings of words with similar frequencies: as the significance figure (.053) shows, the difference between the two groups was not significant. This might be because there is a list of words at the end of both *English for Pre-University Students* (Pre-University book) and *Bridging the Gap*, which could have encouraged students to memorise words. However, it would be interesting to see in the interviews whether the students memorised because of the test, the textbooks, or other factors.

The students also reported using new words (Item 18) and using an English-to-English dictionary (Item 26). However, SPE students did these activities more often than GE students, and significantly differently. The question that might be asked in the interview involved why some students used new words in sentences that they made, and used an English-to-English dictionary, but others did not.

#### 6.1.2.2. Specific grammar learning activities

The following items asked about the frequency with which the students did various activities for learning grammar. I again used frequency, mean, and Independent-Samples T- Test to analyse these items.

*5. I practice grammar by making sentences*
*11. I memorise grammatical rules*
*15. I use texts to learn grammar*
**Figure 6-8 Questionnaire Items for Learning Grammar**

**Table 6-9 Frequency with which SPE and GE groups reported doing various grammar activities**

| | A good idea/Never | | Rarely-Sometimes | | Often-Always | | Mean | | Significance of the means |
|---|---|---|---|---|---|---|---|---|---|
| | SPE | GE | SPE | GE | SPE | GE | SPE | GE | |
| **5** | 18.3 | 39.4 | 44.5 | 41.1 | 35.0 | 17.8 | 2.86 | 2.19 | .000 |
| **11** | 16.7 | 15.3 | 27.2 | 30.6 | 53.4 | 52.8 | 3.34 | 3.36 | .965 |
| **15** | 4.0 | 13.9 | 23.2 | 27.1 | 70.9 | 56.8 | 3.98 | 3.43 | .000 |

According to Table 6.9, a greater number of SPE students learned grammar by making sentences than GE students and they were significantly different in this respect. However, there was a large number of students in both groups who did not do so. Therefore, it would be interesting to see why some students practiced grammar by making sentences but some others did not.

Based on the significance figure for Item 11, the two groups of students memorised grammatical rules with no significant difference. As Table 6.3 showed, they were not significantly different in memorisation of word meanings either. Nevertheless, it would be interesting to investigate in the interviews the reasons why the students memorised or did not memorise grammar or vocabulary.

As for Item 15, the two groups of students were significantly different in using texts to learn grammar. The number of SPE students who 'often or always' did this activity was more than the GE students, while the GE students were greater in number in terms of not doing this activity or doing it less often.

#### 6.1.2.3. Specific reading activities

The following items asked about the frequency with which the students reported doing various reading activities.

*1. I read English newspapers and magazines*
*7. I translate texts into Farsi while reading*
*13. I try to guess the meaning of new words from context*
*16. I read English story books*
*17. I pay attention to the topic of the text when I read*
*19. I read the text first to get a general idea and then go back to read it more carefully*
*21. I read various English texts as much as I can*

*25. I make summaries of the information that I read in English*
**Figure 6-9 Questionnaire Items for Reading Activities**

**Table 6-10 Frequency with which the SPE and GE groups reported doing various reading activities**

| | A good idea/Never | | Rarely-Sometimes | | Often-Always | | Mean | | Significance of the means |
|---|---|---|---|---|---|---|---|---|---|
| | SPE | GE | SPE | GE | SPE | GE | SPE | GE | |
| 1 | 37.2 | 73.0 | 49.9 | 22.3 | 11.1 | 4.3 | 2.04 | 1.43 | .000 |
| 7 | 6.7 | 15.9 | 27.2 | 41.2 | 63.9 | 41.1 | 3.77 | 3.10 | .000 |
| 13 | 3.8 | 15.0 | 23.7 | 36.1 | 70.9 | 48.0 | 3.97 | 3.23 | .000 |
| 16 | 13.5 | 55.8 | 42.9 | 29.1 | 42.0 | 13.5 | 3.18 | 1.87 | .000 |
| 17 | 2.2 | 6.7 | 10.5 | 21.1 | 84.9 | 68.4 | 4.37 | 3.86 | .000 |
| 19 | 3.2 | 13.2 | 15.4 | 26.8 | 80.1 | 58.8 | 4.21 | 3.55 | .000 |
| 21 | 9.7 | 43.3 | 47.7 | 40.9 | 41.0 | 14.7 | 3.19 | 2.07 | .000 |
| 25 | 19.9 | 35.5 | 40.4 | 41.2 | 36.9 | 22.2 | 2.95 | 2.35 | .000 |

As Table 6-10 shows, SPE and GE students reported doing all the above reading activities significantly differently with SPE students doing these activities more frequently than GE students. All the items in Table 6.10 (except for Item 7) have to do with context of use and using English. It would be interesting to see in the interviews if there is any reported cause-effect relationship between the test on the one hand and the students' use of context (guessing the meaning of new words from context, paying attention to the topic of the text, getting a general idea of the text before careful reading) and using English (Items 1, 16, 21, 25) on the other. It would also be interesting to see why the majority of the students used their mother tongue to understand texts, i.e. translate.

#### 6.1.2.4. Writing Activity

The following item is concerned with learning English through writing diaries or daily notes.

*4. I write things like diaries or daily notes in English*
**Figure 6-10 Questionnaire Item for the Writing Activity**

**Table 6-11 Frequency with which SPE and GE groups did the writing activity**

| | A good idea/Never | | Rarely-Sometimes | | Often-Always | | Mean | | Significance of the means |
|---|---|---|---|---|---|---|---|---|---|
| | SPE | GE | SPE | GE | SPE | GE | SPE | GE | |
| 4 | 34.2 | 67.6 | 41.5 | 22.9 | 20.8 | 8.5 | 2.37 | 1.60 | .000 |

The two groups of students wrote diaries or notes significantly differently. The SPE students used English in writing more than the GE students. However, overall the mean scores show that this activity was not very frequent for each group. Therefore, the question to be probed in the interview was why some students did this activity and why some others did not.

### 6.1.2.5. Reading and writing activities through the Internet

The following items asked about the frequency with which the students did Internet activities of reading and writing. As with previous sections, frequency, mean, and significance tests were used to analyse these items.

*9. I use the Internet in English.*
*24. I send and receive emails in English*
**Figure 6-11 Questionnaire Items for Internet Activities of Reading and Writing**

**Table 6-12 Frequency with which SPE and GE groups reported doing Internet activities**

| | A good idea/Never | | Rarely-Sometimes | | Often-Always | | Mean | | Significance of the means |
|---|---|---|---|---|---|---|---|---|---|
| | SPE | GE | SPE | GE | SPE | GE | SPE | GE | |
| **9** | 38.8 | 50.2 | 26.7 | 22.2 | 31.3 | 24.6 | 2.60 | 2.22 | .001 |
| **24** | 44.7 | 55.2 | 20.8 | 22.2 | 30.7 | 20.4 | 2.48 | 2.06 | .001 |

The SPE and the GE students were significantly different in the use of the Internet. SPE Group reported doing these activities more frequently than GE Group. However, it is not clear whether the students who used the Internet more frequently thought that they were useful for the exam or whether those who did this activity less frequently thought that they were not as useful. It is also not clear whether the students who used the Internet were driven by the test to use it or they used it because they simply liked to use the Internet.

### 6.1.2.6. Specific oral Activities

#### 6.1.2.6.1. Listening and speaking activities

Item 8 asked how often the students listened to English programs on the radio and Item 12 asked how often they approached people to talk to in English.

*8. I listen to English programs on the radio*
*12. I look for people I can talk to in English*
**Figure 6-12 Questionnaire Items for Listening and Speaking Activities**

**Table 6-13 Frequency with which SPE and GE groups reported doing listening and speaking activities**

| | A good idea/Never | | Rarely-Sometimes | | Often-Always | | Mean | | Significance of the means |
|---|---|---|---|---|---|---|---|---|---|
| | SPE | GE | SPE | GE | SPE | GE | SPE | GE | |
| 8 | 44.2 | 71.7 | 42.9 | 20.1 | 10.5 | 6.1 | 1.98 | 1.48 | .000 |
| 12 | 12.4 | 43.9 | 32.3 | 35.7 | 53.6 | 18.7 | 3.49 | 2.16 | .000 |

There was a significant difference between SPE and GE students in listening to English programs on the radio. Only 10.5% of the SPE and 6.1% of the GE students did this activity, but the great majority from both groups did not do it at all, which confirms the results in Table 6.2.

However, concerning speaking, the majority of SPE students (53.6%) reported looking for people to speak English with, while only a small minority of GE students did this activity. However, it is not clear whether the students did this activity because of the test or factors such as interest. The results also show that more students from both groups did the speaking activity compared to the listening activity. In this regard, the results agree with those in Table 6.2 where students ranked 'speaking' higher than 'listening'.

#### 6.1.2.6.2. Audiovisual activities

The following items asked about the frequency with which the students reported doing various activities involving audios and videos.

*2. I use tapes or CD's to practice English*
*6. I watch English films or programmes*

*10. I learn English through computer games*

**Figure 6-13 Questionnaire Items for Activities involving Audios and Videos**

**Table 6-14 Frequency with which the SPE and GE groups reported doing various audiovisual activities**

| | A good idea/Never | | Rarely-Sometimes | | Often-Always | | Mean | | Significance of the means |
|---|---|---|---|---|---|---|---|---|---|
| | SPE | GE | SPE | GE | SPE | GE | SPE | GE | |
| **2** | 19.7 | 63.9 | 50.9 | 26.8 | 27.2 | 8.4 | 2.74 | 1.64 | .000 |
| **6** | 12.1 | 33.3 | 43.9 | 37.8 | 41.8 | 27.6 | 3.12 | 2.53 | .000 |
| **10** | 56.9 | 52.6 | 28.3 | 28.0 | 13.2 | 17.2 | 1.87 | 2.02 | .086 |

The SPE and GE students were significantly different in using tapes or CDs to practice English and in watching English films or programmes. SPE students reported doing these activities more frequently than GE students. However, the two groups did not use computer games significantly differently, although according to the descriptive statistics it was the GE Group who used computer games more frequently. However, as with other activities discussed in this section, it is not clear whether the students who reported doing these activities reported doing them because of the test or other factors, although there is some indication of the relationship between the students' knowledge of the test and the activities they report.

### 6.1.2.7. Errors

The following item asked about students' reported activities in terms of whether they gave priority to communication or accuracy.

*23. I encourage myself to write or speak English even when I'm afraid of making mistakes.*

**Figure 6-14 Questionnaire Item for Asking about Priority of Learning**

**Table 6-15 Frequency with which SPE and GE groups used English despite the possibility of errors**

| | A good idea/Never | | Rarely-Sometimes | | Often-Always | | Mean | | Significance of the means |
|---|---|---|---|---|---|---|---|---|---|
| | SPE | GE | SPE | GE | SPE | GE | SPE | GE | |
| **23** | 7.5 | 25.3 | 29.1 | 39.7 | 61.7 | 32.7 | 3.69 | 2.71 | .000 |

Based on the results in Table 6.15, more SPE students used English despite fear of making mistakes than GE students who were mainly concerned about

accuracy. The two groups were statistically different. However, the question to be probed includes why the students were concerned about accuracy i.e. whether it was because of the test.

### 6.1.2.8. Mock exams and practice tests

The following items ask about the frequency with which students took mock exams and practiced with sample test questions.

*14. I take mock SPE/GE Exams*
*22. I practice with sample test questions*
**Figure 6-15 Questionnaire Items for Test Preparation**

**Table 6-16 Frequency with which the SPE and GE groups reported doing exam-related activities**

| | A good idea/Never | | Rarely-Sometimes | | Often-Always | | Mean | | Significance of the means |
|---|---|---|---|---|---|---|---|---|---|
| | SPE | GE | SPE | GE | SPE | GE | SPE | GE | |
| **14** | 19.9 | 25.6 | 22.4 | 28.0 | 56.1 | 45.0 | 3.44 | 3.03 | .000 |
| **22** | 6.2 | 10.6 | 29.9 | 38.4 | 61.2 | 49.5 | 3.72 | 3.29 | .000 |

Table 6.16 shows that significantly more SPE students took mock exams and practiced with sample test questions than GE students. However, it is not clear from the results why the students reported doing these activities. It might be because English was more important to the SPE students than to the GE students i.e. the difference might be related to each group's purpose of test preparation. Therefore, it would be interesting to probe this issue.

### 6.1.2.9. Reviewing

The following item enquired how often students went over the materials they had learned before.

*20. I review and practice what I learned*
**Figure 6-16 Questionnaire Item for Reviewing Learned Materials**

**Table 6-17 Frequency with which SPE and GE groups did reviewing and practicing**

| | A good idea/Never | | Rarely-Sometimes | | Often-Always | | Mean | | Significance of the means |
|---|---|---|---|---|---|---|---|---|---|
| | SPE | GE | SPE | GE | SPE | GE | SPE | GE | |
| **20** | 2.7 | 8.7 | 23.2 | 32.8 | 72.0 | 55.3 | 3.93 | 3.50 | .000 |

Both groups of students went over and practiced what they had learned before. However, SPE students reported doing these activities more frequently than GE students. It would be interesting to see what reasons the students give for reviewing and practicing.

### 6.1.2.10. Summary of the results for unpredicted reported activities

In addition to predicted activities, I also examined unpredicted students' activities, in answer to Research Question 1. A summary of these results are presented in Table 6.18.

**Table 6-18 Summary of the results for unpredicted reported activities**

| Unpredicted activities | SPE | GE | SPE | GE |
|---|---|---|---|---|
| **Specific vocabulary activities** | | | | |
| Memorising words | 3.98 | 3.85 | No difference | |
| Making sentences with new words | 3.19 | 2.29 | Longer | Shorter |
| Using English to English dictionary | 3.65 | 2.43 | Longer | Shorter |
| **Specific grammar activities** | | | | |
| Making sentences to practise grammar | 2.86 | 2.19 | Longer | Shorter |
| Memorising grammar rules | 3.34 | 3.36 | No difference | |
| Using texts for grammar | 3.98 | 3.43 | Longer | Shorter |
| **Specific reading activities** | | | | |
| Reading newspapers & magazines | 2.04 | 1.43 | Longer | Shorter |
| Translating while reading | 3.77 | 3.10 | Longer | Shorter |
| Guessing meaning of new words | 3.97 | 3.23 | Longer | Shorter |
| Reading story books | 3.18 | 1.87 | Longer | Shorter |
| Paying attention to topic of texts | 4.37 | 3.86 | Longer | Shorter |
| Skimming for general meaning | 4.21 | 3.55 | Longer | Shorter |
| Reading various texts | 3.19 | 2.07 | Longer | Shorter |
| Making summaries of reading texts | 2.95 | 2.35 | Longer | Shorter |
| **Diary writing** | | | | |
| Writing diaries, notes, etc | 2.37 | 1.60 | Longer | Shorter |
| **Using the Internet** | | | | |
| Using the Internet in English | 2.60 | 2.22 | Longer | Shorter |
| Sending and receiving emails in English | 2.48 | 2.06 | Longer | Shorter |
| **Listening and speaking** | | | | |
| Listening to the radio | 1.98 | 1.48 | Longer | Shorter |
| Looking for people to speak English with | 3.49 | 2.16 | Longer | Shorter |
| **Audio-visual activities** | | | | |
| Using tapes and CDs | 2.74 | 1.64 | Longer | Shorter |
| Watching English films & TV programmes | 3.12 | 2.53 | Longer | Shorter |
| Using computer games | 1.87 | 2.02 | No difference | |
| **Errors** | | | | |
| Using English despite errors | 3.69 | 2.71 | Longer | Shorter |
| **Using sample tests** | | | | |
| Taking mock exams | 3.44 | 3.03 | Longer | Shorter |
| Using practice tests | 3.72 | 3.29 | Longer | Shorter |

| Reviewing and practice | SPE | GE | SPE | GE |
|---|---|---|---|---|
| Reviewing and practicing | 3.93 | 3.50 | Longer | Shorter |

The results show that in general SPE students reported spending longer on all the activities except memorisation and the use of computer games, where the two groups were not statistically different. This could be because of the increase in the weight of English or because the SPE students had better English background. However, although the SPE students generally reported spending longer on the activities, the frequency with which they reported doing some of these activities was not very high. Some of the activities which had to do with the use of English, generally considered 'communicative' in the current thinking about language learning, including 'reading newspapers and magazines', 'writing diaries', 'using the Internet and emails in English', 'listening to the radio', 'using tapes and CDs', and 'using computer games' were done infrequently. However, there were other use-oriented activities which were done more frequently such as 'reading story books', 'reading various texts', 'looking for people to speak with', and 'watching films and TV programmes'. Though it is possible to attribute these activities to certain factors (see 8.2), it is not clear from the results why the students reported doing some of these activities with higher and others with lower frequencies. For example, the low frequency of 'listening to the radio' could be because oral skills are not tested in the exam or because radio is not the students' favourite media compared to other media such as television. In fact, 'watching films and TV programmes' was done more frequently, which could be because of the popularity of TV as well as accessibility, interest, etc. There were other activities which were reported being done frequently such as 'using English-to-English dictionary,' 'using texts for grammar practice', 'guessing', 'paying attention to topic', and 'skimming'. The higher frequency of these

activities might be because these activities were recommended to students in their textbooks as learning tips. However, although 'summarising reading texts' was a learning tip, it was done less frequently, which could be because this activity required a higher language ability than the students had. Conversely, translation was done frequently, although it was not encouraged by the students' textbooks. This activity, which is usually considered traditional, might be because of the students' learning beliefs, assuming that the belief was held by the students in advance. Another traditional activity, memorisation (irrespective of the lack of difference between the two groups), was also done frequently by both groups, which again could be because of the students' beliefs. However, memorisation of words was more frequent than memorisation of grammar rules, which might be because of the test as the number of vocabulary items was higher than grammar items. Similarly, 'making sentences for new words' was more frequent than 'making sentences for grammar rules', which again could be because of the differences in the number of vocabulary and grammar items. 'Reviewing and practising' could be because of the higher weight of English or because of the students' belief in the value of practsing in language learning. Finally, the SPE students 'used English despite the possiblity of errors' more frequently than the GE students, which could be because of their belief in the value of communication rather than too much concern about accuracy.

In sum, from the results discussed above, it is only possible to speculate as to the possible (reported) causes of these activities. Some possible causes might include the test, interest, access, beliefs, the textbook, the teacher, English background, etc. I examine possible causes further in the next section.

## 6.2. RQ2: Why do the students report doing these activities in order to prepare for the SPE Test?

In order to account for the students' reported activities, I will examine the students' knowledge of the test, purposes for test preparation, motivation, interest, and anxiety in the following sections.

### 6.2.1. Knowledge about the test

In order to find out whether there could be any relationship between what the students reported they did and the demands of the test, I asked the students how much they knew about the tests. The results are presented in Table 6.19.

Table 6.19 shows that the two groups had a good knowledge of the test with the SPE Group obtaining a score of .75 and GE .77 out of 1. In other words, 71.02% of SPE students and 74.50% of GE students gave correct answers to the test knowledge questions. The two groups were similar in their scores on each item except for Item 5 where about half of the SPE students did not know the exact number of questions in the SPE Test. The high mean scores give us reasonable confidence in attributing the students' test preparation activities to the test.

**Table 6-19 SPE and GE students' test knowledge**

| Items | % of students giving correct answers | | Mean score per item | | Significance of the means |
|---|---|---|---|---|---|
| | SPE | GE | SPE | GE | |
| 1. A section in the SPE/GE Test includes listening and speaking. | 76.8% | 75.0% | .80 | .77 | .326 |
| 2. Each section of the SPE/GE Test has the same number of questions. | 68.5% | 69.1% | .72 | .71 | .191 |
| 3. SPE/GE Test has six sections. | 75.5% | 77.5% | .80 | .80 | .029* |
| 4. Vocabulary and Reading have the same weight. | 74.1% | 73.3% | .78 | .76 | .099 |
| 5. The total number of questions in the SPE/GE Test is 70. | 53.9% | 75.0% | .565 0 | .78 | .000* |
| 6. The majority of the SPE/GE questions come from the Pre-university book. | 67.7% | 72.7% | .72 | .75 | .002* |
| 7. Marks allocated to English are more than non-English subjects. | 79.5% | 80.2% | .84 | .83 | .000* |
| 8. Each question has the same mark as the other, regardless of which section it belongs to. | 72.2% | 69.6% | .75 | .72 | .093 |
| **Groups' Test knowledge** | **71.02%** | **74.50%** | **.75** | **.77** | |

**Note:** maximum score= 1, minimum score= 0

### 6.2.2. Purposes for test preparation

The following item asked about the students' purposes or reasons for test preparation.

*8. Please number from 1 to 5 the most important reasons why you are preparing for the SPE/GE test where 1= the most important and 5= the least important.*

*I am preparing for the SPE Test because……*

| | |
|---|---|
| *a) I will use English resources at university* | |
| *b) I will have more and better opportunities for my job in the future* | |
| *c) I want to pass the SPE Test* | |
| *d) I will be able to communicate in English* | |
| *e) I want to travel abroad* | |
| *f) I will be able to appreciate cultural products such as films, art and literature* | |
| *g) I'm interested in English* | |
| *h) I want to live in an English-speaking country* | |
| *i) I will use English in many ways in future* | |

**Figure 6-17 Questionnaire Item for Reasons for Test Preparation**

Table 6.20 shows the results of the students' ranking their reasons for test preparation.

**Table 6-20 SPE and GE groups' test preparation purposes**

| Purposes/Reasons | Mean | | Sig | | Resultant ranking | |
|---|---|---|---|---|---|---|
| | SPE | GE | | | SPE | GE |
| **English sources** | 3.18 | 3.05 | .385 | | **Interest** | **Passing the exam** |
| **Job** | 2.87 | 2.41 | .000 | | **Passing the exam** | **Job** |
| **Passing the exam** | 2.53 | 2.14 | .001 | | **Job** | **Interest** |
| **Communication** | 2.98 | 3.00 | .847 | | **Communication** | **Communication** |
| **Travelling** | 3.63 | 3.25 | .010 | | **English sources** | **English sources** |
| **Cultural products** | 3.52 | 3.43 | .467 | | **Future use** | **Future use** |
| **Interest** | 1.75 | 2.96 | .000 | | **Cultural products** | **Travelling** |
| **Living in an English-Speaking country** | 3.62 | 3.30 | .109 | | **Living in an English-Speaking country** | **Living in an English-Speaking country** |
| **Future use** | 3.32 | 3.20 | .295 | | **Travelling** | **Cultural products** |

Of the nine items which were ranked, the students were significantly different on four items i.e. 'interest', 'passing the exam', 'job', and 'travelling'. While for the SPE students, 'interest' was the most important reason, for the GE Group 'passing the exam' was the most important. 'Passing the exam' was secondary to the SPE Group,

while it was 'Job' that was secondary to the GE. While 'Job' was of third importance to the SPE, 'interest' had the third importance to the GE. The next significant item, 'travelling', was ranked by the SPE Group as the least important, and of seventh importance by the GE Group. In sum, the results showed that 'interest', 'passing the exam', and 'job' were the first three most important purposes of test preparation for the students. Whether SPE students report doing their activities because of their interest or a future job or because they want to pass the exam, all of these reasons act as motives for the students to make efforts. However, it would be interesting to compare these results with data from other instruments and probe them further to see why the students have these purposes and how they are affecting their activities.

### 6.2.3. Motivation and interest

In this section, first I will present the items which were used to elicit data on the students' motivation and interest. Then I will present the results in Table 6.21 and 6.22.

*1. Success in SPE/GE Exam is important for me.*
*2. SPE/GE Exam makes me try.*
*3. I enjoy learning for SPE/GE Exam.*
*6. I don't like to study for SPE/GE Exam.*
*7. What do you think is your motivation level for learning for SPE/GE Exam? Please tick the appropriate box. (a- Highly motivated, b- Well-motivated, c- Motivated, d- Slightly motivated, e- Not at all motivated)*

**Figure 6-18 Questionnaire Items for Motivation and Interest**

**Table 6-21 SPE and GE groups' exam motivations**

| | Disagree | | No opinion | | Agree | | Mean | | Significance of the means |
|---|---|---|---|---|---|---|---|---|---|
| | SPE | GE | SPE | GE | SPE | GE | SPE | GE | |
| 1 | 0.8 | 5.7 | 2.4 | 10.3 | 96.2 | 83.1 | 4.65 | 4.20 | .000 |
| 2 | 2.7 | 10.5 | 9.4 | 19.5 | 87.1 | 69.1 | 4.25 | 3.79 | .000 |
| 3 | 3.2 | 19.5 | 13.5 | 34.0 | 81.7 | 44.7 | 4.19 | 3.33 | .000 |
| 6 | 87.6 | 66.0 | 7.5 | 18.0 | 4.3 | 15.0 | 1.63 | 2.24 | .000 |

**Table 6-22 SPE and GE groups' evaluation of their exam motivations**

| | Not/Slightly motivated | | Motivated | | Well/Highly motivated | | Mean | Significance of the means | |
|---|---|---|---|---|---|---|---|---|---|
| | SPE | GE | SPE | GE | SPE | GE | SPE | GE | |
| 7 | 4.6 | 29.1 | 28.0 | 41.4 | 66.8 | 29.1 | 3.95 | 3.00 | .000 |

The results in 6.21 show that SPE students were significantly more motivated to learn English for their exams than GE students. There were more SPE students (96.2%) for whom success was important, more SPE students (87.1%) who tried for the exam, more SPE students (81.7%) who enjoyed preparing for the exam, and more SPE students (87.6%) who disagreed with the statement that they did not like to study for the exam.

Item 7 asked the students to self-evaluate their motivation. Again, the SPE students were more motivated than the GE Group, with a significant difference. 66.8% of the SPE students were well/highly motivated, while only 29.1% of GE students were in this category.

Items 3 and 6 confirm the results presented in Table 6.20 where the SPE students' interest was higher than the GE students'. However, the results of Item 1 (which showed that success in the SPE test was more important than in the GE test) contradict the results in Table 6.20 which showed that interest was the more important reason for test preparation than passing the exam. These inconsistencies suggest a need to look at the results from the perspectives of other instruments. Also it is not clear from the results why SPE students reported being more motivated than GE students or how the test is affecting the students' motivation.

### 6.2.4. Test anxiety

The following items examined the students' anxiety during test preparation.

*4. I encounter a lot of stress during test preparation.*
*5. I'm afraid of getting a bad mark on SPE/GE Exam.*
**Figure 6-19 Questionnaire Items for Anxiety**

**Table 6-23 The SPE and GE groups' beliefs about their exam anxiety**

| | Disagree | | No opinion | | Agree | | Mean | | Significance of the means |
|---|---|---|---|---|---|---|---|---|---|
| | SPE | GE | SPE | GE | SPE | GE | SPE | GE | |
| 4 | 28.6 | 30.3 | 15.6 | 19.8 | 55.0 | 48.4 | 3.35 | 3.21 | .080 |
| 5 | 27.2 | 21.7 | 16.2 | 25.2 | 55.5 | 51.6 | 3.41 | 3.35 | .453 |

The results in Table 6.23 show that 55% of the SPE students and 48.4% of the GE students reported encountering a lot of stress during test preparation. Almost the same number of students in each group agreed that they were afraid of getting a bad mark on their exams. However, the p-values show that the two groups were not significantly different in their test preparation anxiety.

Despite the non-significant difference between the two groups, it would still be interesting to examine further the students' anxiety through other instruments as the results might give us insights into test washback in general.

### 6.2.5. Summary of the results for possible (reported) causes of the students' activities

Table 6.24 shows briefly whether students' knowledge of the test, their purposes for test preparation, motivation, interest, and anxiety could possibly account for their reported activities.

**Table 6-24 Summary of the results-possible (reported) causes**

| **Possible causes** | **Questionnaire** | |
|---|---|---|
| **Knowledge of the test** | **SPE** | **GE** |
| 1. A section in the SPE/GE Test includes listening and speaking. | No difference | |
| 2. Each section of the SPE/GE Test has the same number of questions. | No difference | |
| 3. SPE/GE Test has six sections. | *Lower* | *Higher* |
| 4. Vocabulary and Reading have the same weight. | No difference | |
| 5. The total number of questions in the SPE/GE Test is 70. | *Lower* | *Higher* |
| 6. The majority of the SPE/GE questions come from the Pre-university book. | *Lower* | *Higher* |
| 7. Marks allocated to English are more than non-English subjects. | Higher | Lower |
| 8. Each question has the same mark as the other, regardless of which section it belongs to. | No difference | |
| **Total score on knowledge of the test** | **.75 of 1** | **.77 of 1** |
| **Purposes/reasons** | | |
| English sources | No difference | |
| Job | *Lower* | *Higher* |
| Passing the exam | *Lower* | *Higher* |
| Communication | No difference | |
| Travelling | *Lower* | *Higher* |
| Cultural products | No difference | |
| Interest | Higher | Lower |
| Living in an English-Speaking country | No difference | |
| Future use | No difference | |

| Motivation and interest | SPE | GE |
|---|---|---|
| Importance of exam success | Higher | Lower |
| Trying/making effort | Higher | Lower |
| Interest | Higher | Lower |
| Lack of interest | *Lower* | *Higher* |
| Level of motivation | Higher | Lower |
| **Anxiety** | | |
| Stress | No difference | |
| Afraid of bad marks | No difference | |

The students' total score on test knowledge shows that they were aware of the main requirements of the test including weight of each section, weight of English, and the broader curriculum. Since SPE students were statistically different in these areas of activities, I had a basis on which to speculate that these activities could have been due to the test.

Of the 9 reasons for test preparation, the first three reasons, although slightly different from one group to another, were the most important which included interest, passing the exam and future job, all three of which could have motivated the students to try hard (Table 6.20). There was an inconsistency in the results for the importance of 'passing the exam' which suggested looking at the results using different instruments.

Regarding motivation and interest, the results showed that SPE students were motivated and more interested on all the items than GE students. However, the two groups were not different in anxiety. In the end, I suggested that the results could be probed further to explore the relationship between the students' reported activities and their test knowledge and how the test was interacting with the students' purposes, motivation, interest, and anxiety or how it was affecting them.

## 6.3. RQ3: Do SPE students perceive the SPE Test to be more difficult than the GE Test?

In this section as well as in 6.4, I will discuss two more areas of washback: test

difficulty and English ability. The following item asked the students how difficult the SPE/GE test was. Based on the claim made by NOET about the increase in the difficulty of the SPE Test, I predicted SPE students' rating of the difficulty would be higher than GE students'.

*9. Please tick the appropriate box.*
*I think the SPE/GE Test is.....*

| *a) very difficult* | *b) difficult* | *c) of medium difficulty* | *d) easy* | *e) very easy* |
|---|---|---|---|---|
| | | | | |

**Figure 6-20 Questionnaire Item for Test Difficulty**

**Table 6-25 The SPE and GE groups' views of their exam difficulty**

| | Easy | | Medium | | Difficult | | Mean | | Significance of the means |
|---|---|---|---|---|---|---|---|---|---|
| | SPE | GE | SPE | GE | SPE | GE | SPE | GE | |
| 14 | 3.0 | 10.6 | 26.1 | 49.3 | 69.5 | 38.7 | 3.95 | 3.34 | .000 |

Table 6.22 shows that the results confirm my expectation as well as NOET's claim. The results show that while the majority of SPE students (69.5%) perceived their test as 'difficult', the majority of GE students (49.3%) perceived their test of 'medium difficulty'. The mean difficulty of the SPE Group is also higher than the mean of the GE Group.

That SPE students considered the SPE Test as more difficult might be because this test was not solely based on the school textbook. However, it is not clear how the test difficulty is affecting the students' test preparation activities, i.e. whether it makes the students spend more time or less time on studying or whether it encourages the students to spend more time on certain sections but less time on others.

## 6.4. RQ4: Do SPE students have better English backgrounds than GE students?

As mentioned in the test specifications (Appendix 3), one of the purposes of the SPE Test was to admit more proficient applicants. To answer Research Question 4, therefore, I asked the students two questions about their English background, one about the length of their attendance in English institutes and one about their self-

assessed proficiency. My expectation was that SPE students would have better English backgrounds than GE students.

### 6.4.1. Attendance in English language institutes

The following item asked how long the students attended English language institutes.

*4. Have you attended English language institutes so far? Please tick the appropriate box.*

| *Yes* | *No* |
|---|---|
| | |

*If your answer was 'Yes', please tick one of the following boxes to indicate how many terms you have attended English language institutes.*

| *a) 1-5* | *b) 6-10* | *c) 11-15* | *d) 16-20* | *e) 21-25* | *f) 26-30* | *g) Over 30* |
|---|---|---|---|---|---|---|
| | | | | | | |

**Figure 6-21 Questionnaire Item for Period of Attending Language Institutes**

**Table 6-26 Percentage of the SPE and GE groups attending English institutes different lengths of time by terms**

| | 0 | 1-5 | 6-10 | 11-15 | 16-20 | 21-25 | 26-30 | Over 30 | Mean | Sig. |
|---|---|---|---|---|---|---|---|---|---|---|
| **SPE** | 18.9 | 11.6 | 15.1 | 18.9 | 18.9 | 11.1 | 3.5 | 1.9 | 13.20 | .000 |
| **GE** | 58.3 | 14.7 | 10.5 | 8.2 | 4.3 | 0.7 | 1.0 | 0.7 | 4.71 | |

The results fulfilled my expectation. The SPE Group's average number of terms of attendance in English institutes was 13.20, while the GE Group's was 4.71, and here the two groups were significantly different. The results suggest that the SPE Exam was successful in attracting more proficient applicants than GE applicants. However, it would be interesting to probe the issue further, i.e. whether the students continued attending the institutes or stopped attending for the purpose of test preparation, and why.

### 6.4.2. Self-assessed English proficiency

In addition to collecting information about the students' attendance in English institutes, I asked the students to assess their own proficiency in English. The results are presented below.

*6. How do you evaluate your English ability in general?*

| *Weak* | *Below average* | *Average* | *Above average* | *Strong* |
|---|---|---|---|---|
| | | | | |

**Figure 6-22 Questionnaire Item for Self-Evaluation**

**Table 6-27 The SPE and GE groups' assessment of their own language ability**

| | Below average/weak | | Average | | Above average/strong | | Mean | | Significance of the means |
|---|---|---|---|---|---|---|---|---|---|
| | SPE | GE | SPE | GE | SPE | GE | SPE | GE | |
| 4 | 4.9 | 22.9 | 41.2 | 55.8 | 53.4 | 20.8 | 3.97 | 2.95 | .000 |

As I expected, the results showed that SPE students assessed their proficiency higher (3.97 on average) than GE students (2.95 on average). This confirms the results presented in Table 6. 25 as well as suggests that NOET achieved one of the intended purposes of the SPE Test, which was admitting students with a stronger English background.

## 6.5. RQ5: What beliefs do the SPE students report holding about learning English?

As mentioned earlier, SPE students partly rely on out-of-school resources for their test preparation, which probably requires them to make some of the decisions on their own. Therefore, what beliefs they have about language learning may play an important role in the type of washback the test is likely to produce. In this section, I will report their learning beliefs and in the next section (6.6) I will discuss how they interact with the test. I will present the results using Horwitz's (1987; 1999) categories.

### 6.5.1. Foreign Language Aptitude

The following item asked about students' own aptitude in learning foreign languages.

*6. I have a special ability for learning foreign languages*

**Figure 6-23 Questionnaire Item for Aptitude in Learning Foreign Languages**

**Table 6-28 SPE and GE groups' reported beliefs about their own language learning aptitude**

| | Disagree | | No opinion | | Agree | | Mean | | Significance of the means |
|---|---|---|---|---|---|---|---|---|---|
| | SPE | GE | SPE | GE | SPE | GE | SPE | GE | |
| 6 | 2.7 | 25.5 | 33.4 | 40.8 | 57.1 | 31.9 | 3.78 | 3.05 | .000 |

The results in 6.28 show that more SPE students (57.1%) believed that they had a special ability for learning foreign languages than GE students (31.9). The two groups were significantly different on this item. However, the question which cannot be answered by the results presented here is whether the students believe they have that special ability for test preparation as\well or believe that they are less able, and why.

### 6.5.2. Difficulty of language learning

The two items to be presented in this section concern the difficulty of learning English. According to Horwitz (1987), how students judge the difficulty of a language influences their expectations and commitment to language learning. If students look at a language as easy, they are likely to become frustrated when they do not see improvements as they expect. On the other hand, if they perceive the language as too difficult or as, for instance, taking ten years or so to learn, it could be discouraging and make them put in minimal efforts (p 123).

*1. I believe that I will learn to speak English very well.*

**Figure 6-24 Questionnaire Section 2 Item 1 for Belief about Difficulty of language Learning**

**Table 6-29 SPE and GE groups' reported beliefs about difficulty of language learning**

| | Disagree | | No opinion | | Agree | | Mean | | Significance of the means |
|---|---|---|---|---|---|---|---|---|---|
| | SPE | GE | SPE | GE | SPE | GE | SPE | GE | |
| 1 | 0.3 | 15.0 | 7.8 | 29.5 | 85.7 | 54.6 | 4.34 | 3.49 | .000 |

Item 2 examined the students' expectations of their success. As Table 6.29 shows, SPE Group had a higher expectation of success than the GE Group. They were significantly different in their expectations.

*22. The English language is a) very difficult b) difficult c) of medium difficulty d) easy e) very easy*
**Figure 6-25 Questionnaire Section 2 Item 22 for Belief about Difficulty of Language Learning**

**Table 6-30 SPE and GE groups' reported beliefs about difficulty of language learning**

| | Easy | | Medium | | Difficult | | Mean | | Significance of the means |
|---|---|---|---|---|---|---|---|---|---|
| | SPE | GE | SPE | GE | SPE | GE | SPE | GE | |
| 22 | 30.5 | 17.8 | 45.3 | 49.9 | 17.3 | 31.5 | 2.83 | 3.18 | .000 |

Based on the results in Table 6.30, almost half of the students from each group believed that English was of 'medium' difficulty and the other half considered English as either 'easy' or 'difficult'. The question concerning test preparation is whether the students consider the SPE Test of the same level of difficulty, or of a different level of difficulty and, if different, why.

### 6.5.3. Nature of language learning

The following items asked about the nature of language learning including the role of the knowledge of culture in language learning (Item 3), focus of language learning (items 7, 10, 14), difference between learning a language and learning other subjects (Item 13), and the role of memorisation (Item 18).

*3. It is necessary to know about English-speaking cultures in order to speak English.*
*7. The most important part of learning a foreign language is learning vocabulary words.*
*10. The most important part of learning a foreign language is learning the grammar.*
*13. Learning a foreign language is different than learning other academic subjects.*
*14. The most important part of learning English is learning how to translate into Farsi.*
*18. Language learning involves a lot of memorisation.*
**Figure 6-26 Questionnaire Items for Nature of Language Learning**

**Table 6-31 SPE and GE groups' reported beliefs about the nature of language learning**

| | Disagree | | No opinion | | Agree | | Mean | | Significance of the means |
|---|---|---|---|---|---|---|---|---|---|
| | SPE | GE | SPE | GE | SPE | GE | SPE | GE | |
| **3** | 8.6 | 15.9 | 24.5 | 26.5 | 59.3 | 56.7 | 3.77 | 3.60 | .009 |
| **7** | 11.1 | 6.9 | 12.1 | 6.9 | 70.4 | 85.3 | 3.94 | 4.25 | .000 |
| **10** | 37.2 | 46.8 | 24.0 | 24.3 | 31.3 | 27.6 | 2.94 | 2.75 | .008 |
| **13** | 6.5 | 9.3 | 19.9 | 19.9 | 65.8 | 69.3 | 3.95 | 3.85 | .134 |
| **14** | 32.3 | 11.5 | 23.2 | 18.6 | 37.7 | 68.1 | 3.06 | 3.78 | .000 |
| **18** | 19.4 | 19.8 | 19.9 | 19.2 | 53.4 | 59.8 | 3.49 | 3.55 | .375 |

The results show that SPE and GE groups were significantly different on 4 out of 6 items. The average of Item 3 shows that the majority from both groups believed in the value of culture in language learning, but significantly more of the SPE Group

believed in culture than the GE Group. The question which arises concerning test preparation is whether the students believed that knowledge of culture was necessary for the SPE/GE Test as well and why.

Concerning Items 7, 10, and 14, which deal with the focus of language learning, while the majority from both groups agreed that vocabulary should be the focus of language learning, a minority from the two groups agreed with grammar learning. However, the difference between them was that fewer SPE students agreed about vocabulary but more of them agreed about grammar learning. In terms of translation, while the majority of GE students agreed with translation, only a minority from the SPE Group endorsed this belief. However, it is not clear how the test is interacting with (affecting or being affected) the beliefs of the students about what the focus of language learning should be, i.e. whether and to what extent the students would or would not do vocabulary, grammar, and translation for test preparation.

As regards beliefs about memorisation and difference between learning a language and learning other subjects, there was no significant difference between the two groups. However, the majority from both groups agreed that learning a language was different from learning other academic subjects and that a lot of language learning was memorisation. Again, it is not clear from the results whether and to what extent the students considered test preparation different from studying other subjects and whether they used memorisation and why. Interaction of the test with beliefs might be particularly significant in the case of beliefs such as memorisation and translation which are usually considered traditional (Sakui and Gaies, 1999), i.e. whether the test is encouraging beliefs which are traditional or beliefs which are based on current thinking about language learning.

### 6.5.4. Learning and Communication Strategies

The following items asked about learning and communication strategies which, according to Horwitz (1987), concern students' actual language learning practices.

*2. It is important to speak English with an excellent pronunciation.*
*4. I enjoy practicing English with the English speakers I meet.*
*5. It's o.k. to guess if you don't know a word in English.*
*9. I feel timid speaking English with other people.*
*12. It is important to practice with tapes and CDs.*
**Figure 6-27 Questionnaire Items for Strategies**

**Table 6-32 The SPE and GE groups' reported beliefs about the nature of learning and communication strategies**

| | Disagree | | No opinion | | Agree | | Mean | | Significance of the means |
|---|---|---|---|---|---|---|---|---|---|
| | SPE | GE | SPE | GE | SPE | GE | SPE | GE | |
| 2 | 4.6 | 10.0 | 7.5 | 16.9 | 80.3 | 72.3 | 4.39 | 3.94 | .000 |
| 4 | 0.8 | 6.4 | 5.4 | 16.2 | 87.3 | 76.6 | 4.56 | 4.08 | .000 |
| 5 | 6.5 | 17.1 | 11.9 | 18.1 | 75.2 | 64.2 | 4.12 | 3.66 | .000 |
| 9 | 64.4 | 55.9 | 11.6 | 24.0 | 16.4 | 17.8 | 2.18 | 2.40 | .005 |
| 12 | | 4.6 | 2.7 | 12.3 | 90.0 | 82.2 | 4.62 | 4.17 | .000 |

As the p-values in Table 6.32 show, SPE and GE Groups were significantly different on all the items in this section. The results for Item 2 show that the number of SPE students concerned about pronunciation was greater than GE students. Concern about pronunciation was examined because Horwitz (1987; 1999) states in this regard that overconcern with accent may inhibit communication. It is not clear, however, whether the test takers who do speaking are concerned about pronunciation for test preparation as well.

On Item 4, both groups agreed that they would enjoy practicing English with native speakers if they had the chance. However, more students from the SPE Group had this belief than from the GE Group.

On Item 5, the majority of the students from both groups agreed that guessing the meaning of unknown words was a good idea. However, the SPE students endorsed this belief significantly more than the GE Group.

As for Item 9, the students disagreed about their feeling timid when speaking English, with the SPE feeling less timid than the GE. This item (like Item 2) assumes speaking and therefore in a test preparation situation probably cannot be addressed to students who do not do speaking. In addition, belief about timidity as a personality characteristic may be less likely to be affected by a test.

The mean values of Item 12 show that the students believed in the value of practicing with tapes and CDs and therefore showed a high degree of agreement with the SPE Group endorsing the belief more than the GE Group.

However, the questions which arise with respect to the beliefs discussed above are whether in preparation for the SPE Test, speaking with an excellent pronunciation was important, whether the students enjoyed practicing with native speakers if they had the chance, whether they guessed the meaning of words or looked up every single word, whether the students did a lot of repetition and practice, and finally whether they felt timid speaking English to practice for test preparation. It would be interesting to follow up these questions with 'why'.

### 6.5.5. Motivations and expectations

The following items concerned the desires the students had about learning English. It would be interesting to examine them in connection with test preparation as they would give us insights into the aspects of motivation and expectations which could be important for motivating test takers to study for a test. The items and the results are presented below.

*8. People in my country feel that it is important to speak English.*
*11. I would like to learn English so that I can get to know English speakers better.*
*15. If I learn English very well, I will have better opportunities for a good job.*
*16. I want to learn to speak English well.*
*17. I would like to have English-speaking friends.*

**Figure 6-28 Questionnaire Items for Motivations and Expectations**

**Table 6-33 The SPE and GE groups' reported beliefs about language learning motivations and expectations**

| | Disagree | | No opinion | | Agree | | Mean | | Significance of the means |
|---|---|---|---|---|---|---|---|---|---|
| | SPE | GE | SPE | GE | SPE | GE | SPE | GE | |
| **8** | 4.9 | 10.3 | 23.2 | 22.9 | 63.1 | 65.4 | 3.89 | 3.80 | .152 |
| **11** | 27.8 | 32.7 | 35.8 | 30.7 | 28.8 | 34.2 | 3.07 | 3.02 | .465 |
| **15** | 2.4 | 5.2 | 7.5 | 10.6 | 83.6 | 82.9 | 4.42 | 4.31 | .040 |
| **16** | 0.5 | 4.2 | 1.1 | 7.0 | 91.9 | 87.4 | 4.75 | 4.31 | .000 |
| **17** | 2.7 | 13.8 | 14.3 | 30.7 | 76.5 | 52.8 | 4.30 | 3.60 | .000 |

The results show that out of 5 items, the two groups were significantly different on 3 items. On Item 8, the majority from both groups agreed that speaking English was important to the people of their country. On Item 11, only the minority from both groups agreed that their purpose of learning English was to know English speakers, with no significant difference between the groups. On the remaining 3 items, however, the two groups were significantly different with the SPE Group endorsing the beliefs more than the GE students. A greater number of SPE students associated learning English with better job opportunities, and they showed a greater desire to learn to speak well and to have English-speaking friends.

However, the questions which cannot be answered by the results presented in Table 6.33 include whether they (those who agreed to the beliefs just discussed) believe that people in their country also feel that it is important to pass the SPE Test, whether they believe that preparing for the SPE Test would make them know English-speakers better, whether preparing for the SPE Test would provide for better job opportunities, whether they would like to prepare for the SPE Test well, and finally whether having English-speaking friends would help them better prepare for the SPE Test.

### 6.5.6. Relative usefulness of oral versus written skills

The following item asked students if they believed oral skills or written skills were more useful in language learning.

*19. Speaking and listening to English are more useful than reading and writing English.*
**Figure 6-29 Questionnaire Item for Usefulness of Oral and Written Skills**

**Table 6-34 The SPE and GE groups' reported beliefs about relative usefulness of oral versus written skills**

| | Disagree | | No opinion | | Agree | | Mean | | Significance of the means |
|---|---|---|---|---|---|---|---|---|---|
| | SPE | GE | SPE | GE | SPE | GE | SPE | GE | |
| 19 | 11.1 | 16.0 | 20.2 | 21.9 | 62.3 | 61.5 | 3.88 | 3.73 | .028 |

The majority from both groups agreed that oral skills were more useful than written skills, but SPE students agreed significantly more than GE students. However, again it is not clear whether for test preparation also the students prefer oral skills to written skills.

### 6.5.7. Role of time

The following item asked students how long they believed language learning took.

*20. Language learning takes a long time.*
**Figure 6-30 Questionnaire Item for Role of Time**

**Table 6-35 The SPE and GE groups' reported beliefs about the length of time language learning takes**

| | Disagree | | No opinion | | Agree | | Mean | | Significance of the means |
|---|---|---|---|---|---|---|---|---|---|
| | SPE | GE | SPE | GE | SPE | GE | SPE | GE | |
| 20 | 9.7 | 13.5 | 21.3 | 20.2 | 63.1 | 65.5 | 3.75 | 3.74 | .874 |

The majority from both groups believed that language learning took a long time, with non-significant difference. The question to be examined for test preparation would be whether the students' concern would be short-term purposes and activities such as doing practice tests or long-term purposes of developing a good English background (Bailey, 1996: 269).

### 6.5.8. Use of language

The following item asked students if they believed using English was useful in learning English.

*21. It is important to find as many ways as possible to use English.*
**Figure 6-31 Questionnaire Item for Use of Language**

**Table 6-36 The SPE and GE groups' reported beliefs about using English**

| | Disagree | | No opinion | | Agree | | Mean | | Significance of the means |
|---|---|---|---|---|---|---|---|---|---|
| | SPE | GE | SPE | GE | SPE | GE | SPE | GE | |
| 21 | 0.3 | 1.9 | 4.0 | 6.4 | 88.9 | 90.3 | 4.63 | 4.44 | .000 |

The great majority of the students in both groups believed that in order to learn English it was important to use it in various ways. On average, significantly more SPE students agreed with this belief. It would be interesting to see whether the students would act consistently with this belief and would feel a need to use English for test preparation.

### 6.5.9. Summary

In this section, I examined the students' reported beliefs about foreign language learning in general based on the Horwitz's (1987; 1999) questionnaire (mainly) and two other questionnaires. Table 6.37 summarises what types of beliefs the students reported they held.

**Table 6-37 SPE and GE students' reported language learning beliefs**

| | SPE | GE | SPE | GE |
|---|---|---|---|---|
| **Foreign language aptitude** | | | | |
| 6. I have a special ability for learning foreign languages. | Higher | Lower | 3.78 | 3.05 |
| **Difficulty of language learning** | | | | |
| 1. I believe that I will learn to speak English very well. | Higher | Lower | 4.34 | 3.49 |
| 22. The English language is a) very difficult b) difficult c) of medium difficulty d) easy e) very easy | Lower | Higher | 2.83 | 3.18 |
| **Nature of language learning** | | | | |
| 3. It is necessary to know about English-speaking cultures in order to speak English. | Higher | Lower | 3.77 | 3.60 |
| 7. The most important part of learning a foreign language is learning vocabulary words. | Lower | Higher | 3.94 | 4.25 |
| 10. The most important part of learning a foreign language is learning the grammar. | Higher | Lower | 2.94 | 2.75 |
| 13. Learning a foreign language is different than learning other academic subjects. | No difference | | 3.95 | 3.85 |
| 14. The most important part of learning English is learning how to translate into Farsi. | Lower | Higher | 3.06 | 3.78 |
| 18. Language learning involves a lot of memorization. | No difference | | 3.49 | 3.55 |
| **Learning and Communication Strategies** | | | | |
| 2. It is important to speak English with an excellent pronunciation. | Higher | Lower | 4.39 | 3.94 |
| 4. I enjoy practicing English with the English speakers I meet. | Higher | Lower | 4.56 | 4.08 |
| 5. It's o.k. to guess if you don't know a word in English. | Higher | Lower | 4.12 | 3.66 |
| 9. I feel timid speaking English with other people. | Lower | Higher | 2.18 | 2.40 |
| 12. It is important to practice with tapes and CDs. | Higher | Lower | 4.62 | 4.17 |
| **Motivations and expectations** | | | | |
| 8. People in my country feel that it is important to speak English. | No difference | | 3.89 | 3.80 |
| 11. I would like to learn English so that I can get to know English speakers better. | No difference | | 3.07 | 3.02 |
| 15. If I learn English very well, I will have better opportunities for a good job. | Higher | Lower | 4.42 | 4.31 |
| 16. I want to learn to speak English well. | Higher | Lower | 4.75 | 4.31 |
| 17. I would like to have English-speaking friends. | Higher | Lower | 4.30 | 3.60 |
| **Relative usefulness of oral versus written skills** | | | | |
| 19. Speaking and listening to English are more useful than reading and writing English. | Higher | Lower | 3.88 | 3.73 |
| **Role of Time** | | | | |
| 20. Language learning takes a long time. | No difference | | 3.75 | 3.74 |
| **Use of language** | | | | |
| 21. It is important to find as many ways as possible to use English. | Higher | Lower | 4.63 | 4.44 |

Table 6.37 shows that in general SPE students endorsed the beliefs more than GE students. The SPE Group was more concerned about pronunciation than the GE Group, which might act as an obstacle to using language for communication. However, they were less timid than GE students, which may be an advantage for communication. SPE students considered English easier than GE students, and it

would be interesting to explore in terms of whether this belief could affect test preparation or could be reshaped by that. The majority of the SPE students believed in the importance of vocabulary learning, although less than the GE students. Conversely, only a minority of the SPE students considered grammar learning important, although more than the GE students. Therefore, if the students do a lot of vocabulary learning or if they do not emphasise grammar in their test preparation, it could be because of their beliefs in the importance of vocabulary and in the insignificance of grammar as well. Both groups believed in the value of translation, but the SPE Group endorsed this belief less than the GE Group. Both groups also believed in memorization but there was no statistical difference between them. The question which arises here is whether students' memorisation and translation beliefs would make them memorise or translate for their test preparation more or less than the extent to which they believe in them. In other words, the question is whether there is consistency between beliefs and activities. Another question is which types of beliefs are the test preparation activities consistent with? Are they generally considered traditional or contemporary? This is the subject of the next section.

## 6.6. RQ6: Are the SPE students' reported activities consistent with their reported beliefs?

To answer this research question, I will pair reported beliefs and reported activities based on the issues they deal with and then compare them for consistency. I will use Paired-Samples T-Test to compare the reported beliefs and reported activities of each group of SPE and GE. Then I will compare the (in)consistencies of the SPE Group with those of the GE Group. I will only examine the beliefs which have counterparts in 'activities', i.e. beliefs the contents of which match with those of 'activities'.

### 6.6.1. Reported belief about vocabulary and the activity

Through the following pair of items I compare the belief (abbreviated as 'Bel' in this section) about vocabulary learning and the activity of learning vocabulary (abbreviated as 'Act') for consistency.

*Bel: The most important part of learning a foreign language is learning vocabulary words.*
*Act: Please number from one to five the following language skills on which you spend most time where 1= most time and 5 = least time.*

| *Vocabulary* | *Grammar* | *Reading* | *Listening* | *Speaking* |
|---|---|---|---|---|
| | | | | |

**Figure 6-32 Questionnaire Item for Belief on Vocabulary and Activity**

The results are presented in the following table.

**Table 6-38 Consistency between belief and activity concerning vocabulary**

| SPE | | | GE | | |
|---|---|---|---|---|---|
| **Belief ($\bar{X}$)** | **Activity ($\bar{X}$)** | **Sig** | **Belief ($\bar{X}$)** | **Activity ($\bar{X}$)** | **Sig** |
| 3.91 | 3.76 | .136 | 4.23 | 3.56 | .000 |

Table 6.38 shows that for the SPE Group the difference between belief and activity was not significant (.136) while for the GE Group the difference was significant (.000). In other words, the belief and activity were consistent in the case of the SPE Group but inconsistent in the case of the GE Group. Consistencies suggest that the test did not affect beliefs and inconsistencies suggest that the test probably affected beliefs. Therefore, the results suggest that while the SPE students' belief about vocabulary was not affected by the test, the GE Groups' belief was affected. The results also show that the mean value for the GE Group's activity is lower than the mean of belief. This may suggest that the amount of activity encouraged by the GE Test was probably less than the importance the GE students attached to vocabulary for general language learning. However, it is not clear from the results whether the inconsistency is due to the test or other factors and why.

### 6.6.2. Reported belief about grammar and the activity

The following pair of items examines consistency between belief about

grammar learning and the activity of learning grammar.

*Bel: The most important part of learning a foreign language is learning the grammar.*
*Act: Please number from one to five the following language skills on which you spend most time where 1= most time and 5 = least time.*

| *Vocabulary* | *Grammar* | *Reading* | *Listening* | *Speaking* |
|---|---|---|---|---|
| | | | | |

**Figure 6-33 Questionnaire Items for the Belief and Activity of Grammar Learning**

**Table 6-39 Consistency between belief and activity concerning grammar**

| SPE | | | GE | | |
|---|---|---|---|---|---|
| **Belief ($\bar{X}$)** | **Activity($\bar{X}$)** | **Sig** | **Belief ($\bar{X}$)** | **Activity($\bar{X}$)** | **Sig** |
| 2.96 | 3.15 | .071 | 2.80 | 2.83 | .726 |

Table 6.39 shows that the p-values for both groups of SPE and GE are not significant. In other words, belief about grammar was consistent with the activity of grammar learning for both groups, which suggests that the students' belief was not affected by the test.

### 6.6.3. Reported belief about guessing and the activity

*Bel: It's o.k. to guess if you don't know a word in English.*
*Act: I try to guess the meaning of new words from context.*
**Figure 6-34 Questionnaire Items for the Belief and Activity of Guessing**

**Table 6-40 Consistency between belief and activity concerning guessing**

| SPE | | | GE | | |
|---|---|---|---|---|---|
| **Belief ($\bar{X}$)** | **Activity($\bar{X}$)** | **Sig** | **Belief ($\bar{X}$)** | **Activity($\bar{X}$)** | **Sig** |
| 4.13 | 3.91 | .002 | 3.66 | 3.17 | .000 |

The results in Table 6.40 show that for both groups belief about guessing is inconsistent with the activity of guessing for test preparation as the p-values show insignificance. This suggests that the students' belief was probably affected by the test. The results also show that both groups' mean values for activity are lower than the mean of the belief, which suggests that probably the amount of activity engaged in for test preparation was less than the degree to which the students believed guessing was important for language learning in general. However, whether the students used or did not use guessing because of the test or other factors requires further

investigation. The effect of the test might also have been mediated by the students' textbook, as in the Pre-university book guessing is encouraged as a reading skill (Appendix 3).

### 6.6.4. Reported belief about using tapes and CDs and the activity

*Bel: It is important to practice with tapes and CDs.*
*Act: I use tapes or CD's to practice English.*
**Figure 6-35 Questionnaire Items for the Belief and Activity of Using Tapes and CDs**

**Table 6-41 Consistency between belief and activity concerning the use of tapes and CDs**

| SPE | | | GE | | |
|---|---|---|---|---|---|
| **Belief ($\bar{X}$)** | **Activity ($\bar{X}$)** | **Sig** | **Belief ($\bar{X}$)** | **Activity ($\bar{X}$)** | **Sig** |
| 4.62 | 2.63 | .000 | 4.17 | 1.36 | .000 |

The results in Table 6.41 show that there was a significant difference between belief and activity for both groups, which suggests an effect from the tests. The results also indicate that while the mean values of activities are much lower than those of the belief for both groups, the mean value for the SPE Group is higher than that of the GE Group (2.63 versus 1.36). This might mean that both tests had probably negative washback in that they encouraged very little oral work but that the SPE Test had less negative washback in that it encouraged more oral work or more students to do oral work.

### 6.6.5. Reported belief about translation and the activity

*Bel: The most important part of learning English is learning how to translate into Farsi.*
*Act: I translate texts into Farsi while reading.*
**Figure 6-36 Questionnaire Items for the Belief and Activity of Translation**

**Table 6-42 Consistency between belief and activity concerning translation**

| SPE | | | GE | | |
|---|---|---|---|---|---|
| **Belief ($\bar{X}$)** | **Activity ($\bar{X}$)** | **Sig** | **Belief ($\bar{X}$)** | **Activity ($\bar{X}$)** | **Sig** |
| 3.05 | 3.72 | .000 | 3.77 | 3.04 | .000 |

Table 6.42 shows that the belief about and the activity of translation were

inconsistent, which suggests that both tests probably affected the translation belief. However, the effects of the two tests were different. While the SPE students translated for test preparation more than they believed translation was essential for language learning, the GE Group translated less than the degree to which they considered translation essential. As translation is not recommended as a skill in the students' textbook as well as in the language learning/teaching literature, the effect of the SPE Test might be interpreted as negative but the effect of the GE Test as positive.

### 6.6.6. Reported belief about memorisation and the activity

In this section, two activities will be compared with the belief about memorisation. One activity concerns memorisation of word meanings (in Pair 1) and the other concerns memorisation of grammatical rules (in Pair 2). The pairs of items and the results are presented below.

*Pair 1:*
*Bel: Language learning involves a lot of memorisation.*
*Act: I memorise word meanings.*
*Pair 2:*
*Bel: Language learning involves a lot of memorisation.*
*Act: I memorise grammatical rules.*
**Figure 6-37 Questionnaire Items for the Belief and Activities of Memorisation**

**Table 6-43 Consistency between belief and activity concerning memorisation**

| Pairs | SPE | | | GE | | |
|---|---|---|---|---|---|---|
| | Belief ($\bar{X}$) | Activity ($\bar{X}$) | Sig | Belief ($\bar{X}$) | Activity ($\bar{X}$) | Sig |
| Pair 1 | 3.48 | 3.95 | .000 | 3.56 | 3.81 | .000 |
| Pair 1 | 3.48 | 3.28 | .020 | 3.56 | 3.33 | .001 |

The results in Table 6.43 show that there were significant differences between the belief and activity for both groups, which suggest that the students' beliefs were probably affected by the tests. The mean values indicate that the two tests affected the students similarly. While the students memorised word meanings more than they believed in memorisation, they memorised grammar rules less than they thus believed. Since memorisation is not recommended in the Pre-university book and the language

learning literature, the fact that it has higher mean values in the case of vocabulary learning might be considered negative washback of the two tests, and the fact that it has lower mean values in the case of grammar rules might be considered positive washback of the two tests. The increase of memorisation in the case of vocabulary and its decrease in the case of grammar rules might be due to the fact that the students believed in vocabulary learning more than grammar learning (Tables 6.38 and 6.39). This suggests that beliefs are interrelated (see 3.7). I will discuss this issue further in 6.6.8.

### 6.6.7. Reported belief about using English and the activities

In this section I examine the consistency between belief about use of English and the activities in which the students report using English.

*Belief:*
*21. It is important to find as many ways as possible to use English.*
*Activities:*
*1. I read English newspapers and magazines.*
*4. I write things like diaries, notes, etc in English.*
*6. I watch English films or programmes.*
*8. I listen to English programs on the radio.*
*16. I read English story books.*
*21. I read various English texts as much as I can.*
*24. I send and receive emails in English.*

**Figure 6-38 Questionnaire Items for the Belief and Activity of Using English**

**Table 6-44 Consistency between reported belief and reported activity concerning use of English**

| Pairs | SPE | | | GE | | |
|---|---|---|---|---|---|---|
| | Belief ($\bar{X}$) | Activity($\bar{X}$) | Sig | Belief ($\bar{X}$) | Activity($\bar{X}$) | Sig |
| 21 vs 1 | 4.62 | 1.76 | .000 | 4.44 | 1.01 | .000 |
| 21 vs 4 | 4.62 | 2.19 | .000 | 4.44 | 1.36 | .000 |
| 21 vs 6 | 4.62 | 3.05 | .000 | 4.44 | 2.41 | .000 |
| 21 vs 8 | 4.62 | 1.81 | .000 | 4.44 | 1.26 | .000 |
| 21 vs 16 | 4.62 | 3.13 | .000 | 4.44 | 1.68 | .000 |
| 21 vs 21 | 4.62 | 3.11 | .000 | 4.44 | 1.86 | .000 |
| 21 vs 24 | 4.62 | 2.33 | .000 | 4.44 | 1.91 | .000 |

Table 6.43 shows that the reported beliefs and activities of both groups of students were significantly different on all the items, which suggest that the students were acting against their beliefs in their test preparation. Despite the fact that both

groups agreed very highly about the importance of using English, the frequency with which they reported doing the activities was low. However, SPE students were acting less inconsistently with their beliefs than GE students since they reported doing the activities more frequently. What is not clear, however, is whether the inconsistency between the belief and the activities is due to the test or other factors. For example, in connection with SPE Group's 'reading English story books (Item 16)', which was more frequent than other activities and therefore less inconsistent with their belief, factors such as interest might have also been important, or in the case of GE Group's 'watching English films and programmes (Item 6)', factors such as interest as well as access to facilities might have interacted with washback.

The following analyses were done after I conducted the interviews. The reason I returned to this analysis was that some students made a link between translation and 'paying attention to the topic of texts' and 'skimming for a general meaning of the passage' and between memorisation and using texts. In other words, some of them said instead of translation they paid attention to the topic of texts and skimmed for a general meaning of the passage, or instead of memorization they used examples for learning vocabulary and used tests for learning grammar. Therefore, I decided to examine through the questionnaire data whether beliefs in translation and memorisation were inconsistent with the alternative activities the students reported. In the following sections I will compare these beliefs with each of the alternative activities.

### 6.6.8. Reported beliefs in translation and memorisation versus alternative activities

*Bel: The most important part of learning English is learning how to translate into Farsi.*
*Act: I pay attention to the topic of the text when I read.*
*Act: I skim the text first to get a general idea and then go back to read it more carefully.*

*Bel: Language learning involves a lot of memorisation.*
*Act: I learn word meanings with an example in which the new word is used.*
*Act: I use texts to learn grammar.*

**Figure 6-39 Questionnaire Items for Translation and Memorisation vs. Alternative Activities**

**Table 6-45 Consistency between reported beliefs in translation and memorisation and alternative reported activities**

| | SPE | | | GE | | |
|---|---|---|---|---|---|---|
| **Belief and activities** | **Belief ($\bar{X}$)** | **Activity($\bar{X}$)** | **Sig** | **Belief ($\bar{X}$)** | **Activity($\bar{X}$)** | **Sig** |
| Translation & topic | 3.05 | 4.34 | .000 | 3.77 | 3.83 | .370 |
| Translation & skimming | 3.05 | 4.17 | .000 | 3.77 | 3.49 | .000 |
| Memorisation & examples | 3.48 | 2.46 | .000 | 3.56 | 2.07 | .000 |
| Memorisation & texts | 3.48 | 3.91 | .000 | 3.56 | 3.37 | .014 |

The results in Table 6.45 show that beliefs in translation and memorisation were inconsistent with, i.e. contradicted, the alternative activities except in the case of the relationship between translation and attention to topic of texts for the GE Group. In other words, the results indicate that those who believe less in translation may try to benefit more from the context in reading comprehension by paying attention to the topic of the passage and skimming or vice versa, and those who believe less in memorisation may be inclined to make more use of context in learning vocabulary or grammar. This suggests that the relationship among beliefs themselves and between beliefs and activities may be more complex than the preceding analyses suggested (see 3.7).

### 6.6.9. Summary of the results for consistency between reported beliefs and reported activities

Table 6.47 summarises the consistency between reported beliefs and reported activities and shows which beliefs were consistent and which beliefs were inconsistent.

**Table 6-46 Consistency between reported beliefs and reported activities**

| Beliefs versus activities | SPE | | GE | |
|---|---|---|---|---|
| | Belief | Activity | Belief | Activity |
| **Belief in vocabulary learning versus** activity of vocabulary learning | Consistent | | Higher | Lower |
| **Belief in grammar learning versus** activity of grammar learning | Consistent | | Consistent | |
| **Belief in the use of tapes & CDs versus** activity of using tapes &CDs | Higher | Lower | Higher | Lower |
| **Belief in guessing versus** activity of guessing | Higher | Lower | Higher | Lower |
| **Belief in translation versus** activity of translation | Lower | Higher | Higher | Lower |
| **Belief in memorisation versus** activity of memorising vocabulary | Lower | Higher | Lower | Higher |
| **Belief in memorisation versus** activity of memorising grammar rules | Higher | Lower | Higher | Lower |
| **Belief in using English versus** reading newspapers and magazines | Higher | Lower | Higher | Lower |
| **Belief in using English versus** writing diaries, notes, etc | Higher | Lower | Higher | Lower |
| **Belief in using English versus** watching films or TV programmes | Higher | Lower | Higher | Lower |
| **Belief in using English versus** listening to the radio | Higher | Lower | Higher | Lower |
| **Belief in using English versus** reading story books | Higher | Lower | Higher | Lower |
| **Belief in using English versus** reading various texts | Higher | Lower | Higher | Lower |
| **Belief in using English versus** sending and receiving emails | Higher | Lower | Higher | Lower |
| **Belief in translation versus** paying attention to the topic of texts | Lower | Higher | Consistent | |
| **Belief in translation versus** skimming for general meaning of texts | Lower | Higher | Higher | Lower |
| **Belief in memorisation versus** using examples for vocabulary learning | Higher | Lower | Higher | Lower |
| **Belief in memorisation versus** using texts for grammar learning | Lower | Higher | Higher | Lower |

**Note:** 'Higher' and 'Lower' compare the mean values of beliefs and activities of each group

The table shows that the two groups were inconsistent in all of the items except 'grammar' where both groups were consistent and two other items where one group was consistent. In the majority of the cases there were beliefs which were considered important, but their corresponding activities were reduced in frequency in test preparation. This was particularly true of the beliefs about 'using English', 'using tapes and CDs', and 'guessing', which probably indicates the negative washback of the test in that the test did not encourage use of English as well as use of context in the case of guessing. As regards translation, while the SPE Group's activity increases compared to the degree of importance attached to the belief about translation, the GE

Group's activity decreases. However, it is not clear from the questionnaire results whether the increase or decrease was because of the test since it could be due to too much or too little importance attached to the belief as well. Therefore, based on these results, it can only be claimed that the test might interact with beliefs, not necessarily affect beliefs or be affected by the beliefs. This suggests that to produce positive washback, not only test innovations but also what beliefs learners have are important. As far as memorisation is concerned, while the results were the same for the two groups, they were different for grammar and vocabulary, i.e. in the case of vocabulary the activities were higher and in the case of grammar, the activities were lower. This could be because of the test as the results earlier (Table 6.1) showed that the students from both groups spent more time on vocabulary than on grammar. The last four rows in the table showed a more complex relationship between beliefs and activities and suggested that certain beliefs might entail existence or non-existence of certain other beliefs and activities or entail certain degrees of their existence or non-existence.

In short, in this chapter I addressed my research questions through the questionnaire data. I examined the reported activities of the students, possible causes of these activities, test difficulty, students' English background, learning beliefs, and the relationship between beliefs and test preparation activities. However, the main question the questionnaire data was not able to answer was whether there was a (reported) cause-effect relationship between the test and the activities, between the test and beliefs, or between the beliefs and the activities. These results will be compared with the letters in Chapter 7 and the interviews in Chapter 8. In the interviews, I will also explore the type of relationship between the test, activities and beliefs.

# Chapter 7. Results from the Letters

In Chapter 6, I presented the results from the questionnaire data. Although the results enabled me to establish a tentative relationship between the students' activities and the test, they did not establish a (reported) cause-effect relationship or whether there could have been other causes for the students' reported activities. Therefore, in this chapter, I will look at the results from the lens which the letters provided. I will report them based on predicted versus unpredicted washback and in order of the research questions. I will report the frequency of occurrences of the students' recommendations in each group and the significance of the differences between the groups.

## 7.1. RQ1: What activities do the SPE students report doing in order to prepare for the SPE Test?

### 7.1.1. Predicted washback

#### 7.1.1.1. Test sections

Based on the number of items in the two tests (Appendix 3), I expected that more SPE students than GE students would recommend each of the test sections. The following table shows whether the results are in agreement with my expectation.

**Table 7-1 Percentage of students recommending the test sections**

| Test Sections | % GE | % SPE | P-value |
|---|---|---|---|
| **Vocabulary** | 86 | 81 | .581 |
| **Reading** | 39 | 45 | .670 |
| **Grammar** | 51 | 43 | .528 |
| **Cloze** | 2 | 12 | .090 |
| **Language Functions** | 0 | 0 | NA |
| **Sentence Structure** | 0 | 0 | NA |

Table 7.1 shows what percentage of the students from each group recommended which skill and whether the difference between the groups was statistically significant. Some of the statements included 'I suggest that you should learn vocabulary more than grammar' (SPE), 'Reading and cloze tests are among the

very important ones' (SPE), 'Learn vocabulary and reading very well' (GE).

The p-values show that there were no significant differences between the two groups in any of the common test sections. Therefore, my expectation that more SPE students would recommend the test sections was not fulfilled. The results also show that, unlike my expectation, no SPE students recommended the two new sections of Language Functions and Sentence Structure. This suggests that the two sections probably had no washback, which confirms the results of the questionnaire where there was no significant difference between the two groups. The non-significant differences in the new sections as well as the common sections might be because the sample sizes of each group were small (42 in SPE and 49 in GE), out of which a subset of the students made the recommendations. That only some of the students made the recommendations could be because the letters were written under time restrictions and therefore the students did not have enough time to write about everything they wanted to or it could be because they did not simply remember to write about them as the letter instruction was only a general one and did not ask them which aspects of test preparation to write about. It could also be because the students were not really interested in writing, although most of them volunteered to write the letters. Therefore, it is suggested that letter data should not be collected under tight time limitation and should be collected from students who are interested in writing in general.

#### 7.1.1.2. Oral activities

Since oral skills are not tested in the SPE and GE tests, I expected neither group to recommend practice of oral skills. However, a few students did recommend listening and speaking activities for example 'Speaking with an imaginary person' (SPE), 'Speaking English or listening to conversations can be very effective in

learning as both of them improve general English' (GE). The results are presented in Table 7.2.

**Table 7-2 Percentage of students recommending oral activities**

| Oral activities | % GE | % SPE | P-value |
|---|---|---|---|
| Listening | 6 | 5 | .960 |
| Speaking | 18 | 14 | .735 |

The results in Table 7.2 show that there was no significant difference between GE and SPE students in terms of recommending oral language activities. Therefore, my expectation was not fulfilled. The fact that only a few students from each group recommended oral activities might be because of the students' awareness of the test demands (Table 6.1 and 7.6). The results agree with those of the questionnaire reported in Table 6.8, partly agree with those of Table 6.10 and contradict those in Table 6.9.

#### 7.1.1.3. English versus Non-English subjects

Based on the change in the SPE Test where more weight was given to English against Non-English subjects, I expected that SPE students would not recommend studying Non-English courses as much as GE students would. The results are presented below.

**Table 7-3 Percentage of students recommending non-English subjects**

| | % GE | % SPE | P-value |
|---|---|---|---|
| Non-English subjects | 0 | 7 | .094 |

Although the percentage values show that 7 SPE students recommended non-English subjects (e.g. Non-English subjects affect our rankings very much') and no GE student recommended them, the p-value shows that the two groups were not significantly different. Therefore, my expectation about the weight of English versus non-English subjects was not fulfilled. However, the reason the GE students did not mention non-English subjects might be because of the letter writing instruction which

asked the students to write about ways of preparing for the GE (and SPE) Test and therefore they might have considered it irrelevant to talk about non-English subjects.

#### 7.1.1.4. Broader curriculum

The SPE Test is based on a broader curriculum than the GE Test. While the GE Test questions are based on the Pre-university book, the SPE Test is based on both the Pre-university book and extra materials, including *Bridging the Gap*. Therefore, I predicted that the SPE students would rely on the Pre-university book, *Bridging the Gap*, and other extra materials, while the GE students would only use the Pre-university book. Since SPE students were supposed to use materials in addition to the school materials, I expected further that they would recommend classes outside school such as preparation classes or English institutes. The results of the students' recommendations are presented below.

**Table 7-4 Percentage of students recommending materials and classes in and out of school**

| Materials and classes | | | % GE | % SPE | P-value |
|---|---|---|---|---|---|
| Materials | Out of school | Extra materials | 18 | 43 | .012 |
| | | *Bridging the Gap* | 0 | 7 | .094 |
| | In school | Pre-university book | 39 | 29 | .377 |
| Classes | Out of school | Preparation classes | 12 | 24 | .175 |
| | | English institutes | 2 | 0 | 1.000 |
| | In school | School classes | 8 | 10 | 1.000 |

The results in Table 7.4 show that, except for the use of extra materials, there were no significant differences between GE and SPE students in the recommendation of other resources. The fact that 43% of the SPE students recommended extra materials might be because of the test ('Use extra materials particularly *Bridging the Gap* (SPE), 'You should provide books in addition to the school book' (GE) (see also Table 7.6). However, since 18% of the GE students also recommended extra materials, it is worth probing further to see what they really meant by 'extra materials'. The reason why only a few SPE students recommended *Bridging the Gap* could be because the data was collected early in the students' test preparation period

when they might have barely started using the book. The reason for the insignificant differences between the groups in recommending the Pre-university book ('Firstly, you should emphasis the pre-university book') (GE) and school classes might be because they were important parts of test preparation for both groups. This might be the case with preparation classes as well ('Go to a prepration class as school English is very limited') (SPE). The fact that only 2 GE students and no SPE student recommended classes in English institutes might be because, as Bailey (1999: 14) states, the students might have skipped language classes to prepare for the tests.

### 7.1.2. Unpredicted washback

#### 7.1.2.1. Specific learning activities

In answer to Research Question 1, I presented the results related to 'predicted washback' in 7.1.1, and in this section I report the results related to 'unpredicted washback' as they were not necessarily predictable based on the features of the SPE Test. Another difference between them is that while the activities discussed under 'unpredicted washback' were concerned with language skills in general (vocabulary, grammar, reading, listening and speaking), the reported activities to be discussed here are specific activities related to those skills. Some of the SPE students' suggestions included 'You should learn vocabulary within senetences or by making sentences', 'Translate reading texts as far as you can', 'Also memorise and learn English words', 'Memorise many words', 'Using English story books for familiarity with sentence structure', and 'reading newspapers and magazines'. Some of the GE students' suggestions included 'Guessing the meaning of unknown words through the phrases in the text', 'Using flashcards with the words on one side and the meanings on the other', 'Use grammar rules and words in conversations to improve your English', 'Memorising words, them memorizing grammar, then translating reading texts',

'Leraning grammar through several examples', and 'Show interest in newspapers...'.

**Table 7-5 Percentage of students recommending specific skills activities**

| Skills | Specific activities | % GE | % SPE | P-value |
|---|---|---|---|---|
| **Vocabulary** | Learning out of context | 16 | 12 | .764 |
| | Memorising word meanings | 4 | 5 | .551 |
| | Using texts (sentences & passages) | 12 | 17 | .764 |
| | Using words in new contexts | 14 | 7 | .331 |
| **Grammar** | Learning forms/out of context | 6 | 2 | .620 |
| | Memorising grammar rules | 2 | 0 | .186 |
| | Using texts (sentences & passages) | 6 | 2 | .620 |
| | Using grammar in new contexts | 4 | 2 | 1.000 |
| **Reading** | Translation | 12 | 7 | .497 |
| | Guessing the meaning of new words | 4 | 0 | .497 |
| | Newspapers | 0 | 2 | .461 |
| | Magazines | 2 | 2 | 1.000 |
| | Stories | 2 | 0 | 1.000 |
| **Oral** | CDs | 2 | 5 | .59 |
| | Tapes | 2 | 10 | .173 |
| | TV programmes | 0 | 2 | .461 |
| | Films | 14 | 2 | .062 |

The results in Table 7.5 show that there were no significant differences between the two groups in any of the specific skills activities recommended. Nevertheless, it would be interesting to examine why some students recommended these activities and why some others did not.

## 7.2. RQ2: Why do the students report doing these activities in order to prepare for the SPE Test?

In this section, I present the students' statements which they used to justify their recommendations of learning activities. These statements included those which showed their knowledge about the test and those which indicated their purposes for test preparation. The results are presented in Tables 7.6 and 7.7.

**Table 7-6 Students' knowledge of the test**

| Statements indicating knowledge of the test | % GE | % SPE | P-value |
|---|---|---|---|
| Vocabulary is the most important | 20 | 17 | .789 |
| Grammar is the second most important | 12 | 7 | .497 |
| Grammar is the third most important | 2 | 2 | .875 |
| Reading is the second most important | 2 | 5 | .593 |
| Reading is the third most important | 2 | 0 | .186 |
| Cloze is the second most important | 0 | 2 | .461 |
| There is no listening section | 0 | 2 | .461 |
| Questions come from the Pre-university book | 2 | 0 | .121 |
| Questions come from out of the Pre-university book | 0 | 12 | .018* |
| Prior knowledge is necessary | 2 | 5 | .593 |

Based on the results in table 7.6, 20 out of 42 SPE students and 17 out of 49 GE students (almost half of the students from each group) thought that vocabulary was the most important skill and 12 SPE and 7 GE students (almost a quarter of the students from each group) recommended grammar as the second most important. An equal number of SPE students, i.e. 12, but none from the GE Group suggested that 'questions came from out of the Pre-university book'. This showed a statistical difference between the groups as well. The GE students did not mention this probably because they thought it was not relevant to them. However, 2 of them suggested that 'questions came from the Pre-university book'. Similarly the number of students who showed awareness of other test requirements was very few. Nevertheless, these statements were useful in that they gave an indication of possible reasons for the students' reported activities even though there were no statistical differences between the groups in these statements. Next, I look at the students' reasons for test preparation in Table 7.7.

**Table 7-7 Students' purposes for test preparation**

| Purposes | % GE | % SPE | P-value |
|---|---|---|---|
| Passing the exam is important | 31 | 10 | .012 |
| English is necessary for future use | 2 | 0 | .186 |
| English is key to international communication | 0 | 2 | .461 |
| Passing the exam creates job opportunities | 0 | 2 | .461 |

The results in Table 7.7 show that the students mentioned various reasons for

test preparation with different frequencies. The reasons included 'passing the exam', 'future use of English', 'English as the key to international communication', and 'job opportunities'. However, the students were only significantly different in one purpose, 'passing the exam'. More GE students mentioned 'passing the exam' than SPE students, which could mean that passing the exam was more important for GE students than for SPE students. The questionnaire results also showed that 'interest' was more important for SPE students than 'passing the exam' (Table 6.20). This could also mean that SPE students' reported activities might have been driven by factors other than the test.

## 7.3. RQ3: Do SPE students perceive the SPE Test to be more difficult than the GE Test?

As mentioned in the test specifications (Appendix 3), the difficulty level of the SPE Test has increased, according to the test authorities. Therefore, my expectation was that SPE students would consider the SPE Test as more difficult than the GE Test. The results in Table 7.8 show the differences between the SPE and GE tests in terms of perceived difficulty.

**Table 7-8 Students' views of the difficulty of the tests**

| | % GE | % SPE | P-value |
|---|---|---|---|
| SPE is more difficult than GE | 0 | 10 | .041 |
| SPE requires high knowledge | 0 | 2 | .460 |
| GE requires average knowledge | 2 | 0 | 1.000 |

The results in the first row of Table 7.8 show that from the perspective of 10% of the SPE students, the SPE Test was more difficult than the GE Test, which confirms my expectation. 2% of the SPE students comment that the SPE Test is difficult without comparing it with the GE Test and 2% of the GE students comment that the GE Test is of average difficulty without comparing it with the SPE Test. The reason these students did not compare the two tests might be because of the letter

writing instruction which was very general. An additional reason for the GE students (why they did not compare the two tests) might be because, unlike the SPE students, they were only going to take the GE Test not both tests.

## 7.4. RQ4: Do SPE students have better English backgrounds than GE students?

I answer this research question by the background information which I collected from the sample of letter writers. I asked them to write on top of the letters how many terms they had attended language schools. Since the difficulty level of the SPE Test has increased, I expected the SPE students to have a better English background than the GE students. The results are presented below.

**Table 7-9 Attendance in English institutes by terms**

| | SPE | GE | P-value |
|---|---|---|---|
| **Mean** | 13.20 | 4.71 | .004 |
| **SD** | 9.28 | 7.22 | |

As the p-value shows the SPE and GE students were significantly different. The SPE students attended longer (13.2 terms) in English institutes than the GE students (7.22 terms). This confirms the results from the questionnaire data as well (Table 5.26).

## 7.5. RQ5: What beliefs do the SPE students report holding about learning English?

To answer this research question, I examined the letters to see what belief statements the students made about learning. Table 7.10 includes the beliefs stated by some of the SPE and GE students.

**Table 7-10 Students' reported learning beliefs**

| Beliefs | % GE | % SPE | P-value |
|---|---|---|---|
| Language means arranging 'words' for communication | 0 | 2 | .461 |
| Vocabulary facilitates reading | 6 | 2 | .621 |
| Vocabulary and grammar complement each other | 0 | 2 | .461 |
| Vocabulary improves speaking | 2 | 0 | 1.000 |
| Speaking improves vocabulary as it provides for vocabulary practice | 2 | 0 | 1.000 |
| Language learning is gradual and takes time | 6 | 0 | .245 |
| English is elusive | 2 | 2 | 1.000 |

As the results show, 5 out of 7 statements concerned vocabulary, which suggests that the students believed vocabulary learning was important. The first statement in the table is exclusively about vocabulary and the next four concern the relationship between vocabulary and other skills, i.e. whether or how knowledge of vocabulary facilitated the learning of other skills or how the knowledge of other skills facilitated the learning of vocabulary. The last two statements concern, not skills, but English or language as a whole.

The belief that 'language means arranging 'words' for communication' stated by 2% of the SPE students could encourage them to focus on vocabulary rather than other skills. These students probably considered vocabulary as the main building blocks of English. They also said that 'vocabulary and grammar complement each other', which could encourage focus on both grammar and vocabulary. 6% of GE students and 2% of SPE students believed, 'knowledge of vocabulary facilitates reading comprehension'. This belief also might encourage focus on vocabulary because the students probably thought that by studying vocabulary they could also learn the reading skill. 4% of GE students believed in the mutual relationship between vocabulary and speaking. They believed that both skills helped improve each other. Another belief stated by 6% of GE students was 'language learning is gradual and takes time'. Finally, 2% of SPE and 2% of GE students believed that English was elusive. This belief could encourage students to practice and review the materials

often so that they would not forget them.

However, based on the contextual approach to beliefs (3.4.3), what is not clear is where these beliefs originate from, i.e. whether they originate from the test preparation experience or whether they were held in advance. Therefore, unless one talks to the students, it is not possible to determine the possible effect of the test on beliefs or whether the beliefs originated from other sources. For example, one of the letter writers stated that views of language and knowing how to learn were important but that the information they received about these issues was not adequate. Therefore, it would be interesting to talk to the student to see from which source he or she received or expected to receive the information, i.e. whether the source was the test, the teacher, other factors or (probably) all of them.

## 7.6. Summary

This chapter reported the results from the analyses of the letter data which was collected mainly based on the students' agenda, i.e. a general topic only asked the students what they suggested for test preparation. The students' suggestions were examined based on the predicted and unpredicted washback and in order of the research questions. It also tried to account for the reasons for the students' recommended activities through the students' test knowledge and their purposes for test preparation. The results showed that two groups of GE and SPE students were not statistically different on the majority of the recommendations but were significantly different in some of the predicted areas of washback including the use of 'extra materials', their knowledge of the test that 'the questions came from out of the pre-university book', in 'passing the exam' as their main purpose of test preparation, difficulty level of the test, and English background. I stated that the non-significant differences in the majority of the recommendations might be firstly because the

number of the students in each group was small (42 SPE and 49 GE), and secondly only some of the students made the recommendations. I also argued that it could be because of the shortage of time and the memory factor as the letters were written under time restrictions. This might have caused some students not to mention some of the issues they wanted to. It could also have been because the students were not good writers or not interested in writing in general. Therefore, it was suggested that provided the letter data were not collected under time restrictions, the students were good writers and interested in writing, the letters could prove as a useful instrument in washback studies.

The letter data also elicited information about the learning beliefs of the students. Most of the beliefs concerned vocabulary which suggested the importance of vocabulary to the students. However, it was not possible to determine whether there was any (reported) cause-effect relationship between the test and the beliefs, which called for probing the results further.

In short, this chapter reported how:

1. The letter data was based on the students' agenda and was collected under time limits.

2. The data were analysed quantitatively.

3. The results showed non-significant differences between SPE and GE groups except for five areas.

4. The significant results did not contradict those of the questionnaire.

5. The lack of significant results from the letter data could be due to the small sample size and the fact that a subset of the students from each group made the recommendations.

6. Knowledge of the test and purposes for test preparation were suggested as

possible reasons for the students' recommended activities.

7. It was not possible to determine whether the students' learning beliefs originated from the test or from other sources.

8. It was suggested that letters can be used as a data collection instrument in washback studies.

# Chapter 8. Results of the Interviews

In Chapter 6, I presented the results of the questionnaire analysis and in Chapter 7 the results of the letters analysis. In this chapter I will discuss the results of the interview analysis, which were conducted with 9 SPE and 9 GE students. In order to have as representative a sample of the interviewees as possible, I attempted to select students from different levels of English background, from different fields of study, and from different regions (Table 5-5 and 5-8). I used semi-structured interviews and had two purposes for the use of them. One of the purposes was to see the results from the different lens the interviews provided and the other was to explore the nature of the relationship between reported activities, the test, and beliefs. Accordingly, I used two types of questions in the interviews. The first type of question was to confirm the responses on the previously-used instruments and the second included wh-questions, in particular, 'why'. Responses to the former will be used to answer Research Question 1, and responses to the latter will be used to answer Research Questions 2, 5 and 6. However, to answer Research Question 3, I used the background information which I collected from the interviewees. Research Question 4 was not answered by the interviews because of practical problems (see 5.2.2).

As in the previous chapters, I will present the results in the following sections based on the research questions and predicted versus unpredicted washback (6.1.1, 6.1.2, 7.1.1, and 7.1.2).

## 8.1. RQ1: What activities do the SPE students report doing in order to prepare for the SPE Test?

As mentioned in the introduction to this chapter, the results reported in this section are responses the students gave in confirmation or disconfirmation of the responses to previous instruments. The reports are categorised based on predicted and

unpredicted washback.

### 8.1.1. Predicted washback

#### 8.1.1.1. Test sections

The ratings of the amount of time the interviewees in the two groups spent on preparing for the sections common to both tests were similar, i.e. each group spent most time on vocabulary, grammar, reading, and cloze in that order. For the SPE Group the two additional sections of the SPE Test followed next, i.e. they spent less time on these two sections than on the common sections. The results are presented in the following table.

**Table 8-1 Amount of time spent on the test sections in rank order**

| | | Common sections | | | | Additional sections | |
|---|---|---|---|---|---|---|---|
| | **No** | **Vocabulary** | **Grammar** | **Reading** | **Cloze** | **Language Functions** | **Sentence structure** |
| **SPE** | **1** | 2nd | 1st | 3rd | 4th | 5th or 6th | 5th or 6th |
| | **2** | 1st | 3rd | 2nd | 4th | 5th or 6th | 5th or 6th |
| | **3** | 1st | 3rd | 2nd | 4th | 7th | 7th |
| | **4** | 1st | 3rd | 2nd | 4th | 5th or 6th | 5th or 6th |
| | **5** | 1st | 2nd | 3rd | 4th | 5th or 6th | 5th or 6th |
| | **6** | 1st | 4th | 2nd | 4th | 5th or 6th | 5th or 6th |
| | **7** | 1st | 2nd | 3rd | 4th | 5th | 6th |
| | **8** | 1st | 2nd | 3rd | 4th | 7th | 7th |
| | **9** | 1st | 2nd | 3rd | 4th | 7th | 7th |
| | | | | | | | |
| **GE** | **1** | 1st | 3rd | 2nd | 4th | | |
| | **2** | 1st | 3rd | 2nd | 4th | | |
| | **3** | 1st | 3rd | 2nd | 4th | | |
| | **4** | 1st | 2nd | 3rd | 4th | | |
| | **5** | 1st | 3rd | 2nd | 4th | | |
| | **6** | 1st | 2nd | 3rd | 4th | | |
| | **7** | 1st | 2nd | 3rd | 4th | | |
| | **8** | 1st | 2nd | 3rd | 4th | | |
| | **9** | 1st | 2nd | 3rd | 4th | | |
| | | | | | | | |
| **Average Ranks** | **SPE** | 1.11 | 2.44 | 2.55 | 3.88 | 5.94 | 6.05 |
| | **GE** | 1 | 2.44 | 2.55 | 4 | | |
| P-value | | .940 | 1.000 | 1.000 | .966 | NA | NA |

**Note 1:** No = student No/ID, **Note 2:** 1st = most amount of time and 6th = least amount of time, **Note 3:** 5 or 6 = No difference between Language Functions & Sentence Structure, **Note 4:** 7 (as the figure out of the range of 1-6) = no time was spent on the section (if 0 was selected instead of 7, the average would be distorted i.e. less than others)

As Table 8.1 shows, in the sections common to both tests most time was spent on vocabulary, and the least amount of time was spent on 'cloze'. Grammar and

reading are in between and very close to each other.

Concerning the additional sections of the SPE Test, the table shows that the SPE students spent less time on these two sections than on the common sections. There was almost no difference between them in terms of the amount of time which was spent on them i.e. there was only one SPE student who reported spending more time on Language Functions than on Sentence Structure.

This pattern partly confirmed the results of the questionnaire (Table 6.1). It confirmed the results for the SPE Group but not the GE Group as more GE students had reported spending more time on reading than on grammar. The results of the interviews confirmed the results of the letter data both in terms of the common sections and the additional sections as the letters showed no statistically significant differences between the two groups (Table 7.1). I will discuss how the students justified the amount of time they spent on the test sections later in this chapter under Research Question 2 (8.2).

#### 8.1.1.2. Oral activities

Since oral activities included a number of activities, before the interview I made a list of these activities and asked the students whether they did them for their test preparation. These activities included the use of tapes and CDs, watching films and TV programmes, listening in general and listening to the radio, and speaking. The results are presented in Table 8.2.

**Table 8-2 Number of students reporting doing oral activities**

| | # of students out of 9 |
|---|---|
| **SPE** | 3 |
| **GE** | 2 |
| **P-value** | .655 |

As Table 8.2 shows, there was no difference between SPE and GE students in what they said about oral activities. The results of the interview confirmed the results

of the questionnaire for 'listening' and 'speaking' in general (Table 5.2) where both groups of SPE and GE ranked them lower than the 'test sections'. However, the results did not confirm the results from the questionnaire concerning specific oral activities as the questionnaire results had shown that there were significant differences between the two groups in using tapes and CDs, watching films and TV programmes, and listening to the radio (Tables 6.13 and 6.14). More SPE students did them than GE students, although both groups did these activities infrequently.

However, the results of the interview confirmed the results of the letters, as the letter data showed no significant differences between the students in oral activities (Table 6.2).

#### 8.1.1.3. Extra materials

Based on the results presented in Table 8.3, SPE and GE groups were statistically different in the use of extra materials including the use of *Bridging the Gap* but similar in the use of the Pre-university textbook.

**Table 8-3 Number of students reporting using extra materials, Pre-university textbook, and *Bridging the Gap***

| | Extra materials | School textbook | *Bridging the Gap* |
|---|---|---|---|
| **SPE** | 9 | 9 | 7 |
| **GE** | 1 | 9 | 0 |
| **P-value** | .011 | 1.000 | .008 |

These results confirm the results of the questionnaire and letters for extra materials where more SPE students used and recommended them than GE students (Table 5.13 and 6.4). However, the results do not confirm the results from the questionnaire or the letters for the pre-university textbook and *Bridging the Gap* where there were no differences between the two groups (Tables 6.3, 6.4 and 7.4).

#### 8.1.1.4. Preparation class

In attending preparation classes the two groups were statistically different. 8

SPE students attended preparation classes, while only two GE students attended such classes.

**Table 8-4 Number of students reporting attending preparation classes**

| | Preparation class |
|---|---|
| **SPE** | 8 |
| **GE** | 2 |
| **P-value** | .004 |

These results confirm the questionnaire result which showed that more SPE students attended preparation classes than GE students (Table 6.5). However, the result does not confirm the results of the letter data which showed no significant difference between the two groups (Table 7.4).

#### 8.1.1.5. Attending private English institutes

In order to examine what other out-of-school resources the students resorted to, I asked them whether they attended English language institutes to help them in their preparation. Only one SPE student but no GE student reported that they attended English institutes.

**Table 8-5 Number of students reporting attending private English language schools**

| | # of students attending English institutes |
|---|---|
| **SPE** | 1 |
| **GE** | 0 |
| **p-value** | .317 |

Although most SPE students and some GE students reported long attendance in language schools in the questionnaire (Table 6.26), they might have stopped attending because of test preparation (Bailey, 1996: 269). However, the results also confirm those of the letter data which showed no significant difference between the two groups (Table 7.4).

#### 8.1.1.6. English versus Non-English subjects

One of the intentions of the test authorities was to give more weight to English vis-a-vis Non-English subjects. Therefore, I asked the students on which subject they

spent more time. The results are presented in Table 8.6.

**Table 8-6 Number of students reporting spending more or equal amount of time on English and 'non-English subjects'**

| | More time on English | More time on Non-English | Equal amount of time |
|---|---|---|---|
| **SPE** | 4 | 1 | 4 |
| **GE** | 0 | 9 | 0 |
| **p- value** | .046 | .011 | .046 |

Table 8.6 shows the two groups were significantly different. While four SPE students spent more time on English, four others spent equal amount of time on English and Non-English subjects. There was only one SPE student who spent more time on Non-English. However, all GE students spent more time on Non-English subjects. These results confirm the results of the questionnaire which showed that a greater number of SPE students spent more time on English than on Non-English subjects (Table 6.6). However, the results do not confirm those of the letter data which showed no significant difference between the two groups (Table 7.3).

### 8.1.2. Unpredicted washback

#### 8.1.2.1. Written activities

Since written activities, like oral activities, included a number of activities, before the interview I made a list and asked the students whether they did them for their test preparation. The written activities included reading newspapers and magazines, reading stories, and writing daily notes or diaries. The results are presented in Table 8. 7.

**Table 8-7 Number of students reporting doing written activities**

| | # of students out of 9 |
|---|---|
| **SPE** | 3 |
| **GE** | 0 |
| **P- value** | .083 |

Like oral activities, there were no statistically significant differences (.083) between SPE and GE students in these activities. 3 SPE students said they did the written activities, but no GE student reported doing them.

The results of the interview did not confirm the results of the questionnaire as the questionnaire results had shown that there were significant differences between the two groups in the written activities with more SPE students doing them than GE students (Tables 6.11 and 6.12). However, these activities were not mentioned in the letters.

#### 8.1.2.2. Translation activity

There was no significant difference between SPE and GE students in translation. Of the 9 SPE students, one never used translation when reading a passage, but all others did. Similarly, all the GE students, except for one, used translation while reading. These results contradict those of the questionnaire which showed that the SPE group used translation more often than the GE group (Table 6.10). However, they confirm the results of the letters which showed no significant difference between the two groups (Table 7.5).

**Table 8-8 Number of students reporting translating**

| | # of students out of 9 |
|---|---|
| **SPE** | 8 |
| **GE** | 8 |
| **P- value** | 1.000 |

In the interview, some students considered two other activities related to translation: one 'paying attention to the topic of the text' while reading, and one getting a 'general meaning of the text' by skimming, i.e. reading in stages. These skills were the ones which were recommended in their textbooks as well. The following table reports the students' paying attention to the topic.

**Table 8-9 Number of students reporting paying attention to the topic**

| | # of students out of 9 |
|---|---|
| **SPE** | 9 |
| **GE** | 9 |
| **P- value** | 1.000 |

The following table reports the students' reading in stages, i.e. first for general

meaning and second for full understanding.

**Table 8-10 Number of students reporting reading in stages**

| | # of students out of 9 |
|---|---|
| **SPE** | 9 |
| **GE** | 7 |
| **P- value** | .617 |

According to Tables 8.9 and 8.10, the two groups were similar in these activities. These results contradict those of the questionnaire which showed that more SPE students did these activities than GE students (Table 6.10). However, it is not possible to compare these results with those of the letters as the letter writers did not write about these activities.

### 8.1.2.3. Guessing activity

As Table 8.11 shows, in the guessing activity also the two groups were similar.

**Table 8-11 Number of students reporting guessing**

| | # of students out of 9 |
|---|---|
| **SPE** | 8 |
| **GE** | 6 |
| **P- value** | .593 |

One student from the SPE group and three students from the GE group did not guess the meaning of unknown words while reading. These results do not confirm the results of the questionnaire which showed that more SPE students applied the reading skill of guessing (Table 6.10). However, the results do confirm those of the letter data which showed no significant difference between the two groups (Table 7.5).

### 8.1.2.4. Memorisation activity

Next, I asked each student if they memorised grammar rules. As Table 8.12 shows, there was no statistical difference between SPE and GE students in this activity. All the 9 SPE students and 8 GE students reported memorising grammar rules.

Two of the SPE students and one GE student also talked about the

supplementary activity of using texts to help their learning of grammar without my asking them if they did so. However, as for the other students who did not talk about this activity (using texts), I asked directly whether they used texts to supplement their memorisation. All the SPE as well as all the GE students reported doing these activities.

**Table 8-12 Memorisation and use of texts for grammar learning**

| Activities | # of SPE students | # of GE students |
|---|---|---|
| **Memorisation** | 9 | 8 |
| **Texts (sentences & reading passages)** | 9 | 9 |
| **P- value** | 1.000 | .808 |

The interview results concerning memorisation of grammar rules confirm the results of both the questionnaire data (Tables 6.9) and the letter data (Table 7.5) as neither of them showed any significant differences between SPE and GE groups. However, the results concerning the other activities do not confirm the questionnaire but do confirm the letters. The questionnaire had shown that SPE students reported doing the activities of using reading texts and using examples more than GE group (Tables 6.9). The letter data, however, showed no significant differences between the groups (Table 7.5).

### 8.1.2.5. Practice tests

Concerning practice tests, first I asked the students whether they used them before I asked them why. The results were similar for both groups. All the students from both groups reported practicing with sample exam questions. The results are presented in Table 8.13.

**Table 8-13 Number of students using practice tests**

| | # of students out of 9 |
|---|---|
| **SPE** | 9 |
| **GE** | 9 |
| **P- value** | 1.000 |

These results do not confirm the results of the questionnaire which showed that

more SPE students practiced with sample items (Table 6.16).

## 8.2. RQ2: Why do the students report doing these activities in order to prepare for the SPE Test?

After I asked the students to confirm or disconfirm their responses on the preceding instruments, I asked them why they did the activities which they reported they did. This section will focus on the reasons for the predicted washback activities as well as 'guessing' activity.

### 8.2.1. Reasons for test sections

The interview question for the test section activities was 'why do you study each section of the test?' Students gave various reasons to justify the amount of time they reported spending on each section. There were reasons common to both groups as well as reasons exclusive to each group. First I will explain the common reasons and then the exclusive reasons. Table 8.14 shows a summary of the results.

**Table 8-14 Reasons for spending time on each test section/skill**

| | Reasons | Students' codes/IDs | | Sources of effect | | | |
|---|---|---|---|---|---|---|---|
| | | | | Test | | Others | |
| | | SPE | GE | SPE | GE | SPE | GE |
| 1 | Number of items/weight | 1,2,3,4,5,6,7,8,9 | 2,3,4,5,6,8,9 | 9 | 7 | | |
| 2 | Interdependence of skills/view of constructs | 1,2,3,4,5,6,7,8,9 | 2,3,4,5,6,7,8,9 | | | 9 | 8 |
| 3 | Teacher's emphasis | 3,5 | 3 | | | 2 | 1 |
| 4 | English background | 3 | 9 | | | 1 | 1 |
| 5 | Difficulty/ simplicity | 1 | 2,6,9 | 1 | 3 | | |
| 6 | Dis/liking a section | 2 | | | | 1 | |
| 7 | Speed for test taking | 2,4,5 | | 3 | | | |
| 8 | Predictability of questions | 2,3 | | 2 | | | |
| 9 | Exams or questions in school | 3 | | | | 1 | |
| 10 | Belief that grammar is fixed or limited | 3 | | | | 1 | |
| 11 | Weakness in a section | 1 | | | | 1 | |
| 12 | Importance of a section in private institutes | 1 | | | | 1 | |
| 13 | Familiarity with questions (so practice tests) | 2,4,5 | | 3 | | | |
| 14 | Unsuitability of textbooks (so practice tests) | 1 | | | | 1 | |
| 15 | Point of time in preparation period | 1,8,9 | | | | 2 | |
| 16 | Learning materials | | 1,3,5,7,9 | | | | 5 |
| 17 | Usefulness for answering exam questions | | 5,7,9 | | 3 | | |
| 18 | Desire for a good mark | | 6 | | 1 | | |
| 19 | Time to be spent in test taking | | 7 | | 1 | | |
| 20 | Availability of resources | | 8 | | | | 1 |
| Total | | | | 18 | 15 | 20 | 16 |

**Note:** The numbers in the last four columns indicate the number of students/frequency of the reasons.

Reasons 1- 5 were the ones stated by the two groups in common. The table shows that all the 9 SPE students and 7 GE students justified the amount of time they spent on each section based on the number of items (Reason 1) (SPE 1: 'out of 70 questions, about 50 questions are related to vocabulary and only 20 to grammar') and the relationship between the sections/skills (Reason 2). These two reasons were the most frequent. Concerning 'interdependence of the skills' (Reason 2), the students from both groups considered vocabulary and grammar the central sections and the other sections either dependent on vocabulary knowledge or grammar knowledge or both. Typically Reading was related to vocabulary, Language Functions to vocabulary, Sentence Structure to grammar, and Cloze to both vocabulary and

grammar (SPE 1: Cloze includes both vocabulary and grammar', GE 5: 'If you know vocabulary you can answer cloze questions easily'). The teacher's emphasis in class on a particular skill was another reason stated by 2 SPE students and 1 GE student (Reason 3) (SPE 5: 'The teacher said it [cloze] was important'). One SPE and one GE student used the English background to justify why they did not spend much time on a section (Reason 4). For example, SPE 3 said 'Language Functions is conversational and requires idioms, but because I've attended English institutes, I don't have a particular problem with this section'. Concerning reading, she said that she did not spend much time on reading because it was 'heavy and difficult' for her and therefore she studied vocabulary in order to learn it (Reason 5). Similarly, students who considered a section easy did not spend much time on that section as GE 2 for instance reasoned that 'grammar only required basic-level knowledge'.

Reasons 6-15 were only stated by SPE students but not by GE students. SPE 2 said, 'I don't study grammar very often because I don't like grammar' (Reason 6). One of her concerns was saving time on the exam by practicing reading to increase the reading speed (Reason 7). The reason why she said she did not spend much time on 'Sentence Structure' was because she said, 'The questions in this section are predictable' (Reason 8). SPE 3's reason for studying grammar and vocabulary was because they were asked about them in their school class or school exam (Reason 9) ('SPE exam is mainly grammar and vocabulary and also because in school we are asked questions on them'). She also believed that, unlike reading and vocabulary, grammar was fixed i.e. had fixed rules which once you learned, you could apply them in other places (Reason 10) ('Grammar doesn't take much time because it is fixed. There's not much difference between pre-university grammar and university level grammar. There's a base to which, at most, a little material is added'). Reason 11,

'weakness in a section', was mentioned by SPE 1 who spent time on grammar because she said, 'Grammar is my major problem. If I don't know grammar, I cannot do anything'. Concerning the effect of private classes (Reason 12), she stated, ('In the language school which I go to they don't emphasise grammar but how much you know grammar is important there). 3 SPE students reasoned that they did not spend much time on Language Functions and Sentence Structure because they thought that familiarity with the questions of these sections would be enough and only used practice tests for this purpose (Reason 13) (SPE 5: 'I can't say which [Language Functions or Sentence Structure] is more important because only familiarity with these questions and practice test are ok.'). Similarly, one SPE student did not allocate separate time to Language Functions and Sentence Structure because she thought that SPE textbooks were not suitable for these two sections and therefore used practice tests instead (Reason 14). SPE 8 and 9 who attended the same preparation class said they were not currently studying Sentence Structure and Language Functions, but their teacher had a plan for those sections (Reason 15) (SPE 8: 'Our preparation class teacher has a plan but we don't know yet.', SPE 9: 'Our preparation class teacher is going to teach us later').

Table 8.14 also shows that GE students gave some reasons which SPE students did not (Reasons 16-20). Reason 16 (learning materials) means that the students' degree of emphasis on a particular section depended on the textbook. For example, GE 1 said, 'I just study based on the textbook. The textbook emphasises vocabulary and then reading, and grammar is only briefly explained in the book'. Reason 17 (usefulness in answering questions) means that some students thought that practice of a section would make it easier to answer exam questions. Desire for a good mark on the exam (Reason 18) was another motive behind practising a section (GE 6: 'There is

a reading passage in the exam because of which we should learn reading so that we can get a good mark on this section'). The amount of time which a test section required in the exam was another factor determining the amount of time to be spent on the section in test preparation (Reason 19). In this regard, GE 8 said: 'I don't spend much time on reading because in the exam it is time consuming but has the same weight as other sections. You should also spend time reading the passage'. Finally, GE 8 talked about 'availability of resources' (Reason 20) as another determining factor: 'Mainly we practice grammar and meaning for cloze tests because most of the exam is related to meaning. Also there's no particular resource to refer to for cloze tests.'

In summary, although both groups were largely affected by the test, they were also affected by factors other than the test. Among the test factors, both groups were affected by the weight of the test sections, the difficulty level of the sections, and time management in test taking. Weight of the sections had the most effect on both groups. However, while SPE students spent time on the sections depending on the predictability of the questions and the degree to which they were familiar with the questions, GE students were affected by a desire for a good mark and whether spending time on a section made it easier to answer the questions.

Among factors other than the test, both groups were affected by their views of the interdependence of the test sections or constructs, emphasis of a section by the teacher or school, their English ability, and learning materials. The two groups were affected most by their views of the test constructs. From the exclusive reasons each group stated, some SPE students were influenced by whether they liked the skill in general, weakness in a section, English institutes, point of time in the preparation period, and their view of a single construct, while the only exclusive non-test factor

which influenced the GE students was the availability of resources.

### 8.2.2. Reasons for the oral activities

As Table 8.2 showed, there were three SPE students (Students 1, 2, and 4) who reported doing oral activities. They had two kinds of purposes for doing these activities: one was for the SPE Exam and one included purposes other than the exam. The types of reasons the students gave are summarised in Table 8.15.

**Table 8-15 Reasons from SPE students who reported doing oral activities**

| ID | Did because of SPE Test | Did for other reasons | Sources of effect | |
|---|---|---|---|---|
| | | | Test | Others |
| 1 | For structures, vocabulary, and reading | For future use of English, for improving English in general, for institute class, and for speaking to my teacher, classmate, and sister | 3 | 5 |
| 2 | Vocabulary & grammar<br>Sentence structure<br>Provides for practice | For improving English in general, stronger proficiency, pronunciation, communication ability, self-confidence in speaking, oral learning better than learning through textbooks | 4 | 6 |
| 4 | Vocabulary, idioms, grammar,<br>Sentence structure | For improving English in general | 4 | 1 |
| Total | | | 11 | 12 |

**Note:** The numbers in the last two columns indicate the number of reasons for each student.

According to Table 8.15, SPE 1 said that oral activities were useful for the structures, vocabulary and reading of the SPE Test. In addition to the SPE Test, she said she did these activities because of the English institute she went to as they were required by the institute and because she wanted to speak English with her tacher, classmates, and her sister. She also thought they were useful for improving English in general and for her future use of English. However, she said she could not do these activities as often as before because of lack of time. Similarly, SPE 2 talked about the usefulness of oral activities for the SPE Test in terms of grammar and vocabulary and that they provided for more practice. She also said these activities were useful for improving English in general, stronger language proficiency, communication ability and self-confidence. She also believed that learning through oral activities was more

efficient than learning through textbooks. Finally, SPE 4 also had two kinds of reasons for doing oral activities. She said oral activities were useful for learning SPE vocabulary, idioms, grammar, and sentence structure as well as for language learning in general.

**Table 8-16 Reasons from SPE students who did not (reportedly) do oral activities**

| ID | Did not do because of SPE Test | Did not do for other Reasons | Sources of effect | |
|---|---|---|---|---|
| | | | Test | Others |
| 3 | No listening or speaking in SPE Test (for exams like TOEFL) | No interest in radio, no access to satellite, BBC programmes difficult | 1 | 3 |
| 5 | Not essential for SPE Test (for improving English in general) | | 1 | |
| 6 | No listening or interview in SPE Test, not main resources (for English in general such as communication, translation, vocabulary, and idiom) | | 1 | |
| 7 | There are more important things to do than oral | No time | 1 | 1 |
| 8 | SPE Test is mainly vocabulary and grammar (for mastery of English) | No time | 1 | 1 |
| 9 | Not for SPE Test (main purpose is success in the exam) | | 1 | |
| Total | | | 6 | 5 |

**Note:** The numbers in the last two columns indicate the number of reasons for each student.

As Table 8.16 shows, the other 6 students reported not doing oral activities and gave various reasons for this. SPE 3 said there were no listening or speaking sections on the SPE Test. She believed that oral activities were not necessary for the SPE Test but for tests such as TOEFL. Lack of interest, lack of access and difficulty were her other reasons. She said she was not interested in the use of radio, did not have access to satellite TV and that programmes such as those from the BBC were difficult. SPE 5 believed that oral activities were not 'essential' for the SPE Test but 'improving English in general'. SPE 6 thought that oral activities were not the main resources as there was no listening or interview on the SPE Test. However, she believed that they could be useful for improving English in general in terms of vocabulary, idioms, and translation as well as for improving communication ability. SPE 7 believed that there were more important activities to do for the SPE than oral activities. In addition, she

said she did not have time to do oral activities. SPE 8's reasons were based on her knowledge of the SPE Test and lack of time. She stated that because SPE mainly consisted of vocabulary and grammar, oral activities were not necessary. However, she believed that these activities would be useful for mastery of English in general. SPE 9 believed that oral activities were not necessary for the SPE Test, nor were they useful for other purposes as her only purpose was success on the exam.

**Table 8-17 Reasons from GE students who (reportedly) did oral activities**

| ID | Because of GE Test | Did for other reasons | Sources of effect | |
|---|---|---|---|---|
| | | | Test | Others |
| 1 | Not for GE | For interest, improving English in general, speaking to someone in future | | 3 |
| 6 | Useful (but not essential), for vocabulary e.g. uses of words, for sentence structure | English is the international language, for improving English in general, interest | 2 | 2 |
| Total | | | 2 | 5 |

**Note:** The numbers in the last two columns indicate the number of reasons for each student.

As Table 8.17 shows, GE 1 did oral activities not for the GE Test but because she was interested in English and wanted to improve her English in general, and use the ability in case she wanted to speak to somebody in the future. However, GE 6 had the test as well as other purposes in mind. Although he did not see oral activities as essential, he thought they were useful for the GE in that these activities would lead to learning the meaning of words, the uses of words in different contexts, and to the ability to make sentences which would be useful for both the GE Test and for other purposes. He also stated that he did oral activities because they were useful for improving English in general, because he was interested in English and because English was the international language. This student's activities were affected by both the GE Test as well as other factors.

**Table 8-18 Reasons from GE students who did not (reportedly) do oral activities**

| ID | Did not do because of GE Test | Did not do for other reasons | Sources of effect | |
|---|---|---|---|---|
| | | | Test | Others |
| 2 | GE is simpler than oral skills, not essential (for other purposes such as future job, other exams, English courses at university, etc) | | 2 | |
| 3 | Not essential (useful in future e.g. for getting a job) | Teacher doesn't ask | 1 | 1 |
| 4 | Not essential for GE, only indirectly effective i.e. could be useful e.g. if someone can speak they can answer exam questions easily (for job, for traveling abroad, for uses in computer, for improving English) | No time, no access | 1 | 2 |
| 5 | Not essential (for English in general) | No time, time consuming e.g. looking up words, I don't have the ability, I can't say sentences correctly | 1 | 5 |
| 7 | Only could be useful | But no time, my English is not good, access difficult | 1 | 3 |
| 8 | Only could be useful (for mastery of English) | But no time | 1 | 1 |
| 9 | Only could be useful | But no time | 1 | 1 |
| Total | | | 8 | 13 |

**Note:** The numbers in the last two columns indicate the number of reasons for each student.

Based on the results in Table 8.18, GE 2 did not do oral activities because she thought they were not essential and because the GE Test was not at that level of difficulty to require oral language practice. She thought such activities could be useful for exams other than the GE Test, for English courses at university and for future jobs. GE 2's statements about oral activities matched her purposes. GE 3 did not do oral activities because she thought they were not essential and also because their teacher did not ask them to do such activities. However, she believed that speaking fluently would be useful in future for getting a job or for other purposes. GE 4 thought that oral activities were not essential but could have some indirect effect and be useful in that if somebody had the ability to speak they could answer exam questions easily. However, despite the fact that he thought oral activities could be useful, he said he did not have time or access to oral materials in order to engage in oral activities. Although he could not do oral activities for the GE Test, he believed that these activities could

be useful for improving English in general, future jobs, travelling abroad, uses in different areas (such as in computer) or any other desired purposes. GE 5 did not do oral activities because she thought they required activities such as looking up a word which were time-consuming and she did not have enough time. She also said she did not do them because she did not have 'the ability to say sentences correctly'. However, she believed that these activities were useful for improving English in general. Although GE 7, 8, and 9 believed that oral activities could be useful for the GE Test, all three students said they did not have time to do oral activities. GE 7, however, gave additional reasons for not doing oral activities which included her weakness in English and problems of access to oral materials. GE 8 also said that these activities would be useful for mastery of English in general.

In summary, both groups of students including those who did and those who did not do oral activities were affected almost equally (based on the number of reasons) by the test and factors other than the test. SPE students who did oral activities did them mainly for grammar and vocabulary as far as the SPE Test was concerned. However, they also did them for improving their English in general, for English institutes, self-confidence, and because of the efficiency of learning through oral activities rather than written activities. The SPE students who did not do oral activities, reasoned based on the fact that there was no oral component in the test and that it mainly comprised vocabulary and grammar. The reasons not related to the test included lack of interest in certain audio devices (radio), access problem, difficulty of oral language, and time limitations.

The two GE students who did oral activities had reasons different from each other. While one of them did the activities for the test (grammar and vocabulary) as well as purposes not related to the test (interest and 'English is the international

language'), the other did them not for the test but for factors other than the test, which included interest, ability improvement in English in general, and desire to speak to someone in the future. The GE students who did not do oral activities reasoned that although they could be useful, they were not essential, and that oral skills were more difficult than the GE Test required. The non-test reasons they used to justify their not doing oral activities included time limitation, lack of access, weakness, inability to say things accurately, and not being asked by the teacher.

### 8.2.3. Language skills in class and in textbooks

Another issue I was interested in was to what extent washback was mediated by the teacher and learning materials. Table 8.14 showed that some students talked about the degree to which test sections were emphasized in class and in the textbooks in answer to my general question of 'why they spent the amount of time they reported they did on the test sections'. However, in the case of students who did not talk spontaneously about them, I asked the question myself, i.e. to what extent the test sections were emphasized by their teacher and in their learning materials. There was consensus among both SPE students and GE students that the only skills emphasised by teachers and materials were vocabulary, grammar and reading, whether in school class or preparation class, which corresponds to the results in Tables 6.1, 7.1, and 8.1. As I mentioned in the 'Test specifications' (Appendix 3), the (common) Pre-university textbook emphasises the three test sections of vocabulary, grammar and reading, and the book, *Bridging the Gap*, includes reading and vocabulary but no grammar.

### 8.2.4. Reasons for extra materials

This section concerns the use of extra materials versus the use of the Pre-university textbook. In the interview sessions, first I asked the students to confirm their answers on the previous instruments and then I asked them why they spent more

time on one than on the other or an equal amount of time on them. The use of extra materials was one of the new requirements of the SPE Test (Appendix 3). The results are presented in Tables 8.19 and 8.20.

**Table 8-19 SPE students' reasons for reportedly using extra materials versus Pre-University book**

| ID | P | Ext | Pre-university book | Extra materials | Sources of effect | |
|---|---|---|---|---|---|---|
| | | | | | Test | Others |
| 1 | + | ++ | Good passages, grammar explanation, & exercises for GE | For SPE questions, from preparation class, from English institute, commercially available books | 1 | 2 |
| 2 | + | ++ | For GE questions | For developing good background for SPE, Pre-uni book uninteresting, from school teacher | 1 | 1 |
| 3 | + | ++ | For GE questions. Going to study once and review near exam | Pre-uni book hasn't changed but LP increased, contradictory, has nothing to offer for SPE, a TOEFL book & a grammar book recommended by prep class teacher | 1 | 3 |
| 4 | + | ++ | For GE questions, taught at school | Pre-uni book has nothing to do with SPE eg easy readings, learning at school not enough, even if admitted problem at university | 1 | 3 |
| 5 | + | ++ | For GE Section | For SPE Section | 1 | |
| 6 | + | ++ | For GE questions | No particular source for SPE at school, Pre-uni book only good for school | 1 | 1 |
| 7 | + | + | For GE Section, includes very important words because also includes previous years, taught in both school & prep class | For SPE questions, recommended or taught by prep class teacher, of appropriate difficulty, emphasis up to the teacher | 1 | 2 |
| 8 | + | + | For GE Section, taught in both school & prep class, compulsory and we are given tests at school | For SPE questions, from prep class teacher, emphasis up to the teacher | 1 | 1 |
| 9 | + | + | Some questions from Pre-uni book | Some questions from extra materials, Prep class teacher's materials, lecture notes & practice tests | 1 | 1 |

**Note 1.** P: Pre-university book, Ext: Extra materials, ++: more time was spent, +: equal amount of time was spent, **Note 2.** The last two columns are concerned with the use of extra materials

SPE 1-6 stated that they studied extra materials more (++) than the Pre-University textbook. SPE 1 justified the use of the pre-University book by saying 'the pre-university book is suitable for the GE section of the exam ... has good reading passages, explains grammar, and if we do the exercises it's good'. Concerning extra materials, she said that they were useful for SPE questions and that she received them

from three sources: from the preparation class, from the institute she attended, and the commercially available ones. She said the institute materials were 'useful for both the institute and the SPE exam'.

SPE 2 said the Pre-University book was the main source for the GE section of the exam. Concerning extra materials she said they were useful for developing a good background for SPE and that the pre-university book was discouraging and made them lose their interest. As she did not attend a preparation class she asked her school teacher what materials to use.

SPE 3 believed that the Pre-University book was appropriate for the GE section and that she was going to study the book only once and review it near the exam. She also argued in favour of extra materials that the Pre-university book with 50 pages of repetitive materials had nothing to offer, that the book had some problems for example in one chapter it said something and in the next chapter it contradicted it, and that the language proficiency of the students had increased but the pre-university book was the same as before. Concerning the types of materials, she said they were recommended to use a TOEFL book and a grammar book by their preparation class teacher.

Similar to SPE 2 and 3, SPE 4 said that GE questions were based on the pre-university book. She also said that the book was only taught at school. In justification of the extra materials she said the pre-university book had nothing to do with the SPE exam, for example its readings were very easy and that even if students were admitted to university based on Pre-university level knowledge, they could not make any progress at university. However, she criticised school in that they did not learn that much English there, e.g. in terms of vocabulary and reading the SPE exam demanded.

SPE 5 talked about learning materials from the perspective of test demands.

She said the Pre-university book was useful for the GE section of the test and extra materials were useful for the SPE section of the test.

SPE 6 also agreed with the above students that the Pre-university book was 'the primary source for GE'. She added that if she did not know the GE, she could not move on to the SPE. She justified the use of extra materials saying that the Pre-university book was only good for school and that in school no particular sources were taught for the SPE Test.

SPE 7 and 8 reported that both the Pre-university book and extra materials were necessary for the SPE Exam. However, they were not sure which they had to spend more time on and said the emphasis on either type of materials would depend on the teacher. Concerning the necessity of the Pre-university book, SPE 7 said that there were very important words in the Pre-University textbook because it included all the words of the previous school years and that the textbooks of the previous school years were important because 'students should learn English from the basics'. She also stated that the Pre-university book was important because it was taught both in school and in the preparation class. Concerning extra materials, she said she used those materials which their preparation class teacher recommended or taught. She emphasised that she only used materials which her teacher recommended and which were at her level and that if they were at a high level they were not good.

SPE 8 justified the use of the Pre-university book saying that it was compulsory in school and that the school teachers gave them tests on the book. Similar to SPE 7, she said it was important also because it was taught both in school and in the preparation class. Regarding extra materials also her comment was similar to that of SPE 7. She said she received extra materials from the preparation class teacher and that it depended on the teacher whether the Pre-university book or extra

materials should be emphasised.

SPE 9 shared her idea with SPE 7 and 8 that both the Pre-university book and extra materials were important and was not yet sure which one should be emphasised more. She said extra materials included those from the preparation class teacher and the notes taken based on the teacher's lecture. Table 9.23 shows the results for the GE group.

In summary, all the SPE students were affected by the requirement of the SPE Test for extra materials. They were also affected by factors other than the test except for one student (SPE 5). The non-test factors included effect from the class or teacher, from English institutes, whether the learning materials were interesting, and of appropriate difficulty, whether they were updated in terms of difficulty to catch up with the students' increased proficiency, insufficiency of school materials and classes, and finally the consequences of using extra materials or their anticipated usefulness after admission to university. Next, the results for the GE students are presented in Table 8.20.

**Table 8-20 GE students' reasons for (reportedly) using extra materials versus Pre-University book**

| ID | P | Ext | Pre-university book | Extra materials | Sources of effect | |
|---|---|---|---|---|---|---|
| | | | | | Test | Others |
| 1 | ++ | + | Main resource for GE | Open our mind i.e. improve our English in general | 1 | 1 |
| 2 | + | - | Enough for GE, more important than extra materials | Not needed because GE is not SPE | 1 | |
| 3 | + | - | Enough if with practice tests | | 1 | |
| 4 | + | - | For majority of the GE questions | Maybe only 1% of questions | 1 | |
| 5 | + | - | Enough if studied completely | | 1 | |
| 6 | + | - | GE from Pre- uni book , the book with practice tests | Exam fairer without extra materials | 1 | |
| 7 | + | - | Enough because includes previous years as well | | 1 | |
| 8 | + | - | Enough, but with supplementary materials | Supplementary materials for exercises needed | 1 | |
| 9 | + | - | Enough because includes previous years as well | | 1 | |

**Note:** ++: more time was spent, +: was used, -: was not used, P: Pre-university book, Ext: Extra materials

All the GE students agreed that the Pre-university book was the main resource for the GE Test. However, a few students said it should be used with supplementary materials and practice tests and one of them used extra materials as well. Although GE 1 emphasised the Pre-university book more than extra materials, she also used extra materials because she thought they 'opened her mind' or improved her English in general. GE 3 and 6 stated that practice tests should also be used with the Pre-university book and GE 8 stated that supplementary materials should be used for exercises in the book, e.g. those commercially available materials for grammar which categorised and explained the grammatical points of the Pre-university. GE 6 added that the GE Test would be fairer without extra materials as there were some students who did not have facilities and could not attend a private language institute and therefore could not answer questions out of the Pre-university book.

In summary, the results of the interview showed the washback of the SPE Test in terms of requiring students to use materials in addition to the school textbook. All the SPE students used extra materials because of their test. However, one GE student also used extra materials not because of her test but because of her own interest. Among SPE students, there were also factors other than the test which affected the choice of extra materials. These factors included the class or the teacher, private language school, attractiveness of learning materials, difficulty level, updatedness of the materials, insufficiency of school materials and classes, and finally usefulness of materials for future.

### 8.2.5. English versus Non-English subjects

As mentioned before, one of the changes introduced in the SPE Test was that

the weight of Non-English subjects was kept constant and instead more weight was given to English (Appendix 3). In the interviews, first I asked the students whether they spent more time on English or Non-English subjects (Table 8.6) and then I asked them why.

**Table 8-21 SPE reasons for the time they (reportedly) spent on English versus 'non-English subjects'**

| ID | E | N | Reasons |
|---|---|---|---|
| 1 | ++ | + | English has more weight<br>Non-English subjects for success and raising our rank |
| 2 | + | + | English and Non-English subjects are equal<br>My teacher says Non-English subjects are very important. |
| 3 | + | ++ | At present more time on Non-English subjects because I'm weaker in these subjects.<br>Later will spend more time on English as it has more weight |
| 4 | + | + | Non-English subjects as important as English because weakness in SPE but good performance in Non-English subjects may result in admission (however problem at university). |
| 5 | + | + | Both are important. According to my friends, not knowing SPE very much but good performance on Non-English subjects results in admission |
| 6 | ++ | + | Priority is with English, but at least we should answer 50% of Non-English section correctly.<br>Non-English subjects are important because the sources are the same books taught in school. |
| 7 | + | + | Almost half of the SPE Exam is Non-English subjects.<br>A Non-English subject at which most students are weak is more important even with low weight. |
| 8 | ++ | + | If 50% correct on SPE, may succeed but if 50% on Non-English subjects, unlikely to succeed.<br>Non-English subjects could lead to a good rank i.e. below hundred. |
| 9 | ++ | + | English should be emphasized more.<br>Non-English subjects also can compensate for low marks on SPE to great extent. |

**Note:** E= English, N = Non-English subjects, ++ = more time was spent, one plus sign in each column = equal amount of time was spent

As shown before in Table 8.6, four students (SPE 1, 6, 8, and 9) spent more time on English, while four others (SPE 2, 4, 5, and 7) spent an equal amount of time on English and Non-English subjects. SPE 3, however, was the only student who spent more time on Non-English subjects. In justifying the emphasis on either set of subjects, all the students used their test knowledge explicitly or implicitly, for example SPE 1 said, 'English has more weight than Non-English subjects' and SPE 8 said, 'If you get 50% of answers correct on SPE, you may succeed but if you get 50% of answers correct on Non-English subjects, you are unlikely to succeed'. However,

other reasons were also mentioned by some students: SPE 2 said the reason why she emphasised Non-English subjects was because of her teacher's advice and SPE 5 said it was because of her friends' advice. SPE 3 reasoned because she was weak in Non-English subjects, temporarily she focused on Non-English subjects and that later she was going to spend more time on English. SPE 6's reason was accessibility of the learning materials. She said because Non-English subjects were taught in school, it was important to study them. SPE 7 stated weakness of their peer test takers on a subject was more important than the weight of the subject, which would make them overtake their peers. Finally, SPE 9 justified studying Non-English subjects as a compensation for inadequate knowledge of SPE.

In summary, all the students were affected by the SPE Test as all of them referred to their knowledge of the test in terms of which subjects to emphasise. However, there were also other factors which made the students make particular decisions. These factors included teachers' advice, friends' advice, weakness in a particular subject, access to the materials, weakness of rivals, and compensation for a subject of which there was inadequate knowledge.

The fact that almost half of the SPE students considered Non-English subjects of equal importance to English might mean reconsideration of the weighting by the test authorities. However, it is also possible that the students' views about weighting might change later in the preparation period. Table 8.22 presents the results for the GE Group.

**Table 8-22 GE students' reasons for the time they (reportedly) spent on English versus Non-English subjects**

| ID | E | N | Reasons |
|---|---|---|---|
| 1 | + | ++ | Non-English subjects are more important because I don't want to study English at university. |
| 2 | + | ++ | I don't spend much time on GE because I'm not counting on it very much as I'm weak in English and have little interest in English. |
| 3 | + | ++ | I spend more time on specialised subjects as GE is one subject but specialised subjects are many.<br>But I study GE more than other general courses because interested in English more than the other general courses<br>Our preparation class teacher gives us a test on GE every week but the other teachers don't do this. |
| 4 | + | ++ | Specialised courses are more important than GE.<br>GE increases the chance of success as someone may get a lower mark than me on the GE and I may answer the GE questions better<br>In general also it is useful because if your English is good you will have a better chance for employment. |
| 5 | + | ++ | I study specialised more than GE as they are more important.<br>I study all general courses (GE and General Non-English subjects) in turn and I may decide to study the other courses later. |
| 6 | + | ++ | Spend dead time on English but more time on other subjects. |
| 7 | + | ++ | We'd better spend time on specialised subjects than on GE. I study 'general Non-English subjects' to make up for the low mark on GE. |
| 8 | + | ++ | The weight of GE is much lower than those of the specialised subjects.<br>I study a general Non-English subject most students are weak at even if it has low weight. |
| 9 | + | ++ | GE weight is much lower than specialised subjects e.g. mathematics 5, English 2<br>I study a general Non-English subject students are weak at. |

**Note:** E= English, N = Non-English subjects, ++ = more time was spent

All the GE students unanimously said that they studied Non-English subjects more than English because of the heavier weight of the Non-English subjects i.e. they reasoned based on their knowledge of the test. However, they gave reasons other than the weight of the subjects as well. GE 1 made a connection between her emphasis on Specialised, i.e. Non-English subjects and her purpose of test preparation that the emphasis was because she wanted to study a non-English major at university. GE 2 said she spent more time on Non-English subjects because she was weak in English and had little interest in it. GE 3, however, said because of her interest in English she spent more time on GE than on other General subjects but less time than on Specialised subjects. She said another reason was that her preparation class teacher gave them a test on GE every week. GE 4 said studying GE had some advantages in

itself, i.e. it put him in a better position to get ahead of his peers and gave him a better chance for employment. When comparing GE and other General courses, GE 5 said it was just a matter of time not a matter of emphasis of one over the other. She said she studied all the General subjects in turn. The reason why GE 7 said he did not emphasise GE was that the GE mark could be compensated for by other General courses. Comparing GE and other General courses, both GE 8 and 9 said what mattered was which subjects the rivals were weak at so that they could learn those subjects more.

Like the SPE Group, all the GE students were affected by the test as well as other factors. Most of the reasons stated by the GE Group were similar to those of the SPE Group. The common reasons for emphasising one subject over the other included the reason based on their 'knowledge of the test', 'effect of the teacher', 'weakness in a subject', 'weakness of rivals', 'point of time in the preparation period', and 'compensation for inadequate knowledge in a subject'. The different reasons included 'effect of friends' and 'accessibility of learning materials' which were stated by the SPE Group and 'purpose of test preparation' and 'interest' stated by the GE Group. In sum, the test authorities' decision to give more weight to English had positive washback in that more SPE students focused on English than on non-English subjects. However, there were still quite a few SPE students who emphasised non-English subjects. Although the students' views might change later in their preparation, this may also suggest reconsideration of the weighting by the test authorities.

## 8.3. RQ4. Do SPE students have better English backgrounds than GE students?

I measured the quality of English background by the length of attendance in English institutes. In the sample of the interviewees, the SPE students had 13.20 terms

of attendance and the GE students 7.22 terms on average, which suggested that the SPE students had a better English background than the GE students. This was confirmed by the results of both the questionnaire and the letters (Tables 6.26 and 7.4).

## 8.4. RQ6. Are the SPE students' reported activities consistent with their reported beliefs?

This section is concerned with consistency between the students' beliefs and their activities for test preparation or, in other words, with whether the test interacted with the students' learning beliefs.

As mentioned before, since the students were close to end-of-term examinations, I had only limited time to interview them. Therefore, I could probe only a limited number of belief items, which I will report in the following sections. They include beliefs about language skills and guessing the meanings of new words.

### 8.4.1. Reported beliefs about language skills

This section is concerned with consistency between the students' beliefs about vocabulary, grammar, and the use of tapes and CDs and the corresponding activities. After I listened to the students' reasons for doing grammar, vocabulary, and using tapes and CDs for their test preparation (8.2.1 and 8.2.2), I asked them whether doing these skills were necessary for language learning in general as well. In other words, I probed the three questionnaire items which dealt with the students' beliefs about grammar, vocabulary, and the use of tapes and CDs. The results are presented in Tables 8.23 and 8.24. In examining the consistency, I will ignore differences in the rankings of grammar and vocabulary activities but only whether they were ranked (i.e. done).

**Table 8-23 SPE students' consistency between reported beliefs and reported activities of vocabulary, grammar and the use of tapes and CDs**

| | Beliefs | | | | Activity | | |
|---|---|---|---|---|---|---|---|
| | Voc | Gr | Tp,CD | Reasons | Voc | Gr | Tp,CD |
| 1 | + | + | + | All skills are important | 2nd | 1st | + |
| 2 | + | - | + | Vocabulary 1st, also other skills but not sure about grammar | 1st | 3rd | + |
| 3 | + | + | + | All skills but grammar not very important | 1st | 3rd | - |
| 4 | + | - | + | All skills, except grammar, like environment or institute | 1st | 3rd | + |
| 5 | + | + | + | Vocabulary more important than grammar, speaking most important but grammar less important | 1st | 2nd | - |
| 6 | + | + | + | Vocabulary 1st, grammar next, speaking for using speaking skills and getting used to it, grammar for speaking correctly | 1st | 4th | - |
| 7 | + | + | + | All skills e.g. vocabulary, grammar, listening, film, translation, etc | 1st | 2nd | - |
| 8 | + | + | + | All skills, but vocabulary more important than grammar, translation, listening, speaking | 1st | 2nd | - |
| 9 | + | + | + | All skills, but vocabulary and translation 1st, speaking, but grammar not very important | 1st | 2nd | - |

**Note:** Voc= Vocabulary, Gr= Grammar, Tp,CD= Tape & CD, += belief in a skill or doing the skill, -= lack of belief in a skill or not doing the skill

Table 8.23 shows that there was consistency between the students' belief about vocabulary learning and their activity of vocabulary learning. The students' reasons' in the 'Reasons' column show that all SPE students believed in vocabulary as one of the essential skills in language learning. They also rated 1st or 2nd the amount of time they spent on vocabulary. In other words, all the SPE students both believed in vocabulary learning as well as did learn vocabulary for their test preparation. However, concerning grammar, 2 students' beliefs were affected. While SPE 2 was not sure and SPE 4 did not believe that grammar was primary in language learning, they studied grammar as their test preparation activities. For the 7 other students, however, there was consistency between their belief in grammar and their activity of grammar. The effect of the test on oral skills, in this case using tapes and CDs, was almost opposite to the effect of the test on grammar, i.e. 6 students (SPE 3, 5, 6, 7, 8, and 9) believed in the use of tapes and CDs but did not use them in their test preparation. There were three others (SPE 1, 2, and 4) who both believed in the value

of tapes and CDs and used them in their test preparation. In sum, while the SPE students were consistent in terms of vocabulary, 2 of them were inconsistent in grammar and 6 of them were inconsistent in the use of tapes and CDs.

Concerning grammar, the effect of the test can be considered as positive in that the two students who did not believe in grammar did grammar learning activities for test preparation. However, the effect can be considered negative in the case of tapes and CDs as 6 students who believed in the use of tapes and CDs did not use them for test preparation.

The results agree with the questionnaire results for vocabulary (Table 6.38) but not completely for grammar where in both cases there were consistencies between belief and activity (Table 6.39). The results also agree concerning the use of tapes and CDs where there were inconsistencies between the belief and activity (Table 6.41).

**Table 8-24 GE students' consistency between belief and activity of vocabulary, grammar and the use of tapes and CDs**

| | Beliefs | | | | Activity | | |
|---|---|---|---|---|---|---|---|
| | Voc | Gr | Tp,CD | Reasons | Voc | Gr | Tp,CD |
| 1 | + | - | + | All skills: vocabulary, speaking, film, magazines, like environment, not grammar | 1st | 3rd | + |
| 2 | + | + | + | All skills are important | 1st | 3rd | - |
| 3 | + | + | + | All skills are important | 1st | 3rd | - |
| 4 | + | + | + | Vocabulary 1st, grammar, translation, speaking, like environment | 1st | 2nd | - |
| 5 | + | - | + | All skills, vocabulary, speaking, etc but grammar not very important | 1st | 3rd | - |
| 6 | + | + | + | vocabulary, grammar, translation, reading, newspaper, magazines, film, programmes | 1st | 2nd | + |
| 7 | + | + | + | vocabulary 1st, translation, speaking, grammar | 1st | 2nd | - |
| 8 | + | + | + | Vocabulary 1st, sentence structure, newspapers, TV programmes, speaking | 1st | 2nd | - |
| 9 | + | + | + | All skills, vocabulary, grammar, speaking, as in environment | 1st | 2nd | - |

**Note:** Voc= Vocabulary, Gr= Grammar, Tp,CD= Tape & CD, += belief in a skill or doing the skill, -= lack of belief in a skill or not doing the skill

The consistency of the GE students in terms of vocabulary and grammar was the same as that for the SPE students, i.e. there was consistency between vocabulary beliefs and activity of all the GE students but inconsistency for two students in

grammar. The results for the oral skills were opposite to the grammar results, i.e. there was inconsistency between the belief and the activity for 7 students. Although these students believed in the usefulness of tapes and CDs, they did not use them for test preparation.

The results do not agree with the questionnaire results for vocabulary (Table 6.32) but almost agree for grammar (Table 6.33). However, the results agree concerning the use of tapes and CDs where there were inconsistencies between the belief and activity (Table 6.35).

In sum, the consistency of GE and SPE students in terms of vocabulary, grammar, and the use of tapes and CDs were similar. There were no inconsistencies for vocabulary but two cases of inconsistency for grammar and inconsistencies for the majority of the students in both groups (6 SPE and 7 GE) for the use of tapes and CDs. Both tests had no effect on vocabulary belief but affected two students positively in terms of grammar and the majority of the students negatively in terms of tapes and CDs. The effects on grammar belief were considered positive as the two tests encouraged the two students to practice a skill which was tested in the two tests. However, the effects of the two tests on the oral skill were considered negative as the majority of the students were discouraged to do orals despite the fact that they believed they were useful.

### 8.4.2. Reported beliefs about guessing

This section discusses whether the students' belief about the value of guessing was affected by the test. As I was asking the students why they used guessing in their reading activity, at the same time I asked them whether the reasons they gave applied to the situation where there was no exam as well. The reasons the students gave for the two situations are presented in Table 8.25. The reasons specific to test preparation

are in bold.

**Table 8-25 SPE students' consistency between the belief and activity of guessing**

| | Belief | Reasons for & against guessing as a belief and as an activity | Activity |
|---|---|---|---|
| **1** | + | Unknown words in texts | + |
| **2** | + | Not possible to learn all words | + |
| **3** | + | Unknown words in texts | + |
| **4** | + | Not possible to learn all words | + |
| **5** | + | Not possible to learn all words. **Dictionary not allowed in exam** | + |
| **6** | + | Unknown words in texts | + |
| **7** | + | Useful for comprehension | + |
| **8** | + | Unknown words in texts | + |
| | - | Not suitable for difficult words | - |
| | - | Not suitable for key words | - |
| **9** | + | Guessing is for students with good English but I use dictionary because weak | - |

Students gave various reasons for and against the value of guessing. Among SPE students, four of them (1, 3, 6, 8) argued that they needed to develop their guessing skill because there was always the possibility that there might be a word or words in a text which would be unknown to them. SPE 2, 4, and 5 reasoned that guessing was useful because it was not possible to learn all the words. SPE 5 said they needed to develop their guessing skill because they were not allowed to use a dictionary in the exam. SPE 7 argued 'guessing is important for understanding the meaning of a sentence because there might be a word in the sentence which has an important role'. However, two SPE students gave reasons against guessing. One of these students was SPE 8 who gave a reason in favour of guessing as well. She stated that on some occasions she did not guess including when the word was very difficult and when the word was important in conveying the meaning of the sentence, in which case she referred to the dictionary. The other student was SPE 9. She approved of guessing only for students with good English and said that she did not use guessing at all and used the dictionary instead because she was weak in English. In sum, except for SPE 5 who gave an additional reason for exam preparation, the other students said

their belief about guessing applied to both exam and non-exam situations i.e. the majority of the students' beliefs were stable. Table 8.26 presents the results for the GE Group:

**Table 8-26 GE students' consistency between belief and activity of guessing**

| | Belief | Reasons for & against guessing as a belief and as an activity | Activity |
|---|---|---|---|
| **1** | + | Not possible to learn all words | + |
| **2** | + | Teacher taught us guessing: general meaning and affixes help in guessing. | + |
| **3** | + | No need to look up the new word if you have got the general meaning of the passage | + |
| **4** | + | Results in longer retention of words in mind | + |
| | - /+ | Sometimes use dictionary when difficult to guess | -/+ |
| **5** | + | Learning words better. **Exact meaning is important for the exam, so dictionary** | - |
| **6** | + | Not possible to learn all words, useful for translation | + |
| **7** | - | Guessing only when no access to someone or a dictionary | - |
| | - | Learning wrong meanings if wrong guesses | - |
| **8** | + | Guessing is good when exact meaning is not important. **I use dictionary for exact meanings as choices are close.** | - |
| **9** | + | Guessing is good because not always access to dictionary. Difficult to guess due to weakness, so use dictionary. | - |

GE 1 and 6, similar to some SPE students above, considered guessing useful because they said it was not possible to learn all the words. GE 6 also said guessing helped translation of a passage. GE 2 talked about the source of his belief (teacher) through suffixes and prefixes and general meaning of the passage. In this case, the effect of the GE Test on the belief as well as the activity was mediated by the teacher. GE 3 also said there was no need for a dictionary if she had a general meaning of the passage in mind. GE 4 thought that guessing resulted in keeping words in mind longer than when he looked up a word in a dictionary and then memorised it, in which case it would soon be forgotten. However, when he found it difficult to guess, he used a dictionary. Similarly, GE 5 stated that if one guessed the meaning of a new word, they learned it better. However, she had a different view for the exam. She said the use of dictionary was necessary as knowing the exact meaning of words was crucial in the exam. Although she did not show any concern for accuracy in her belief, she was concerned about it in her test preparation, which indicates that her belief was affected

negatively by the GE Test. GE 7 neither believed in guessing nor used guessing when reading a passage as he confined guessing to only when there was no access to someone or a dictionary and said that when there was access he looked up the word (and then memorised it). He said he looked up the word in case his guesses might be wrong, in which case it would be difficult to get rid of them from the memory. He was concerned about accuracy both in his belief and in his activity. GE 8 believed in guessing provided that the exact meaning was not important. However, he said the exact meanings of words were important for the exam as choices in the exam were close to each other. Therefore, he did not use guessing in his test preparation, which is another instance of the negative effect of the GE Test on belief. Like GE 5, he was not concerned about accuracy in his belief but in his test preparation. GE 9 found it difficult to guess because of his weakness and used dictionaries, although he generally believed in guessing on the grounds that dictionaries were not always available. In this case, since the inconsistency between the belief and the activity is not due to the test but due to the weakness in language ability, the negative effect on the belief cannot be attributed to the test.

In sum, most of the reasons the students gave about guessing concerned both exam and non-exam situations. In other words, there were largely consistencies between beliefs and activities in both groups, which means that their beliefs were affected similarly. However, there were two inconsistencies in the GE Group (GE 5 and 8) and one inconsistency in the SPE Group which were considered the negative effects of the tests.

The analysis of beliefs in 8.4.1 and 8.4.2 showed that the stability of beliefs varied. The resuls showed that the belief about vocabulary was the most stable, and the belief about the use of tapes and CDs was the least stable, with beliefs about

grammar and guessing in between. The stability also slightly varied from one group to another.

A summary of the whole beliefs results are presented in Tables 8.28 and 8.29.

**Table 8-27 The number of the interviewees who reported beliefs in the importance of vocabulary, grammar, use of tapes and CDs, and guessing skill**

| | | |
|---|---|---|
| The most important part of learning a foreign language is learning vocabulary words. | 9 | 9 |
| The most important part of learning a foreign language is learning the grammar. | 7 | 7 |
| It's o.k. to guess if you don't know a word in English. | 9 | 8 |
| It is important to practice with tapes and CDs. | 9 | 9 |

**Table 8-28 Consistency between reported beliefs and reported activities**

| Themes | SPE | | GE | |
|---|---|---|---|---|
| | Belief | Activity | Belief | Activity |
| **Vocabulary** | Consistent | | Consistent | |
| **Grammar** | Lower (7) | Higher (9) (incons for 2) | Lower (7) | Higher (9) (incons for 2) |
| **Using tapes & CDs** | Higher (9) | Lower (3) (incons for 6) | Higher (8) | Lower (2) (incons for 7) |
| **Guessing** | Higher (9) | Lower (8) (incons for 1) | Higher (8) | Lower (5) (incons for 2) |

**Note:** incons= inconsistency

## 8.5. Summary

The interviews confirmed that the SPE Test showed washback in the areas where changes were introduced by the test authorities including increase in the weight of English, use of extra materials, and attraction of more proficient applicants to study English at university. The interview also confirmed that the newly-added sections had the least washback.

In the analysis of the effect of the test on beliefs, first I compared the individual student's beliefs with his/her activities and then I contrasted the SPE with the GE group. This type of contrast, i.e. between individual students' beliefs and activities might be useful for situations where there is no appropriate contrast group or a baseline study or a situation in which unintended washback is going to be studied.

The analysis of beliefs showed that GE and SPE students' beliefs were affected similarly. Neither test affected vocabulary belief but affected two students

positively in terms of grammar and the majority of the students negatively in terms of the use of tapes and CDs. As regards the guessing belief, 1 SPE student and 2 GE students were affected negatively. The results suggested that some beliefs were less affected by the test, i.e. were more stable than others. The stability also varied from one group to another.

# Chapter 9. Summary of the Findings from All the Instruments

In the last three chapters, I presented the results from each instrument individually. However, in this chapter I will summarise the results from all the three instruments and will present the main findings based on the predicted and unpredicted washback and the research questions.

## 9.1. 'Predicted washback' addressed by RQs 1, 3, and 4

In this section, I present a summary of the combined results from the three instruments concerning the 'predicted washback' addressed by RQ 1 (What activities do the SPE students report doing in order to prepare for the SPE Test?), RQ3 (Do SPE students perceive the SPE Test to be more difficult than the GE Test?), and RQ4 (Do SPE students have better English backgrounds than GE students?). The summary of the results is presented in Table 9.1.

**Table 9-1 Summary of the results: predicted washback (reported activities, test difficulty, English ability**

| **Predicted washback** | **Questionnaire** | | **Letters** | | **Interview** | |
|---|---|---|---|---|---|---|
| **Activities- weight** | **SPE** | **GE** | **SPE** | **GE** | **SPE** | **GE** |
| Vocabulary | Longer | Shorter | No difference | | No difference | |
| Grammar | Longer | Shorter | No difference | | No difference | |
| Reading | Longer | Shorter | No difference | | No difference | |
| Reading | Contradictory | | NA | | NA | |
| Cloze | Longer | Shorter | No difference | | No difference | |
| Sentence Structure | No difference | | Not Mentioned | | No difference | |
| Language Functions | No difference | | Not Mentioned | | No difference | |
| Weight of English | Longer | Shorter | Not Mentioned | | Longer | Shorter |
| Weight of Non-English | *Shorter* | *Longer* | *Shorter* | *Longer* | *Shorter* | *Longer* |
| **Activities- school** | | | | | | |
| Pre-university book | Shorter | Longer | No difference | | No difference | |
| **Activities- out of school** | | | | | | |
| Extra materials | Longer | Shorter | Longer | Shorter | Longer | Shorter |
| *Bridging the Gap* | No difference | | No difference | | Longer | Shorter |
| Preparation class | Longer | Shorter | No difference | | Longer | Shorter |
| Classes in English institutes | ----- | ----- | No difference | | No difference | |
| **Test difficulty** | | | | | | |
| Test Difficulty | Higher | Lower | Higher* | Lower* | ----- | ----- |
| **English background** | | | | | | |
| Total attendance in English institutes | Longer | Shorter | Longer | Shorter | Longer | Shorter |
| Self-assessment | Higher | Lower | Not Mentioned | | ----- | ----- |

**Note 1:** The results in the 'Interviews' column do not concern 'reasons' but whether the students reported doing or not doing the activities. **Note 2:** Where GE is 'higher' the result is in italics. **Note 3:** *

indicates that only 10 SPE students compared SPE and GE tests.

The questionnaire results showed that SPE students were statistically different from GE students in all the areas of predicted washback except the new sections and use of *Bridging the Gap*. They reported spending longer on the sections which the SPE Test had in common with the GE Test and on out-of-school resources including extra materials and preparation classes and spent longer on English than on non-English subjects. SPE students also considered the SPE Test more difficult than the GE Test (RQ 3) and had better English background than GE students (RQ 4). It was speculated that the lack of differences in the use of *Bridging the Gap* and the time spent on the new sections might be because of factors other than the test. For example, because questionnaire data was collected at the beginning of the test preparation period, the students might have not yet started with them. Concerning *Bridging the Gap*, this speculation is confirmed by the results of the interview data which was collected at the end of the data collection period. However, the new sections still showed non-significant results in the letters and the interviews, which probably indicate that the new sections had no washback. The results also showed that three activities, namely, weight of English versus non-English, use of extra materials, and the requirement of stronger English background, were affected by washback most of all as they showed significant differences between SPE Group and GE Group on all the three instruments. This is what Cheng (1997) refers to as 'washback intensity', i.e. 'the degree of washback effect in an area or a number of areas of teaching and learning affected by an examination' (p. 43). The next most 'intense' washback was on attendance in preparation classes and the perception of test difficulty, where the results were statistically significant in two instruments. Based on these results, I have reasonable confidence in claiming that the SPE Test had washback in these five areas.

The results also indicate that the letters and interviews showed fewer statistically significant results than the questionnaire. These inconsistencies could be due to the changes in the students' attitudes during the data collection period, or because the sample of letter writers and the interviewees was much smaller than the questionnaire respondents.

## 9.2. 'Unpredicted washback' addressed by RQ1

The majority of the results from the questionnaire showed statistically significant differences between the SPE and the GE groups. However, in the letters the issues were either not mentioned or showed no significant differences between the two groups. Similarly, the interviews showed no differences between the groups. However, as far as the questionnaire results are concerned, the results showed that, in general, the SPE students reported spending longer on all the activities with the exception of memorisation and the use of computer games, where the two groups were not statistically different. However, although the SPE students generally reported spending longer on the activities, the frequency with which they reported doing these activities varied from one activity to another. They also did some activities which were 'traditional' including translation and memorisation. However, it was not clear from the results why the students did these activities or why they did them with varying frequencies. Nevertheless, it was speculated that some possible causes might be the test, interest, access, beliefs, the textbook, the teacher, and/or English background. The reasons for the students' reported activities were identified in order to address RQ2 (9.3).

**Table 9-2 Summary of the results for unpredicted reported activities**

| Unpredicted washback | Questionnaire | | Letters | | Interview | |
|---|---|---|---|---|---|---|
| | SPE | GE | SPE | GE | SPE | GE |
| Memorising words | No difference | | No difference | | ----- | ----- |
| Making sentences for new words | Longer | Shorter | No difference | | ----- | ----- |
| English to English dictionary | Longer | Shorter | No difference | | ----- | ----- |
| Making sentences for grammar | Longer | Shorter | No difference | | ----- | ----- |
| Memorising grammar rules | No difference | | No difference | | No difference | |
| Using texts for grammar | Longer | Shorter | No difference | | No difference | |
| Reading newspapers & magazines | Longer | Shorter | No difference | | No difference | |
| Translating reading texts | Longer | Shorter | No difference | | No difference | |
| Guessing meaning of new words | Longer | Shorter | No difference | | No difference | |
| Reading story books | Longer | Shorter | No difference | | No difference | |
| Attention to topic of texts | Longer | Shorter | Not Mentioned | | No difference | |
| Skimming | Longer | Shorter | Not Mentioned | | No difference | |
| Reading various texts | Longer | Shorter | Not Mentioned | | ----- | ----- |
| Summarising reading texts | Longer | Shorter | Not Mentioned | | ----- | ----- |
| Writing diaries | Longer | Shorter | Not Mentioned | | No difference | |
| Using the Internet | Longer | Shorter | Not Mentioned | | ----- | ----- |
| Using emails | Longer | Shorter | Not Mentioned | | ----- | ----- |
| Listening to the radio | Longer | Shorter | Not Mentioned | | No difference | |
| Looking for people for conversation | Longer | Shorter | Not Mentioned | | ----- | ----- |
| Using tapes and CDs | Longer | Shorter | No difference | | No difference | |
| Watching films & TV programmes | Longer | Shorter | No difference | | No difference | |
| Using computer games | No difference | | Not Mentioned | | ----- | ----- |
| Using English despite errors | Longer | Shorter | Not Mentioned | | ----- | ----- |
| Taking mock exams | Longer | Shorter | No difference | | ----- | ----- |
| Using practice tests | Longer | Shorter | No difference | | No difference | |
| Review & practice | Longer | Shorter | No difference | | ----- | ----- |

----- = not analysed

## 9.3. Why do the students report doing these activities in order to prepare for the SPE Test? (RQ2)

I addressed students' knowledge of the test, their reasons for test preparation, motivation, interest, and anxiety in the questionnnaire in order to account for their reported activities. I also looked to see what reasons the students gave in the letters and asked them about their reasons for each activity in the interviews.

### 9.3.1. Possible reasons for reported activities based on the questionnaire and letters

Table 9.3 presents the results from the questionnaire and the letters. Since in the interviews (reported) 'causes' were examined for 'predicted activities' only, they are not reported in this table but will follow the discussion of the results from the questionnaire and the letters.

The SPE and GE students' total scores on test knowledge shows that they were aware of the main requirements of the test including weight of each section, weight of English, and the broader curriculum. Since SPE students were statistically different in the corresponding activities as well, I had a basis on which to speculate that the related activities which they reported, particularly the predicted areas of washback, could have been due to the test.

As far as the results from the letters are concerned, most suggestions about test knowledge centred around vocabulary and grammar as the most important skills and the use of extra materials versus the Pre-university book. Very few students, however, commented on other aspects of test knowledge such as cloze, and reading.

Concerning reasons for test preparation, 'interest', 'passing the exam' and 'future job' were the first three most important reasons given in response to the questionnaire, all three of which could have motivated the students to try harder (see also Table 6.20). The letter data also was useful in identifying some reasons for test preparation (see also Table 7.7) including 'passing the exam', 'future use of English', 'English as the key to international communication', and 'job opportunities'. However, the students were only significantly different in one purpose, 'passing the exam'.

**Table 9-3 Summary of the results- possible (reported) causes of students' activities**

| Possible causes | Questionnaire | | Letters | |
|---|---|---|---|---|
| **Knowledge of the test** | **SPE** | **GE** | **SPE** | **GE** |
| 1. A section in the SPE Test includes listening and speaking. | .80 | .77 | | |
| 2. Each section of the SPE Test has the same number of questions. | .72 | .71 | | |
| 3. SPE Test has six sections. | .80* | .80* | | |
| 4. Vocabulary and Reading have the same weight. | .78 | .76 | | |
| 5. The total number of questions in the SPE Test is 70. | .56* | .78* | | |
| 6. The majority of the SPE questions come from the Pre-university book. | .72* | .75* | | |
| 7. Marks allocated to English are more than non-English subjects. | .84* | .83* | | |
| 8. Each question has the same mark as the other, regardless of which section it belongs to | .75 | .72 | | |
| **Total score on knowledge of the test** | **.75 of 1** | **.77 of 1** | | |
| 1.Vocabulary is the most important | | | 20 | 17 |
| 2.Grammar is the second most important | | | 12 | 7 |
| 3.Grammar is the third most important | | | 2 | 2 |
| 4.Reading is the second most important | | | 2 | 5 |
| 5.Reading is the third most important | | | 2 | 0 |
| 6.Cloze is the second most important | | | 0 | 2 |
| 7.There is no listening section | | | 0 | 2 |
| 8.Questions come from the Pre-university book | | | 2 | 0 |
| 9.Questions come from out of the Pre-university book | | | 0* | 12* |
| 10.Prior knowledge is necessary | | | 2 | 5 |
| Total number of students | | | 42 | 52 |
| **Purposes** | | | | |
| English sources | No difference | | Not mentioned | |
| Job | *Lower* | *Higher* | No difference | |
| Passing the exam | *Lower* | *Higher* | Higher | Lower |
| Communication | No difference | | No difference | |
| Travelling | *Lower* | *Higher* | Not mentioned | |
| Cultural products | No difference | | Not mentioned | |
| Interest | Higher | Lower | Not mentioned | |
| Living in an English-Speaking country | No difference | | Not mentioned | |
| Future use | No difference | | No difference | |
| **Motivation and interest** | | | | |
| Importance of exam success | Higher | Lower | ---------- | |
| Trying/making effort | Higher | Lower | ---------- | |
| Interest | Higher | Lower | ---------- | |
| Lack of interest | *Lower* | *Higher* | ---------- | |
| Level of motivation | Higher | Lower | ---------- | |
| **Anxiety** | | | | |
| Stress | No difference | | ---------- | |
| Afraid of bad marks | No difference | | ---------- | |

Asterisks in 'Test Knowledge' section indicate statistical differences between the groups.
----------= not analysed

Regarding motivation and interest, the questionnaire results showed that the SPE students were more motivated and more interested than the GE students. However, the two groups were not different in anxiety. I suggested that the results

could be probed further to explore whether there was a (reported) cause-effect relationship between the students' activities and their test knowledge and how the test was affecting or interacting with the students' purposes, motivation, interest, and anxiety. However, statements about motivation and interest in the letters were not analysable as they were very general.

### 9.3.2. Reasons for activities based on the interviews

Concerning the amount of time they spent on the test sections, both groups were affected by the test as well as by factors other than the test. Among the test factors, both groups were affected by the weight of the test sections, the difficulty level of the sections, and time management in test taking. The weight of the sections had the most effect on both groups. However, while SPE students spent time on the sections depending on the predictability of the questions and the degree to which they were familiar with the questions, GE students were affected by a desire for a good mark and whether spending time on a section made it easier to answer the questions.

Among factors other than the test, both groups were affected by their views of the interdependence of the knowledge of the test sections or constructs, emphasis of a section by the teacher or school, their English ability, and learning materials. The two groups were affected most by their views of the test constructs. There were reasons which were unique to each group as well. While some SPE students were influenced by whether they liked the skill in general, by weakness in a skill/section, attending English institutes, point of time in the preparation period, and their view of a single construct, one GE student was influenced by the availability of resources.

In terms of oral activities, both groups of students were affected by the test as well as by non-test factors. SPE students who reported doing oral activities did them mainly for grammar and vocabulary. However, they also reported doing them for

improving their English in general, for English institutes, self-confidence, and because of the preference for oral activities over written activities. The SPE students who reported not doing oral activities, reasoned that there was no oral component in the test and that it mainly comprised vocabulary and grammar. The reasons not related to the test included lack of interest in certain audio devices (radio), problems of access, difficulty of oral language, and lack of time.

The two GE students who reported doing oral activities had reasons different from each other. While one of them reported doing the activities for the test (grammar and vocabulary) as well as purposes not related to the test (interest and 'English is the international language'), the other did them not for the test but for factors other than the test, which included interest, ability improvement in English in general, and desire to speak to someone in English in future. The GE students who reported not doing oral activities reasoned that although they could be useful, they were not essential and that oral skills were more difficult than what GE Test required. The non-test reasons they used to justify their not doing oral activities included time limitations, lack of access, weakness, inability to say things accurately, and not being asked by the teacher.

The students' spending time on the test sections was also mediated by the teacher and the textbooks. There was a consensus among both SPE and GE students that the only skills emphasised by teachers and materials were vocabulary, grammar and reading, whether in school class or preparation class (see also the textbook descriptions).

As for the use of extra materials, SPE students were affected by the requirement of SPE Test for extra materials as well as by non-test factors. The non-test factors included the class or teacher, English institutes, whether the learning materials were interesting, and of appropriate difficulty, whether they were updated,

inadequacy of school materials and classes, and finally usefulness of extra materials for future, i.e. after admission to university. However, the only one GE student who used extra materials did not use them for the test but for her own interest.

In terms of whether to emphasise English or non-English subjects, the reasons which the two groups stated in common included those based on their 'knowledge of the test', 'effect of the teacher', 'weakness in a subject', 'weakness of rivals', 'point of time in the preparation period', and 'compensation for inadequate knowledge in a subject'. The different reasons which the SPE Group stated included 'effect of friends' and 'accessibility of learning materials' and the different reasons which the GE Group stated included 'purpose of test preparation' and 'interest'.

## 9.4. What beliefs do the students report holding about learning? (RQ5)

Since the SPE test preparation was done partly independently, I decided to examine the students' learning beliefs as they might be crucial in this independent learning situation. Table 9.4 briefly shows what types of beliefs the students had based on the three instruments.

**Table 9-4 SPE and GE students' reported language learning beliefs**

| | Questionnaire | | Letters | | Interviews (out of 9[1]) | |
|---|---|---|---|---|---|---|
| **Beliefs in the questionnaire** | **SPE** | **GE** | **SPE** | **GE** | **SPE** | **GE** |
| **Foreign language aptitude** | | | | | | |
| 9. I have a special ability for learning foreign languages. | Higher | Lower | | | | |
| **Difficulty of language learning** | | | | | | |
| 2. I believe that I will learn to speak English very well. | Higher | Lower | | | | |
| 28. The English language is a) very difficult b) difficult c) of medium difficulty d) easy e) very easy | *Lower* | *Higher* | | | | |
| **Nature of language learning** | | | | | | |
| 4. It is necessary to know about English-speaking cultures in order to speak English. | Higher | Lower | | | | |
| 10. The most important part of learning a foreign language is learning vocabulary words. | *Lower* | *Higher* | | | 9 | 9 |
| 15. The most important part of learning a foreign language is learning the grammar. | Higher | Lower | | | 7 | 7 |
| 18. Learning a foreign language is different than learning other academic subjects. | No difference | | | | | |
| 19. The most important part of learning English is learning how to translate into Farsi. | *Lower* | *Higher* | | | | |
| 23. Language learning involves a lot of memorization. | No difference | | | | | |
| **Learning and Communication Strategies** | | | | | | |
| 3. It is important to speak English with an excellent pronunciation. | Higher | Lower | | | | |
| 7. If I meet English speakers, I will enjoy practicing English with them. | Higher | Lower | | | | |
| 8. It's o.k. to guess if you don't know a word in English. | Higher | Lower | | | 9[2] | 8 |
| 11. It is important to repeat and practice a lot. | Higher | Lower | | | | |
| 13. I feel timid speaking English with other people. | *Lower* | *Higher* | | | | |
| 17. It is important to practice with tapes and CDs. | Higher | Lower | | | 9 | 9 |
| **Motivations and expectations** | | | | | | |
| 12. People in my country feel that it is important to speak English. | No difference | | | | | |
| 16. I would like to learn English so that I can get to know English speakers better. | No difference | | | | | |
| 20. If I learn English very well, I will have better opportunities for a good job. | Higher | Lower | | | | |
| 21. I want to learn to speak English well. | Higher | Lower | | | | |
| 22. I would like to have English-speaking friends. | Higher | Lower | | | | |
| **Relative usefulness of oral versus written skills** | | | | | | |
| 24. Speaking and listening to English are more useful than reading and writing English. | Higher | Lower | | | | |
| **Role of Time** | | | | | | |
| 25. Language learning takes a long time. | No difference | | | | | |
| **Use of language** | | | | | | |
| 26. It is important to find as many ways as possible to use English. | Higher | Lower | | | | |
| **Beliefs elicited in the letters** | | | | | | |
| 1.Language means arranging 'words' for communication | | | No difference | | | |
| 2. Vocabulary facilitates reading | | | No difference | | | |
| 3. Vocabulary and grammar complement each other | | | No difference | | | |
| 4. Vocabulary improves speaking | | | No difference | | | |
| 5. Speaking improves vocabulary as it provides for vocabulary practice | | | No difference | | | |
| 6. Language learning is gradual and takes time[3] | | | No difference | | | |
| 7. English is elusive | | | No difference | | | |

**Note 1:** Numbers in the 'Interviews Column' refer to the number of interviewees in each group, **Note 2:** One in each group did not completely believe in guessing, **Note 3:** This was the only belief which was close to the belief addressed in the questionnaire (Item 25).

In general SPE students more than GE students, but both groups had a number

of traditional as well as contemporary beliefs. The SPE Group was more concerned about pronunciation than the GE Group, which might act as an obstacle to using language for communication. However, they were less timid than GE students, which might be an advantage for communication. While the majority from both groups believed in the importance of vocabulary learning, only a minority considered grammar learning important. Both groups believed in translation, but the SPE Group endorsed this belief less than the GE Group. Both groups also believed in memorisation but there was no statistical difference between them. The question which arose was how the students' beliefs would interact with the students' test preparation activities. In other words, the question was whether there were consistencies or inconsistencies between the students' reported beliefs and reported activities.

The letters were useful in eliciting beliefs; these were all different from those of the questionnaire, except one. This suggested that letters were a useful instrument for examining beliefs as well. Most of the beliefs elicited concerned vocabulary, which confirmed that the students believed vocabulary learning was important. However, based on the contextual approach to beliefs (3.4.3), unless one talked to the students, it was not clear whether the beliefs originated from test preparation experience or whether they were held in advance. I did not examine the interaction between the beliefs elicited in the letters with the test but rather with some of the beliefs in the questionnaire.

## 9.5. Are the SPE students' reported activities consistent with their reported beliefs? (RQ6)

After I examined what types of beliefs the students had, I decided to examine how they were interacting with the test. To this end, first I compared the individual

group's beliefs with their activities in the case of the questionnaires and individual students' beliefs with their activities in the case of the interviews. Then, I contrasted the SPE with the GE group. Table 9.5 summarises the consistency between beliefs and activities and shows which beliefs were consistent and which beliefs were inconsistent.

**Table 9-5 Consistency between reported beliefs and reported activities**

| Themes | Group | Questionnaire | | Interviews | |
|---|---|---|---|---|---|
| | | Belief | Activity | Belief | Activity |
| **Vocabulary** | **SPE** | Consistent | | Consistent | |
| | **GE** | Higher | Lower | Consistent | |
| **Grammar** | **SPE** | Consistent | | Lower (7) | Higher (9) (incons for 2) |
| | **GE** | Consistent | | Lower (7) | Higher (9) (incons for 2) |
| **Using tapes & CDs** | **SPE** | Higher | Lower | Higher (9) | Lower (3) (incons for 6) |
| | **GE** | Higher | Lower | Higher (8) | Lower (2) (incons for 7) |
| **Guessing** | **SPE** | Higher | Lower | Higher (9) | Lower (8) (incons for 1) |
| | **GE** | Higher | Lower | Higher (8) | Lower (5) (incons for 2) |
| **Translation** | **SPE** | Lower | Higher | | |
| | **GE** | Higher | Lower | | |
| **Memorisation- vocabulary** | **SPE** | Lower | Higher | | |
| | **GE** | Lower | Higher | | |
| **Memorisation- grammar** | **SPE** | Higher | Lower | | |
| | **GE** | Higher | Lower | | |
| **Belief in use of English VS:** | | | | | |
| Newspapers and magazines | **SPE** | Higher | Lower | | |
| | **GE** | Higher | Lower | | |
| Diaries, notes | **SPE** | Higher | Lower | | |
| | **GE** | Higher | Lower | | |
| Films or programmes | **SPE** | Higher | Lower | | |
| | **GE** | Higher | Lower | | |
| Radio | **SPE** | Higher | Lower | | |
| | **GE** | Higher | Lower | | |
| Story books | **SPE** | Higher | Lower | | |
| | **GE** | Higher | Lower | | |
| Read various texts | **SPE** | Higher | Lower | | |
| | **GE** | Higher | Lower | | |
| Emails | **SPE** | Higher | Lower | | |
| | **GE** | Higher | Lower | | |
| **Belief in translation VS:** | | | | | |
| Topic of texts | **SPE** | Lower | Higher | | |
| | **GE** | Consistent | | | |
| General meaning | **SPE** | Lower | Higher | | |
| | **GE** | Higher | Lower | | |
| **Belief in memorisation VS:** | | | | | |
| Examples for vocabulary | **SPE** | Higher | Lower | | |
| | **GE** | Higher | Lower | | |
| Texts for grammar | **SPE** | Lower | Higher | | |
| | **GE** | Higher | Lower | | |

The questionnaire results showed that the two groups were inconsistent in the

majority of the items. In the majority of the cases there were beliefs which were considered important, but their corresponding activities were reduced in frequency in test preparation. This was particularly true of the beliefs about 'using English', 'using tapes and CDs', and 'guessing', which probably indicates the negative washback of the test in that the test did not encourage use of English as well as use of context in the case of guessing. However, it was not clear from the questionnaire results whether the increase or decrease was because of the test or other factors.

In the interviews, the consistency of GE and SPE students in terms of vocabulary, grammar, and the use of tapes and CDs were similar. There were no inconsistencies for vocabulary but two inconsistencies for grammar and inconsistencies for the majority of the students in both groups (6 SPE and 7 GE) for the use of tapes and CDs. Both tests had no effect on vocabulary belief but affected two students positively in terms of grammar and the majority of the students negatively in terms of tapes and CDs. The effects on grammar belief were considered positive as the two tests encouraged the two students to do a skill which was tested in the two tests. However, the effects of the two tests on the oral skill were considered negative as the majority of the students were discouraged to do oral activities despite the fact that they believed they were useful.

Concerning guessing, most of the reasons the students gave concerned both exam and non-exam situations, i.e. there were largely consistencies between the reported beliefs and the reported activities. However, 3 GE students were inconsistent in their beliefs and activities. Two of the inconsistencies were due to the test and the other was due to the learner's characteristic i.e. weakness. However, of the 9 SPE students only one student was affected by the test. In other words, SPE students' beliefs about guessing were more stable than GE students'. Belief about guessing was

also more stable than belief about using tapes and CDs. However, belief about vocabulary was the most stable. This suggested that stability of beliefs varied.

In short, the analysis of beliefs showed that GE and SPE students' beliefs were affected similarly. Neither test affected vocabulary belief but affected two students positively in terms of grammar and the majority of the students negatively in terms of the use of tapes and CDs. As regards the belief about guessing, 1 SPE student and 2 GE students were affected. The results suggested that some beliefs were less affected by the test i.e. were more stable than others. The stability also varied from one group to another.

# Chapter 10. Conclusions

This study was inspired by the claim that English-major students admitted to university since 2002 have been more successful in their university studies than pre-2002 students (Chapter 1). The study was also motivated by the claims in the washback literature about the influence of high stakes tests on the one hand and lack of sufficient evidence for these claims on the other, particularly in the area of washback on learning (Chapter 2).

I examined the reported learning activities of the students in the two contexts of out of class as well as in the classroom to see what they did in order to prepare for the test (Chapters 6, 7, and 8). I also tried to understand the reasons why the students did what they reported they did. In this connection, I also examined the reported learning beliefs of the learners as a factor with which the test might interact. The reason I considered beliefs was because the SPE test preparation was partly done out of school i.e. independently and therefore I wanted to investigate whether the students' own criteria for learning had a significant role and how the test was interacting with those criteria. Thus, reported activities and learning beliefs were the main themes of this study.

I collected the data from two contrasting groups- students who were going to take the SPE Test and students who were going to take the GE Test. I used three instruments to collect the data: questionnaire, letters, and interviews. I developed the questionnaire items based on the results of two focus groups, the learning tips in the Pre-University textbook and *Bridging the Gap*, and some existing questionnaires. A sample of students was asked to write a letter to their friends and recommend activities they thought were necessary to do to prepare for the test. Finally, follow-up interviews were used to explore the reasons for the students' reported activities. 1038

students responded to the questionnaire, 91 students wrote the letters, and 18 students participated in the interviews. A full discussion of the research process can be found in Chapters 4 and 5.

The results of the study showed that the SPE Test had washback on the reported use of extra materials, the length of time the students reported spending on each section of the test, and the length of time the students reported spending on learning English as a whole as opposed to the non-English subjects. However, the *new* test sections did not show any washback. The students also said why they did or did not do these activities, which I discuss in 10.1.1 in connection with the theoretical contributions of the study. The results also showed that the SPE students had better English backgrounds and perceived the test to be more difficult than the GE students. These results were obtained in answer to Research Questions 1, 2, 3, and 4, which all mainly dealt with predicted washback.

As regards beliefs, in answer to Research Question 5 (What beliefs do the SPE students report holding about learning English?), the questionnaire results showed that SPE students reportedly held more contemporary than traditional beliefs, compared to the GE students. However, it was not clear from the results whether the differences in beliefs of the two groups were due to the test or whether the beliefs were affecting the students' test preparation activities. The questionnaire analysis was taken further in order to answer Research Question 6 (Are the SPE students reported activities consistent with their reported beliefs?). The analysis showed that the majority of the students' beliefs were inconsistent with their test preparation activities. This also suggested that there might be a mutual relationship between the test preparation activities and the beliefs, i.e. the beliefs might affect the learning activities or the learning activities might affect the beliefs. Therefore, I hoped the interviews

would show whether the students made any link between what they did for their test preparation and their beliefs.

The interview results showed that beliefs about guessing were consistent with guessing activities for test preparation, i.e. the students gave the same reasons for why they believed in guessing in general as well as why they used guessing for their test preparation. However, as far as beliefs about language skills are concerned, there were consistencies for all skills except oral skills, i.e. while all the students gave reasons why they believed in the importance of oral skills for language learning in general, the majority (6 SPE and 8 GE students) gave reasons why they did not actually do oral activities for test preparation, and only the minority in each group both believed in and did oral activities.

In this chapter, I will discuss the contributions, implications, and limitations of this study, generalizability of the findings, and finally will make some suggestions for future studies.

## 10.1. Contributions of this study

### 10.1.1. Theoretical contributions

In this section, I show how my research has contributed to a better understanding of how washback works. The study showed that some of the students' test preparation activities were affected by the test as well as other factors. This confirms that the nature of test washback is not straightforward but complex, as noted by Alderson and Wall (1993), Bailey (1996), Wall (2005) and others. Based on the results, the various factors that need to be taken into account when promoting positive washback include learners, teachers/classrooms/ schools, extra private classes, test design/test constructors, learning materials/materials developers/curriculum design, teaching and learning plan, peers, and learning beliefs.

As far as 'learners' are concerned, their views of 'what constituted the test constructs' affected the amount of time they spent on the test sections. They considered vocabulary and grammar the central constructs on which all the test sections depended to the extent that for some of the sections, particularly Language Functions and Sentence Structure, they did not do particular learning activities except using practice tests only to become familiar with the questions. To one of the learners, grammar was limited as it comprised fixed rules, which led her to spend less time on it than on vocabulary and reading. To some of the learners, speed was one of the constructs measured by the SPE Test. For example, some students practiced reading in order to increase their reading speed. If the learners had good 'English backgrounds' in a particular section, they spent less time on that section than on other sections and vice versa. Their activities were also affected by their 'perception of the difficulty' of the test sections. If they considered a section too easy or too difficult, they spent less time learning them (this agreed with Watanabe's (2001) finding). Similarly, if a learner 'disliked a section', they spent less time on that section. The learners' perception of 'the amount of time a section required' was another factor which affected their learning activities, i.e. for some learners, there was a trade-off between the amount of time a section required in test taking and the amount of time spent on preparing for the section. If answering a section required longer than the other sections, less time was spent on learning the section in test preparation and vice versa. Depending on what 'purpose' the learners had, their learning activities were different. Some learners did oral activities in order to improve their general language ability which they thought would lead to successful performance on the different test sections. Other learners did oral activities because of their 'interest', 'status of English' as the international language, preference for oral materials over written

materials, or a preference for a particular audio device such as the radio, despite the fact that the test did not test oral abilities.

This study also showed that 'the teacher's emphasis on a test section' in the classroom, their 'teaching plan', and 'adequacy/inadequacy of the school' influenced the students' learning activities. The teacher's emphasis on the test sections was reflected by their 'informal questions in the classrooms', 'quizzes', 'end-of-term exams' and 'school-leaving exams', which led the students to emphasise the sections accordingly. Some 'teachers prioritized the test sections in their teaching schedule', which led the students to study some sections earlier in the test preparation period and other sections later, although the learners had their own priorities as well. Some learners 'sought their school teachers' advice on the selection of learning materials, on the relative emphasis they should put on the test sections', etc. However, other learners attended 'private preparation classes' because they considered the school inadequate. 'Private English institutes' had their own impact as well. For example, one student justified her spending time on grammar by saying that in the English institute she went to, the extent to which they knew grammar was important.

'The design of the test' was another factor which interacted with the effect of the test. In justification for using extra materials, one student expressed doubt on the predictive validity of the SPE Test in that, even though it was possible to pass the test by studying the Pre-University Book and the Non-English subjects, there was no guarantee that the students would do well at university. 'Predictability of the items' was another factor which reportedly affected the students' activities. One student reasoned that she did not spend much time on 'Sentence Structure' because the questions in this section were predictable.

As regards learning materials, 'mismatch between the textbooks and the test

sections', 'the extent to which the materials emphasized the test sections', 'availability of the materials', 'appeal of the textbooks', 'whether the textbooks were updated or not', and 'core/ general subjects in the curriculum' also reportedly affected the learners' activities'. One SPE student used practice tests for Language Functions and Sentence Structure because she thought that SPE textbooks were not suitable for these two sections. Some learners thought that the degree to which they emphasized the test sections depended on the textbooks. For example one student said that the Pre-University textbook put most emphasis on vocabulary, then on reading and then on grammar. To some learners the Pre-University book was uninteresting, and included repetitive materials and contradictory explanations. Concerning the Cloze section, one student focused on vocabulary and grammar because she said no materials were available for this section. One student also said that the language proficiency of the students had increased in recent years but the books were not updated to catch up with the increased ability of the students. The 'general non-English subjects in the curriculum', which were required by the University Entrance Examinations for all majors, led to a compromise on the extent to which the learners spent time on English. Some students said they thought that focus on these subjects would help them raise their ranking in the Examination and would compensate for their weakness in English.

Finally, the learners' peers also affected their test preparation activities. Some learners spoke English with their 'classmates' and 'siblings' which they thought was useful for their test preparation. Some of them obtained information about test preparation from their 'friends', e.g. about the weight of English and Non-English sections. The ability of the learners' peers also had an impact on their activities. Some learners reportedly spent more time on the test sections which they thought their peers were weak at.

Of the learner factors mentioned above, this study paid special attention to the learners' own learning criteria or beliefs. The interview results showed that while the reported beliefs about guessing, grammar and vocabulary were consistent (stable) with the corresponding test preparation activities for the majority of the students in both groups, the beliefs about oral skills were inconsistent. For instance the minority of the students who did do oral activities, did them because of the test as well as their beliefs. This leads to the conclusion that the relationship between the reported beliefs and the reported activities was mutual, i.e. the test preparation activities of the students both affected their beliefs and were affected by them. This is in line with Barcelos's (2003) claim that not only may beliefs drive actions, but also actions and reflections on experiences may lead to changes in beliefs. The findings are also in line with Alderson and Wall's (1993) Washback Hypotheses 6 (A test will influence how learners learn), 10 (A test will influence the degree and depth of learning), and 11 (A test will influence attitudes to the content, method, etc of ... learning). Beliefs about guessing, which is related to the use of context in language learning and oral skills, which is in turn related to the use of language, were contemporary beliefs which led the students to do the corresponding activities and at the same time were reinforced and validated by the activities leading to positive washback. However, this was not the case for the majority of the students who did not do oral activities and the minority of the students who either did not believe in the value of guessing or believed in the value of guessing but did not use guessing in their test preparation. The general conclusion is that depending on what kinds of beliefs the learners have, positive or negative washback could be expected. Beliefs based on current thinking about language learning (i.e. contemporary beliefs) could lead to positive washback and traditional beliefs could lead to negative washback. Similarly, tests which induce activities which validate

contemporary beliefs but invalidate traditional beliefs are expected to result in positive washback.

The study also showed that not only were different beliefs affected differently but also the effects were different for the groups and the individuals within each group. The belief about the value of oral skills was the least stable (compared to beliefs about vocabulary, grammar, and guessing), but it was less stable (i.e. more inconsistent) for the GE Group than for the SPE Group. That beliefs were affected differently sheds light on the controversy about the stability of beliefs. Contrary to approaches which view beliefs as either stable or unstable, this study showed that these views are not generalizable to all beliefs as some are more stable and some are less stable. More research is needed to determine which beliefs are predictably stable and which are predictably unstable. The fact that the groups and the individuals were affected differently is consistent with Washback Hypothesis 15 ('Tests will have washback effects for some learners and some teachers, but not for others) (Alderson and Wall, 1993). Depending on the degree of the stability of the learners' beliefs, the degree of the effect of the test may vary and it may vary from one individual or group to another individual or group.

In sum, this study has made several theoretical contributions: 1) it has identified factors contributing to the complexity of washback, 2) it was a step towards understanding interactions between the test and independence / learning beliefs as called for by Bailey (1999), and 3) it has contributed to a further specification of 'washback hypotheses' (Alderson and Hamp-Lyons, 1997: 295- 6), i.e. that 'a test may affect learning beliefs'.

### 10.1.2. Methodological contributions

This study made three methodological contributions. It introduced a new

instrument in washback studies, i.e. letters. It showed that the instrument can be used in grounded research as well as in a study in which the researcher has a pre-determined agenda, i.e. theory-driven research.

The study also showed how beliefs can be researched in a washback study. As mentioned in 9.1.4, in the analysis of the consistency between reported beliefs and reported activities, first I compared the individual group's beliefs with their activities in the case of the questionnaires and individual students' beliefs with their activities in the case of the interviews. Then, I contrasted the SPE with the GE group. This type of contrast between beliefs and activities of individual groups or individual students might be useful for a situation where there is no appropriate contrast group or a baseline study. It may also be useful for a situation in which unintended washback is going to be studied or when washback is examined through a communicative language learning framework (Bailey, 1996: 260). Finally, this study can serve as a baseline study for possible future innovations in the SPE Test.

## 10.2. Implications of the findings

### 10.2.1. Implications of the findings for testing practitioners

Awareness needs to be raised among test developers about the nature of testing and nature of evidence. As the study showed that learners' beliefs could be a potential factor interacting with the test, test developers need to develop tests which encourage current thinking about language learning. They also need to work closely with language teachers and curriculum designers who can help learners develop contemporary beliefs.

As the majority of the students in this study did not do any oral skills activities, it is suggested that an oral subtest should be in the SPE Test so that the learners practice oral skills as well. However, due to the large number of applicants and the

consequent problems of scoring, the SPE Test can be administered by two subtests in two stages: a written test in the first stage and an oral test in the second stage. Twice as many applicants as the university can accommodate can be provisionally admitted through the written test and can then be screened through the oral test. It is also possible to include a listening subtest in the first stage as long as the subtest is objective and is scored by machine. However, the oral test in the second stage does not have to be as objective because of the smaller population.

Since quite a few SPE students emphasised non-English subjects in their test preparation, despite the heavier weight of English, it is suggested that more weight should be given to English (including the 'new section') with more items.

The results showed that most of the students focused on grammar and vocabulary. Therefore, it is suggested that these sections should be given less weight or be tested indirectly through the four language skills and communicatively.

### 10.2.2. Implications of the findings for materials developers

Concerning the mediation of learning materials in the washback process, the results of the study showed that the textbooks were not appropriate for the Language Functions section and did not catch up with the increased proficiency of the students. Therefore, it is suggested that materials writers develop materials which match the contents of the test and are regularly updated.

## 10.3. Limitations of this investigation

This study had several limitations. Since I did not have access to baseline data, I had to adopt a contrasting groups design, i.e. students who were going to take the SPE Test and students who were going to take the GE Test. These two groups were not completely similar in their characteristics.

Although one of the foci of the study was learning behaviour, the study was

based on self-report data, not observed behaviour. This was because the teachers whom I talked to were reluctant to be observed (see also 10.5).

The data was collected at the beginning of the test preparation period and some students talked about activities which they had not yet started. Therefore, full washback might not have appeared yet.

I had a practical problem with the interviews which did not allow for more probing of the issues raised in the interviews. Because it was close to end-of-term examinations, the students had to be present in classes to catch up with the materials. Therefore, it was not easy to get permission from the teachers for the students' interviews. I was usually given only half an hour to conduct the interviews.

My pilot study experience showed that when I asked the students to take the questionnaires home to answer, the return rate was low. Drawing on this experience, I had to ask the students to write the letters under time constraints, i.e. after responding to the questionnaires, instead of asking them to write the letters at home at their convenience. Consequently, the letters they wrote were generally short.

## 10.4. Generaizability of the study results

The results of the questionnaire can fairly confidently be generalised to the whole country particularly as far as the results for the 'predicted washback' are concerned. This is because over 1000 students were sampled from different regions, different schools, different fields of study, and different levels of English background; inferential statistics showed significant differences between SPE and GE students; and because the major part of the results relating to predicted washback was confirmed by the other instruments. The results can be generalised to the Pre-university students who are preparing for the SPE and GE tests.

## 10.5. Future directions

This study can be replicated with the data collected at different points of time in order to get a fuller picture of the SPE Test washback.

I also suggest this study be replicated using classroom observation. Although it was not practical in this research to observe the out-of-class self-study activities of the learners, classroom observation would give a better picture of washback by showing the possible mediation of the teacher in the washback process.

I suggest another study in which letters and interviews will be used to further examine the efficiency of letters as an instrument. Letters can be tried without time limitation and with students interested in writing.

Beliefs can be addressed in a more direct way in washback studies. Instead of using a separate 'beliefs questionnaire', it can be adapted for the study. For example, the item, 'the most important part of language learning is learning the vocabulary words' can be adapted for SPE test preparation: 'the most important part of preparing for SPE Test is learning the vocabulary words'.

Finally, I suggest that the relationship between washback and other aspects of independence including motivation and self-assessment be investigated.

## 10.6. Final remarks

This study has contributed to a better understanding of the complexity of test washback. It has identified factors other than the test which might interact with the test to produce certain kinds of washback. Among these factors, the study paid particular attention to the relationship between beliefs about learning and the test. It has also shown how beliefs can be investigated within a washback framework. Additionally, the study has made clear potential contributions to the work of testing practitioners, teachers and materials developers. In terms of the research methodology,

the study has demonstrated how various instruments from those closer to a 'grounded' approach (the letters) to those closer to a 'theory-driven' approach (the questionnaire) can be used in washback studies. It also introduced an instrument which had not been used in washback studies before, i.e. letters. Finally, this research has contributed to my own experience as a researcher in the area of language testing in terms of how to design and adapt various data collection instruments, how to use them in data collection, how to analyse the results and use the relevant software, how to tackle theoretical dilemmas (e.g. the issue of the stability of beliefs) and methodological dilemmas (e.g. whether to use observation or not), and in general how to think as a researcher in the entire research process.

# Bibliography


Alderson, J. C. (1992). Validating Questionnaires. *CRILE Working Papers, 15*, Lancaster: LAMEL, Lancaster University.

Alderson, J. C. & Wall, D. (1993). Does washback exist? *Applied Linguistics, 14(2)*, 115-129.

Alderson, J. C. and Hamp-Lyons, L. (1996). TOEFL Preparation Courses: A Study of Washback. *Language Testing, 13(3)*, 280-297.

Alderson, J. C., & Banerjee, J. (1996). *How Might Impact Study Instruments Be Validated?* Department of English and Modern English Language, Lancaster University.

Alderson, J. C., & Banerjee, J. (2001). Impact and Washback Research in Language Testing. In C. Elder, A. Brown, E. Grove, K. Hill, N. Iwashita, T. Lumley, T. McNamara & K. 0' Loughlin (Eds.), *Experimenting with Uncertainty. Essays in Honour of Alan Davies* (pp. 150-161). Cambridge: CUP.

Alderson, J. C., & Banerjee, J. V. (2001). State-of-the-Art Review of Language Testing and Assessment (Part I). *Language Teaching, 34, 213-236.*

Alexander, P. A. & Dochy, F.J. (1995). Conceptions of Knowledge and Beliefs: A Comparison across Varying Cultural and Educational Communities. *American Educational Research Journal, 32*, 413-442.

Allen, L. (1996). The Evolution of a Learner's Beliefs about Language Learning. *Carleton Papers in Applied Language Studies, 13*, 67-80.

Andrews, S. (1994a). The Washback Effect of Examinations: Its Impact Upon Curriculum Innovation in English Language Teaching. *Curriculum Forum, 4(1)*, 44-58.

Andrews, S. (1994b). Washback or Washout? The Relationship between Examination Reform and Curriculum Innovation. In D. Nunan & R. Berry (Eds.), *Bringing About Change in Language Education* (pp. 67-8 1). Hong Kong: Department of Curriculum Studies. University of Hong Kong.

Andrews, S., Fullilove, J., & Wong, Y. (2002). Targeting Washback - a Case Study. *System, 30*, 207-223.

Argyris, C. and Schon, D.A. (1974). *Theory in Practice.* San Francisco, CA: Jossey Bass.

Argyrous, G. (2005). *Statistics for Research with a Guide to SPSS.* Sage Publications.

Ary, D, Jacobs, L. C., Razavieh, A., Sorensen, C. (2006). *Introduction to Research in Education.* Thomson Wadsworth.

Bachman, L. F., & Palmer, A. (1996). *Language Testing in Practice: Designing and Developing Useful Language Tests.* Oxford: OUP.

Bailey, K. M. (1996). Working for Washback: A Review of the Washback Concept in Language Testing. *Language Testing, 13(3)*, 257-279.

Bailey, K. M. (1999). Washback in Language Testing. *TOEFL Monograph Series.* Report Number: RM-99-04. TOEFL-MS-15. Princeton. NJ: Educational Testing Service. Retrieved 11/3/2001, from http://www.ets.org

Banerjee, J. V. (2003). *Interpreting and Using Proficiency Test Scores.* Unpublished PhD thesis, Department of Linguistics and Modern English Language, Lancaster

University, Lancaster, England.

Barcelos, A. M. F. (2000). *Understanding Teachers' and Students' Language Learning Beliefs in Experience: A Deweyan Approach (John Dewey).* Unpublished doctoral dissertation. The University of Alabama, Tuscaloosa.

Barcelos, A. M. F. (2003). Researching beliefs about SLA: A critical review. In P. Kalaja and A. M. F. Barcelos (Eds.), *Beliefs about SLA: New Research Approaches* (pp. 7-33). Dordrecht: Kluwer Academic Publishers.

Barker, C., Pistrang, N., and Elliott, R. (2002). *Research Methods in Clinical Psychology* (2nd ed.). John Wiley and Sons.

Benson, P. & Lor, W. (1999). Conceptions of Language and Language Learning. *System, 27(4)*, 459-472.

Best, J. W. and Kahn, J, V. (2006). *Research in Education.* Pearson.

Breen, M. P. (Ed.) (2001). *Learner Contributions to Language Learning. New Directions in Research.* Harlow, Essex: Pearson Education Limited.

Chan, V., Spratt, M., & Humphreys, G. (2002). Autonomous Language Learning: Hong Kong Tertiary Students' Attitudes and Behaviours. *Evaluation and Research in Education, 16(1)*, 1-18.

Cheng, L. (1997). How Does Washback Influence Teaching? Implications for Hong Kong. *Language and Education, 11,* 38-54.

Cheng, L. (1998). Impact of a Public English Examination Change on Students' Perceptions and Attitudes towards their English Learning. *Studies in Educational Evaluation, 24(3)*, 279-301.

Cheng, L. (1999). Changing Assessment: Washback on Teacher Perspectives and Actions. *Teaching and Teacher Education, 15*, 253-271.

Cheng, L. and Y. Watanabe (2004) (Eds.). *Washback in Language Testing: Research Contexts and Methods.* Mahwa, New Jersey: Lawrence Erlbaum Associates.

Cohen, A. D. (1998). *Strategies in Learning and Using a Second Language.* Addison Wesley Longman Limited.

Cohen, A. D. and Chi, J. C. (2002). Language Strategy Use Inventory and Index. In R. M. Paige, A. D. Cohen, J. Lassegard, J. C. Chi, & B. J. Kappler. *Maximizing Study Abroad: A Students' Guide to Strategies for Language and Culture Learning and Use.* Minneapolis, MN: Center for Advanced Research on Language Acquisition (CARLA).

Cohen, L., Manion, L. and Morrison, K. (2000). *Research Methods in Education.* London and New York: Routledge/Falmer

Colman, A. M. (2001). *A Dictionary of Psychology.* OUP

Colman, A. M. (2006). *A Dictionary of Psychology* (2nd ed.). OUP

Cotterall, S. M. (1995). Readiness for Autonomy: Investigating Learner Beliefs. *System, 23(2)*, 195-205.

Cotterall, S. M. (1999). Key Variables in Language Learning: What Do Learners Believe about them? *System, 27*, 493-513

Davies, A., Brown, A., Elder, C., Hill, K., Lumley, T., & McNamara, T. (1999). *Dictionary of Language Testing*. Cambridge: CUP/UCLES.

Davis, A. (2003). Teachers' and Students' Beliefs Regarding Aspects of Language Learning. *Evaluation and Research in Education, 17 (4)*, 207-222.

Dewey, J. (1933). *How We Think.* Lexington, MA: D.C. Heath.

Dornyei, Z. (2003). *Questionnaires in Second Language Research*. Mahwah, N.J.: Lawrence Erlbaum.

Dornyei, Z. and Skehan, P (2003). Individual Differences in Second Language Learning. In J. D. Doughty, and M. H. Long (Eds.), *The Handbook of Second Language Acquisition.* Blackwell.

Doughty, J. D. and Long, M. H. (Eds.) (2003). *The Handbook of Second Language Acquisition.* Blackwell.

Ellis, R. (2001). The Metaphorical Constructions of Second Language Learners. In M.

Breen (Ed.), *Learner Contributions to Language Learning: New Directions in Research* (pp. 65- 85). Singapore: Pearson

Ferman, I. (2004). The Washback of an EFL National Oral Matriculation Test to Teaching and Learning. In L. Cheng, L and Y. Watanabe (Eds.), *Washback in Language Testing: Research Contexts and Methods.* Mahwa, New Jersey: Lawrence Erlbaum Associates.

Field, A. (2005). *Discovering Statistics Using SPSS* (2nd ed.). Sage Publications.

Flavell, J. H. (1979). Metacognition and Cognitive Monitoring: A New Area of Cognitive Development Inquiry [Special Issue]. *American Psychologist, 34(10)*, 906-911.

Flavell, J. H. (1981). Monitoring Social Cognitive Enterprises: Something Else That May Develop in the Area of Social Cognition. In J. H. Flavell, & L. Ross (Eds.), *Social Cognitive Development: Frontiers and Possible Futures* (pp. 272-287). New York: Cambridge University Press.

Frederiksen, J. R., & Collins, A. (1989). A Systems Approach to Educational Testing. *Educational Researcher, 18(9),* 27-32.

Frederiksen, N. (1984). The Real Test Bias: Influences of Testing on Teaching and Learning. *American Psychologist, 39(3),* 193-202.

Fullilove, J. (1992). The Tail That Wags. *Institute of Language in Education Journal, 9*, 131-147.

Gardner, R. C. (1985). *Social Psychology and Second Language Learning: The Role of Attitudes and Motivation.* London: Edward Arnold Publishers.

Gillham, B. (2002). *Developing a Questionnaire.* Continuum.

Goh, C. C. M. (1997). Metacognitive Awareness and Second Language Listeners. *ELT Journal,* 51*(4)*, 361-369.

Goh, C. C. M. & Lin, L. X. (1999). Learning Environments and Metacognitive Knowledge about Language Learning. *Hong Kong Journal of Applied Linguistics, 4 (2)*, 41-56.

Gosa, C. M. C. (2004). *Investigating washback: A Case Study Using Student Diaries.* Unpublished PhD dissertation. Lancaster University, UK.

Guba, E. G., and Lincoln, Y. S. (1981). *Effective Evaluation.* San Francisco: Jossey-Bass.

Haladyna, T. M., Nolen, S. B., & Haas, N. S. (1991). Raising Standardized Achievement Tests Scores and the Origins of Test Score Pollution. *Educational Researcher, 20(5),* 2-7.

Hamp Lyons, L. (1997). Washback, Impact and Validity: Ethical Concerns. *Language Testing, 14(3),* 295-303.

Heaton, J. B. (1990). *Classroom Testing.* Harlow: Longman.

Henrichsen, L. E. (1989). *Diffusion of Innovations in English Language Teaching: The ELEC Effort in Japan,* 1956-1968, New York: Greenwood Press.

Herington, R. (1996). *Test-Taking Strategies and Second Language Proficiency: Is There a Relationship?* Unpublished dissertation. Lancaster University, UK.

Holec, H. (1981). *Autonomy and Foreign Language Learning.* Oxford. Pergamon.

Horwitz, E. K. (1985). Surveying Student Beliefs about Language Learning and Teaching in the Foreign Language Methods Course. *Foreign Language Annals, 18(4),* 333-340.

Horwitz, E. K. (1987). Surveying Student Beliefs about Language Learning. In A. L. Wenden & J. Robin (Eds.), *Learner Strategies in Language Learning* (pp. 119-132). London: Prentice Hall.

Horwitz, E. K. (1988). The Beliefs about Language Learning of Beginning Foreign Language Students. *Modern Language Journal, 72(3),* 283-294.

Horwitz, E. K. (1999). Cultural and Situational Influences on Foreign Language Learners' Beliefs about Language Learning: A Review of BALLI Studies [Special Issue]. *System, 27(4),* 557-576.

Hughes, A. (1994). *Backwash and TOEFL 2000.* Unpublished manuscript, commissioned by Educational Testing Service (ETS). University of Reading.

Hughes, A. (2003). *Testing for Language Teachers* (2nd ed.). Cambridge: CUP.

Huhta, A., Kalaja, P., & Pitkanen-Huhta, A. (2006). Discursive Construction of a High-Stakes Test: the Many Faces of a Test Taker. *Language Testing, 23(3),* 326-350.

Johnson, K. and Johnson, H. (1998) (Eds). *Encyclopedic Dictionary of Applied Linguistics.* Blackwell publishers.

Kalaja, P. (1995). Student Beliefs (or Metacognitive Knowledge) about SLA Reconsidered. *International Journal of Applied Linguistics, 5(2),* 191-204.

Kalaja, P. (2003). Research on students' beliefs about SLA within a discursive approach. In P. Kalaja & Barcelos, A. M. F. (Eds.), *Beliefs about SLA: New Research Approaches* (pp. 87-108). Dordrecht: Kluwer Academic Publishers.

Kalaja, P., Alanen, R., & Dufva, H. (forthcoming). Teacher Trainees' Beliefs about EFL Learning in the Light of Narratives. In S. Breidbach, D. Elsner, and A. Young (Eds.), *Anthology of Language Awareness in Teacher Education.*

Krueger, R. A. Casey, M. A. (2000). *Focus Groups: A Practical Guide for Applied Research.* (3rd ed.). Sage Publications.

Kvale, S. (1996). *Interviews: An Introduction to Qualitative Research Interviewing.* Thousand Oaks, California: Sage Publications.

Li, X. (1990). How Powerful Can a Language Test Be? The MET in China. *Journal of Multilingual and Multicultural Development, 11(5),* 393-404.

Lightbown, P. M. and Spada, N. (1993). *How Languages are Learned.* Oxford: Oxford University Press.

Luxia, Q. (2005). Stakeholders' Conflicting Aims Undermine the Washback Function of a High-Stakes Test. *Language Testing, 22(2),* pp. 142-173.

Madaus, G. F. (1988). The Influence of Testing on the Curriculum. In L. N. Tanner (Ed.), *Critical Issues in Curriculum. 87th Yearbook for the National Society for the Study of Education* (pp. 83-121). Chicago: University of Chicago Press.

Marshall, C. & Rossman, G. B. (2006). (4th edition). *Designing Qualitative Research.* Sage Publications.

Miller, L., & Ginsberg, R. B. (1995). Folklinguistic Theories of Language Learning. In B. F. Freed (Ed.), *Second Language Acquisition in a Study Abroad Context* (pp. 293-315). Amsterdam: John Benjamins.

Morgan, D. L. (1997). *Focus Groups as Qualitative Research.* Qualitative Research Methods Series, Volume 16. Sage Publications.

Morrow, K. (1986). The Evaluation of Tests of Communicative Performance. In M. Portal (Ed.), *Innovations in Language Testing. Proceedings of the IUS/NFER Conference* (pp. 1-1 3). London: NFER/Nelson.

Mousavi, S. A. (Ed.) (1999). *A Dictionary of Language Testing* (2nd ed.). Tehran: Rahnama Publications, Iran.

Muhr, T. (2003-2004). *Atlas.Ti: The Knowledge Workbench.* Berlin: Scientific Software Development.

Oxford, R. L. (1990). Strategy inventory for language learning. In Oxford, R. L. (Ed.), *Language Learning Strategies: What Every Teacher Should Know* (pp. 283–300). Boston: Heinle & Heinle.

Pallant, J. (2007). *SPSS Survival Manual.* Open University Press.

Patton, M. Q. (2002). *Qualitative Research and Evaluation Methods* (3rd ed.). Newbury Park, California: Sage Publications.

Popham, W. J. (1987). The Merits of Measurement-Driven Instruction. *Phi Delta Kappan, 68,* 679-682.

Reber, A. S. and Reber, E. S. (2001). *The Penguin Dictionary of Psychology.* (3rd edition). Penguin Books.

Riley, P. (1994). Aspects of Learner Discourse: Why Listening to Learners is So Important. In E. Esch (Ed.), *Self-access and the Adult Language Learner* (pp. 7-18). London: Centre for Information on Language Teaching.

Riley, P. (1997). The Guru and the Conjurer: Aspects of Counseling for Self-Access. In P. Benson, & P. Voller (Eds.), *Autonomy and Independence in Language Learning* (pp. 114-131). New York: Longman.

Sakui, K. & Gaies, S. J. (1999). Investigating Japanese Learners' Beliefs about Language Learning. [Special Issue]. *System, 27,* 473-492.

Saljo, R. (1979). Learning in the learner's Perspective- Differences in Awareness. *Reports from the Institute of Education, No 76*, University of Goeteborg.

Seidman, I. (2006). *Interviewing as Qualitative Research: A Guide for Researchers in Education and the Social Sciences* (3rd ed.). New York: Teachers College Press.

Shohamy, E. (1993). *The Power of Test: The Impact of Language Testing on Teaching and Learning*. National Foreign Language Center Occasional Papers. Washington, DC: National Foreign Language Center.

Shohamy, E., Donitsa-Schmidt, S. and I. Ferman (1996). Test Impact Revisited: Washback Effect over Time. *Language Testing 13(3)*. 298-317.

Silverman, D. (2001). *Interpreting Qualitative Data: Methods for analysing talk, text and interaction (2nd ed.)*. London: Sage.

Smith, M. L. (1991a). Meanings of Test Preparation. *American Educational Research Journal, 28*, 52 1-542.

Smith, M. L. (1991b). Put to the Test: The Effects of External Testing on Teachers. *Educational Researcher, 20(5),* 8-11.

Sudman, S. and Bradburn, N. M. (1983). *Asking Questions*. San Francisco, CA: Jossey-Bass.

Sunderland, J. (2010). Research questions. In Litosseliti, L. (Ed.) *Research Methods in Linguistics*. London: Continuum

Tsagari, C. (2006). *The Wasback Effect of the FCE Test in a Private Language School in Greece*. Unpublished PhD dissertation. Lancaster University, UK.

Van Lier, L. (1988). *The Classroom and the Language Learner: Ethnography and the Second Language Classroom Research*. London: Longman.

VandenBos, G. R. (Ed.) (2007). *APA Dictionary of Psychology*. Wasington D. C.: American Psychological Association.

Victori, M. (1992). *Investigating the Metacognitive Knowledge of Students of English as a Second Language*. Unpublished Master's Thesis. University of California, Los Angeles.

Victori, M. (1999). An Analysis of Writing Knowledge in EFL Composing: A Case Study of Two Effective and Two Less Effective Writers. *System, 27(4)*, 537-555.

Victori, M. and Lockhart, W. (1995). Enhancing Metacognition in Self-Directed Language Learning. *System, 23(2)*, 223-234.

Wall, D. (1997). Impact and Washback in Language Testing. In C. Clapham & D. Corson (Eds.), *Encyclopaedia of Language and Education*. Vol. 7. Language Testing and Assessment (pp. 291-302). Dordrecht: Kluwer Academic.

Wall, D. (1999). *The Impact of High-Stakes Examinations on Classroom Teaching: A Case Study Using Insights from Testing and Innovation Theory*. Unpublished PhD thesis, Department of Linguistics and Modern English Language, Lancaster University, Lancaster, England.

Wall, D. (2000). The Impact of High Stakes Testing on Teaching and Learning: Can This Be Predicted or Controlled? *System, 28*, 499-509.

Wall, D. (2005). *The Impact of High-Stakes Examinations on Classroom Teaching: A Case*

*Study Using Insights from Testing and Innovation Theory*. Cambridge: CUP.

Wall, D. and Alderson, J. C. (1993). Examining Washback: The Sri Lankan Impact Study. *Language Testing, 10(1),* 41-69.

Wall, D. and Horak, T. (2005). *TOEFL Impact Study*. Language Testing Research Group (LTRG). Lancaster University, UK.

Wall, D., & Horak, T. (2006). The Impact of Changes in the TOEFL® Examination on Teaching and Learning in Central and Eastern Europe: Phase 1, the Baseline Study. TOEFL Monograph Series. Report Number: RR-06-18, TOEFL-MS-34. Princeton, NJ: Educational Testing Service. Retrieved 29/6/2006, from http://www.ets.org.

Wall, D. and Horak, T. (2007). Using baseline studies in the investigation of test impact. *Assessment in Education, 14(1),* 99-116.

Watanabe, Y. (1992). Washback Effects of College Entrance Examination on Language Learning Strategies. *JACET Bulletin, 23*, 175-94.

Watanabe, Y. (1996) Does Grammer-Translation Come from the Entrance Examination? Preliminary Findings from Classroom-Based Research. *Language Testing, 13(3),* 319-333.

Watanabe, Y. (2001). Does the University Entrance Examination Motivate Learners? A Case Study of Learner Interviews. In Akita Association of English Studies (Ed.), *Trans-Equator Exchanges: A Collection of Academic Papers in Honor of Professor David Ingram* (pp. 100-110). Faculty of Education and Human Studies, Akita University.

Watanabe, Y. (2004). Methodology in Washback Studies. In L. Cheng and Y. Watanabe (Eds), *Washback in Language Testing: Research Contexts and Methods*, Mahwa, New Jersey: Lawrence Erlbaum Associates.

Wenden, A. (1986). What Do Second Language Learners Know about their Language Learning? A Second Look at Retrospective Accounts. *Applied Linguistics, 7*, 186-205.

Wenden, A. (1987). How to Be a Successful Language Learner: Insights and Prescriptions from L2 Learners. In A. Wenden & J. Rubin (Eds.), *Learner Strategies in Language Learning* (pp. 103-117). London: Prentice Hall.

Wenden, A., (1991). *Learner Strategies for Learner Autonomy*. Hertfordshire: Prentice Hall International.

Wenden, A. (1999). An Introduction to Metacognitive Knowledge and Beliefs in Language Learning: Beyond the Basics [Special Issue]. *System, 27*, 435-441.

Wenden, A. (2001). Metacognitive Knowledge. In M. P. Breen (Ed.), *Learner Contributions to Language Learning. New Directions in Research* (pp. 44-64). Harlow, Essex: Pearson Education Limited.

Wesdorp, H. (1982). Backwash Effects of Language-Testing in Primary and Secondary Education. *Journal of Applied Language Study, 1(1)*, 40-55.

White, C. (1999). Expectations and Emergent Beliefs of Self-Instructed Language Learners. *System, 27(4)*, 443-457.

Wong. Y. (2001). *The Impact of a New Oral Exam on Students' Spoken Language Performance.* Unpublished BEd dissertation. University of Hong Kong.

Yang, D. (1999). The Relationship between EFL Learners' Beliefs and Learning Strategy [Special Issue]. *System, 27(4)*, 515-535.

# APPENDICES

## Appendix 1: A sample of the Specialised English Test (SPE)

صفحه ۱ — زبان انگلیسی (اختصاصی)

**PART A: Grammar**

**Directions: Questions 101-110 are incomplete sentences. Beneath each sentence you will see four words or phrases, marked (1), (2), (3), and (4). Choose the one word or phrase that best completes the sentence. Then mark the correct choice on your answer sheet.**

**101-Laura ---------- the classroom late because she ---------- with one of her classmates.**
1) had left, would argue 2) would leave, argued 3) left, had been arguing 4) was leaving, was arguing

**102-Don't call me at the usual time tomorrow evening. I ---------- dinner.**
1) will take 2) am taking 3) will be taking 4) am going to take

**103-Uncle Joe was so careless -------- money that he spent £1,000 ---------- clothes --------- the holiday weekend.**
1) of, for, in 2) for, on, at 3) on, for, on 4) with, on, over

**104- ---------- the wallet, ---------- the head.**
1) Fatter, emptier 2) The fatter, the emptier
3) The fattest, the emptiest 4) It's the fattest, it's the emptiest

**105-Voters are faced with a lot of parties ---------- to choose.**
1) of which 2) of them 3) from them 4) from which

**106-Betty did not regret ---------- that she did not let her son ---------- graduate school.**
1) saying, enter 2) saying, entering 3) to say, to enter 4) say, to enter

**107-The baby ----------. Why don't you check to make sure about it?**
1) could have slept 2) might be sleeping 3) must have slept 4) should be sleeping

**108-If you feel sick, you -------- go to school.**
1) had better not 2) had not better 3) would rather not 4) would not rather

**109-We moved very slowly --------- frighten the animals.**
1) so as not to 2) in order to not 3) so that we didn't 4) in order that we didn't

**110-Little ---------- I'd become president of the company.**
1) I thought 2) did I think 3) I had thought 4) had I thought

**PART B: Vocabulary**

**Directions: Questions 111-125 are incomplete sentences. Beneath each sentence you will see four words or phrases, marked (1), (2), (3), and (4). Choose the one word or phrase that best completes the sentence. Then mark the correct choice on your answer sheet.**

**111-She possesses complete --------- of three languages.**
1) capacity 2) quality 3) supply 4) mastery

**112-A ---------- is the place in the ground where a dead body is buried.**
1) layer 2) limb 3) grave 4) memorial

**113-The low rate of unemployment is the ---------- of a healthy economy.**
1) reflection 2) instruction 3) promotion 4) identification

**114-A -------- is an idea that is suggested as an explanation for something, but that has not been proved to be true.**
1) pattern 2) feature 3) research 4) hypothesis

**115-The drug is effective but has a --------- to cause headaches.**
1) pressure 2) requirement 3) tendency 4) function

**116-We need to take a more negative ---------- toward physical punishment at schools.**
1) arrangement 2) stance 3) influence 4) measure

**117-The host's face looked tired, so that was our --------- to leave the party.**
1) cue 2) stage 3) desire 4) purpose

**118-The nurse is ---------- from duty at seven o'clock.**
1) devoted 2) dislocated 3) operated 4) released

**119-The government ---------- thousands of passports each year.**
1) restates 2) issues 3) specifies 4) speculates

**120-The little boy --------- when he said there were a million cats in the backyard.**
1) expressed 2) attributed 3) exaggerated 4) announced

121-The police said the accident -------- about 4:30 p.m.
1) occurred 2) disturbed 3) estimated 4) recorded

122-Oliver made a lot of money by -------- in furniture.
1) handling 2) taking part 3) investing 4) pulling out

123-A(n) -------- crowd waited for the arrival of their popular president.
1) public 2) dense 3) extensive 4) sufficient

124-At night, people have to read by -------- light because there is no sunshine.
1) radiated 2) artificial 3) equipped 4) surrounding

125-For many students, illness is a(n) -------- excuse for being absent from school.
1) valid 2) entire 3) ignorant 4) disappointed

**PART C: Sentence Structure**
**Directions: Choose the sentence with the best order for each of the following series. Then mark the correct choice on your answer sheet.**

126-1) It seems strange now but you'll soon be used to working here
2) It seems strange now despite you'll use to work here soon.
3) It seem strangely now but you'll use working here soon.
4) It seems strangely now despite you'll soon be used to work here.

127-1) He finds it hardly keeping the issues in his private life separate from his public life.
2) He finds it hard keeping the issues in his private life separately from his public life.
3) He finds it hard to keep the issues in his private life separate from those in his public life
4) He finds it hardly to keep the issues in his private life separately from those in his public life.

128-1) The luggage sank in middle of the ocean, sweeping by the tide.
2) The luggage sank in the middle of the ocean, sweeping on the tide.
3) The luggage sank in middle of the ocean after sweeping away on the tide.
4) The luggage sank in the middle of the ocean after it was swept away by the tide.

129-1) His mother refused to give him more independence despite his wishes.
2) His mother refused giving him more independence despite his wishes.
3) His mother refused to give him more independently in spite of his wishes.
4) His mother refused giving him more independently in spite of his wishes.

130-1) Having been bitten by the dog, the robber ran along the street from the old lady.
2) Bitten by the dog, the robber had run along the street from the old lady.
3) Bitten by the dog, the robber ran along the street away the old lady.
4) Having been bitten by the dog, the robber had run along the street away the old lady.

**PART D: Language Functions**
**Directions: Read the following conversations between two people and answer the questions about the conversations by choosing one of the choices (1), (2), (3), or (4). Then mark the correct choice on your answer sheet.**

A: I'm not ready yet, and it's going to (131) -------- longer.
B: You'd better hurry. Take five minutes too long and you'll miss the bus.

131-1) be at times 2) be in a minute 3) take the same time 4) take me a while

A: Hi, Jack. It's good to see you again. Are you ready to (132) -------- business again after this break?
B: Not really, but I guess I don't really have too much choice.

132-1) carry out 2) settle down 3) get down to 4) look forward to

A: Pat needs to get to the airport by 3:00. Can you (133) --------?
B: No problem. I haven't got anything else planned for tomorrow afternoon.

133-1) drive her out 2) let her by car 3) make her a ride 4) give her a lift

A: I just got a call from Lance. He wants us to be there at 4:00.
B: (134) --------, I can't make it by then!

134-1) In a flash 2) Round the clock 3) You must be kidding 4) You're fooling mouth

A: That's a fantastic stereo system! It must have been expensive!
B: I bought it (135) --------.

135-1) at shop 2) on sale 3) for a sale 4) in a shop window

A: It's going to take me several days to revise this research paper.
B: Unfortunately, (136) -------- tomorrow.

136-1) it is time 2) it is due 3) time is behind 4) time is on your side

A: Gary was not prepared for the exam.
B: (137) -------- that he failed it.

137-1) That's the way 2) He broke his heart 3) He left short of time 4) It is not surprising

A: I've told you (138) -------- to stop. Cigarettes can ruin your health.
B: I know. As usual, I didn't listen to you.

138-1) the other day 2) at the moment 3) over and over 4) for the moment

A: I've been studying the piano for ten years, since I was nine years old.
B: (139) --------! I didn't know that. Ten years!

139-1) In good heart 2) To a large extent 3) On the right track 4) For goodness sake

A: A number of the classes I want are already full. Let's speak to the professors directly.
B: Okay, I guess (140) --------.

140-1) it's worth a try 2) they'll count on it 3) it's taken into consideration 4) they'll pay lip service to it

**PART E: Cloze Test**

**Directions: Read the following passage and decide which choice (1), (2), (3), or (4) best fits each space. Then mark the correct choice on your answer sheet.**

In strictly practical terms, schooling yields three rewards, and the amount of each reward increases in (141) ---------- to the amount of schooling. First, the individual who is well schooled (142) ---------- the best chance of getting any job, other things (143) ---------- equal. Thus, the chance of unemployment is (144) ----------. Second, the individual with a good (145) ---------- is the one chosen for advancement and promotion, (146) ---------- enabling him or her to earn more over the long (147) ----------. Third, because of rewards one and two, the (148) ---------- individual has more personal freedom. Such a person will have more job (149) ----------, is less threatened with unemployment, and can be freer economically (150) ---------- his or her higher earning power. The decision in favor of further schooling needs to be encouraged if only for the above listed pragmatic reasons.

| | | | |
|---|---|---|---|
| 141-1) registration | 2) proportion | 3) material | 4) occasion |
| 142-1) raises | 2) conducts | 3) stands | 4) realizes |
| 143-1) are | 2) will be | 3) to be | 4) being |
| 144-1) refused | 2) reduced | 3) replaced | 4) rejected |
| 145-1) application | 2) proficiency | 3) fortune | 4) background |
| 146-1) while | 2) so that | 3) thus | 4) in case |
| 147-1) range | 2) distance | 3) pace | 4) run |
| 148-1) protected | 2) educated | 3) appointed | 4) civilized |
| 149-1) collections | 2) suggestions | 3) opportunities | 4) trainings |
| 150-1) beside | 2) so as to | 3) because of | 4) along with |

**PART F: Reading Comprehension**

**Directions: In this part of the test, you will read three passages. Each passage is followed by a number of questions. Answer the questions by choosing the best choice (1), (2), (3), or (4). Then mark the correct choice on your answer sheet.**

Ranked as the number one beverage consumed worldwide, tea takes the lead over coffee in both popularity and production, with more than 5 million metric tons of tea produced annually. Although much of this tea is consumed in Asian, European, and African countries, the United States drinks its fair share. According to estimates by the Tea Council of the United States, tea is enjoyed by no less than half of the U.S. population on any given day. Black tea or green tea—iced, spiced, or instant—tea drinking had spurred a billion-dollar business with major tea producers in Africa, South America, and throughout Asia.

Tea is made from the leaves of an evergreen plant, *camellia sinensis*, that grows tall and lush in tropical regions. On tea plantations, the plant is kept trimmed to approximately four feet tall, and as new buds, called flush, appear they are plucked off by hand. Even in today's world of modern agricultural machinery, hand harvesting continues to be the preferred method. Ideally, only the top two leaves and a bud should be picked. This new growth produces the highest quality tea.

No one knows when or how tea first became popular, but legend has it that tea as a beverage was discovered in 2737 B.C. by Emperor Shen Nung of China when leaves from a Camellia plant dropped into his drinking water as it was boiling over a fire. As the story goes, Emperor Shen Nung drank the resulting liquid and proclaimed the drink to be most nourishing and fortifying. Though this account cannot be documented, it is thought that tea drinking probably originated in China and spread to other parts of Asia, then to Europe, and ultimately to the American colonies around 1650.

With about half the caffeine content as coffee, tea is often chosen by those who want to reduce, but not necessarily eliminate, their caffeine intake. Some people find that tea is less acidic than coffee and therefore easier on the stomach. Others have become interested in tea drinking since the National Cancer Institute published its findings on the antioxidant properties of tea. But whether tea is enjoyed for its perceived health benefits, its flavor, or as a social drink, teacups continue to be filled daily with the world's most popular beverage.

**151-The passage is mainly about ----------.**
1) what popular beverages are
2) tea consumption and production
3) what the benefits of tea drinking are
4) the advantages of tea over coffee

**152- The word "it" in line 14 refers to ----------.**
1) camellia 2) water 3) tea 4) plant

**153-As to the harvesting of tea, we can understand from the passage that ----------.**
1) modern machinery has facilitated it
2) the method has not changed over time
3) it is not done in China any longer
4) the branches of tea should be removed

**154-According to the passage, drinking tea ----------.**
1) first became widespread at an unknown time
2) started in China according to documents
3) dates back to the Shen Nung dynasty
4) started before drinking coffee in most countries

**155-One reason that some people prefer tea to coffee is that it ----------.**
1) is easier to digest 2) has more caffeine 3) functions against cancer 4) has a higher nutritional value

**156- The author might include statistics on tea production and consumption to ----------.**
1) show the popularity of tea
2) impress the reader with the facts about beverage
3) explain why coffee is not popular
4) explain the cost of tea production

**157- The word "spurred" in line 5 means ----------.**
1) stimulated 2) operated 3) managed 4) announced

**158-According to the passage, it is NOT that ----------.**
1) tea contains some caffeine
2) people like to drink tea for three reasons
3) both black and green tea can be instant
4) the highest quality tea is produced in China

In 1876 George Bernard Shaw, a young man of 20, came to London from his birthplace in Dublin, Ireland, seeking his fortune as a writer. He was to become the most famous playwright of his time. His first attempts were not at plays, however, and when his name did begin to be known it was for his essays, pamphlets, and speeches on behalf of the Fabian Society. As a young man Shaw saw the need for changes and reforms in government, and when he became a socialist he joined the Fabian Society, which had come into being to make the ideas of socialism more widely known and better understood.

Politics were not Shaw's only interest, for he loved art, music, and drama and wrote about these subjects for newspapers and magazines. He himself had a very clear mind, and he wanted to make other people think as clearly as he did about all kinds of subjects. It was for this that he began writing plays, for he felt that he could best do it through them. The first one was performed in 1892, and at once many people began to condemn him. This play —and all his later ones too—was not written just to tell a story but to prove Shaw's arguments, and these arguments upset people's beliefs and made them feel shocked and uncomfortable. Nevertheless, the plays were far too clever to remain unknown, and several of them became great successes on the stage in Europe and America as well as in Britain. They included *Arms and the Man*, *Candida*, *The Devil's Disciple*, *Caesar and Cleopatra*, *Man and Superman*, and *Pygmalion*.

By the end of World War I (1914-18), a great many more people had begun to think that the old ways and ideas needed altering if the world was to be made better, and Shaw's "plays of ideas" became much more popular than they had been before. *St. Joan* (first performed in 1924) was thought to be the best play of its time. In it Shaw told, in his own particular way, the story of the heroic life and death of Joan of Arc. He was awarded the Nobel Prize for Literature in 1925.

Shaw lived to be 94 and went on writing almost to the time of his death. His works include long prefaces to the plays, which are as interesting as the plays themselves.

**159-The main topic of the passage is ----------.**
1) Shaw as a great politician
2) Shaw's professional career
3) the autobiography of George Bernard Shaw
4) the main plays written By George Bernard Shaw

**160-The increasing popularity of "plays of ideas" ----------.**
1) altered the trend of the war in 1918
2) showed the importance of Joan of Arc
3) was due to people's changing attitudes
4) was the result of people's interest in war

**161-The word "altering" in line 17 is closest in meaning to ----------.**
1) opposing
2) discussing
3) changing
4) rejecting

**162-Shaw's original orientation was largely ----------.**
1) artistic
2) political
3) literary
4) linguistic

**163-Shaw's writings ----------.**
1) included arguments which some people did not favor
2) remained unknown until the end of the war
3) upset people in Europe not as much as in Britain
4) were first performed in 1924

**164-According to the passage, Shaw ----------.**
1) wrote his first play in 1892
2) continued to write in his nineties
3) was awarded the Nobel Prize for *St. Joan*
4) lived in the district called the Fabian Society

Language is a fundamental part of total human behavior, and behaviorists have examined it as such and sought to formulate consistent theories of first language acquisition. The behaviorist approach focuses on the immediately perceptible aspects of linguistic behavior—the publicly observable responses—and the relationships or associations between those response and events in the world surrounding them. A behaviorist might consider effective language behavior to be the production of correct responses to stimuli. If a particular response is reinforced, it then becomes habitual, or conditioned. Thus children produce linguistic responses that are reinforced. This is true of their comprehension as well as production responses, though to consider comprehension is to wander just a bit out of the publicly observable realm. One learns to comprehend an utterance by reacting appropriately to it and by being reinforced for that reaction.

One of the best-known attempts to construct a behaviorist model of linguistic behavior is embodied in B F Skinner's (1957) classic, *Verbal Behavior*. Skinner is commonly known for his experiments with animal behavior in "Skinner's boxes", but he has also gained recognition for his contributions to education through teaching machines and programmed learning. Skinner's theory of verbal behavior was an extension of his general theory of learning by operant conditioning. Operant conditioning refers to conditioning in which the organism (in this case, a human being) emits a response, or operant (a sentence or utterance), without necessarily observable stimuli; that operant is maintained (learned) by reinforcement (for example, a positive verbal or nonverbal response for another person). If a child says "want milk" and a parent gives the child some milk, the operant is reinforced and, over repeated instance, is conditioned. According to Skinner verbal behavior, like other behavior, is controlled by its consequences. When consequences are rewarding behavior is maintained and is increased in strength and perhaps frequency. When consequences are punishing, or when there is lack of reinforcement entirely, the behavior is weakened and eventually extinguished.

**165-The best topic for the passage would be ----------.**
1) The Features of Linguistic Behavior
2) Theories of First Language Acquisition
3) Language as a Main Part of Behavior
4) The Behaviorist Approach to L1 Acquisition

**166-According to behaviorists, language acquisition ----------.**
1) is reinforced by positive responses
2) conditions total human behavior
3) is a kind of habit formation
4) can be explained through different formulas

**167-Skinner used boxes to ----------.**
1) study the behavior of animals
2) prove classical theories of learning
3) experiment with verbal behavior
4) teach some animals to use language

**168-The word "emits" in line 15 means ----------.**
1) observes
2) picks up
3) sends out
4) matures

**169-It is understood from the passage that Skinner's verbal behavior ----------.**
1) was strong and frequent
2) had rewarding consequences
3) gained recognition in 1957
4) was the same as linguistic behavior

**170-According to the passage, it is true that ----------.**
1) an operant is either a response or a stimulus
2) conditioning is a way to reject operants
3) reinforcement can make behavior conditional
4) uncontrolled behavior should be punished

# Appendix 2: A sample of the General English Test (GE)



**Part A: Grammar**

**Directions:** Questions 76-80 are incomplete sentences. Beneath each sentence you will see four words or phrases, marked (1), (2), (3), and (4). Choose the one word or phrase that best completes the sentence. Then mark your answer sheet.

76- She was wearing a ——— shirt.
1) white and black silk beautiful 2) silk beautiful black and white
3) black and white beautiful silk 4) beautiful black and white silk

77-Tom selected a book which he knew was very complex ——— he wanted to leave a good first impression on the professor.
1) since 2) although 3) so that 4) whereas

78- Our teacher usually makes all the students ——— a lecture in the class.
1) give 2) to give 3) be given 4) that they give

79-Ted was frightened when he got closer to the mountain; it was ——— high.
1) such 2) so 3) enough 4) such a

80-I left a message on your answer phone last night. So you ——— informed about the meeting.
1) should have 2) might have 3) must have been 4) should have been

**Part B: Vocabulary**

**Directions:** Questions 81-90 are incomplete sentences. Beneath each sentence you will see four words or phrases, marked (1), (2), (3), and (4). Choose the one word or phrase that best completes the sentence. Then mark your answer sheet.

81-You will lose weight and become weak unless you receive all the ——— your body needs.
1) qualities 2) functions 3) devices 4) nutrients

82- So far the police can only ——— on the possible reason for the killing.
1) research 2) influence 3) speculate 4) operate

83- They have ——— the game; it is going to be ——— at another place later on.
1) take off – taken 2) put off – held
3) turned down – thought of 4) snack in – waited for

84- He did his best to ——— his arm to reach the apple, but he couldn't.
1) raise 2) attach 3) stretch 4) release

85-Two people talking together to reach an agreement can more easily achieve their goal if they are a bit ———.
1) flexible 2) mental 3) economical 4) emotionless



86-The new social behavior among young people is in part ——— by the recent changes in the ——— of family life.
1) launched – layer 2) composed – feature
3) estimated – experience 4) caused – pattern

87- I know you weren't there but, to ———, the meeting was useful.
1) compare 2) summarize 3) organize 4) forecast

88- He could have survived the ——— if he had his seat belt on.
1) accident 2) valley 3) extraction 4) damage

89-She says that living in an urban area or a rural one does not ——— a difference to her.
1) do 2) bring 3) make 4) have

90-The man told me that the two computer programs were ——— and that they could be used ———.
1) necessary – firmly 2) fashionable – immediately
3) voluntary – efficiently 4) similar – interchangeably

**Part C: Cloze Test**

**Directions:** Questions 91-95 are related to the following passage. Read the passage and decide which choice (1), (2), (3), or (4) best fits each space. Then mark your answer sheet.

Parents are advised to read to their children at preschool and early primary levels. When they read to a child from story books, comic strips of the better type, and children's magazines, he becomes (91) ——— that books, magazines, and newspapers hold something (92) ——— and amusement for him. He also comes to (93) ——— that this "something" which he enjoys (94) ——— within black and white symbols, and that these symbols can be discovered only when one knows how to read. This awareness and interest form a place from (95) ——— he can jump into learning to read in school, and into the reading of self-selected books.

| | | | |
|---|---|---|---|
| 91-1) aware | 2) amused | 3) responsible | 4) concerned |
| 92-1) interest | 2) interested | 3) interesting | 4) of interest |
| 93-1) respect | 2) realize | 3) recall | 4) rely |
| 94-1) being included | 2) including | 3) is included | 4) includes |
| 95-1) it | 2) that | 3) which | 4) where |

**Part D: Reading Comprehension**

**Directions: In this part of the test, you will read a passage. The passage is followed by five questions. Answer the questions by choosing the best choice (1), (2), (3), or (4). Then mark your answer sheet.**

**Studies show that rapid increase in population growth creates problems for developing countries. So why don't people have fewer children? The information collected from the developed countries suggests that it is only when people's living standards begin to rise that birth rates begin to fall. There are good reasons for this. Poor countries cannot afford social services and old age pensions – the money they get form the government when they become old and can't work – and people's incomes are so low they have nothing to save for the future. As a result, people look to their children to provide them with financial help in their old age. Having a large family can be a form of insurance. And even while they are still quite young, children can do a lot of useful jobs on a small farm. So poor people in a developing country will need to see clear signs of much better conditions ahead before they can think of having smaller families. But their conditions cannot be improved unless there is a reduction in the rate at which population is increasing. This will depend on a very much wider acceptance of family planning and this, in turn, will mean basic changes in people's ideas.**

**96- What is the best title for the passage?**

1) Reasons for Population Increase
2) Children's Function in Poor countries
3) People's Income in Developing Countries
4) Family Planning Programs in Developing Countries

**97-According to the writer, people in poor countries have a lot of children mainly because ———.**

1) they receive government support for every child
2) they rely on them for support during their old age
3) methods of family planning are not being practiced in those countries
4) they usually become involved with farming, which makes it easier for them to get their food

**98- The writer mentions that, before deciding upon family planning, ———.**

1) people in poor countries should be moved to urban areas
2) governments of poor countries should change people's income and ideas
3) governments of developing countries must forbid parents to use child labor on farms
4) developed countries should help developing ones to improve their standards of living

**99- The word "they" in line 10 refers to ———.**

1) parents 2) people 3) family 4) children

**100-For which of the following has the writer given a definition?**

1) Pension 2) Insurance 3) Family planning 4) Population increase

## Appendix 3: SPE Test Specifications/Description

The following test description is based on the frameworks in Alderson, Clapham, and Wall (1993) and Bachman and Palmer (1996). Where necessary, I will also describe the GE Test to show how the SPE Test developed from that test.

**1. The purpose of the SPE Test**

The SPE Test is a norm-referenced test which has been used to select the most proficient applicants for BA programmes in English literature, Teaching English as a Foreign Language (TEFL) and translation since 2002.

**2. Target language situation**

Examinees will be required to handle advanced oral and written materials of general, and specialised nature i.e. materials in English literature, language teaching, and translation.

**3. Characteristics of the setting**

This national test is usually conducted in schools in towns or cities across the country and a number of foreign countries at the same time. Attempts are made to keep the noise, temperature, humidity, seating conditions, lighting, etc. at an ideal level. It is a paper and pencil test that the students are already familiar with.

**4. Description of the test takers**

Applicants should have high school and pre-university certificates. Normally, they are at the age of 18.

**5. Test level**

Based on my brief conversations with two of the authorities in the National Organisation for Educational Testing (NOET), they claimed that the SPE Test was more difficult than the GE Test. The annual reports of the NOET (2002; 2003; 2004; 2005) also confirmed the higher difficulty of the SPE Test. They showed the national

averages of the students admitted to English departments from 2001-2004 (Table 1). However, they did not show whether the differences between the averages were statistically significant.

**Table 1: Mean scores on the GE and SPE tests of English students admitted to university from 2001-2004**

| Before SPE was introduced | | After SPE was introduced | | | | | |
|---|---|---|---|---|---|---|---|
| 2001 | | 2002 | | 2003 | | 2004 | |
| GE | SPE | GE | SPE | GE | SPE | GE | SPE |
| 72.9 | - | 85 | 51.9 | 74 | 65.5 | 69.9 | 60.4 |

Table 1.3 shows that the national averages on the SPE Test in the years 2002, 2003 and 2004 were lower than the averages on the GE Test. They were also lower than the GE in 2001 when the SPE had not been introduced yet.

## 6. Language elements and skills to be tested

The SPE Test has 6 sections including Vocabulary, Grammar, Reading, Cloze, Language Functions, and Sentence Structure. The first four sections were retained from the GE Test but the last two sections were added as new sections. However, there is no empirical evidence to show whether the SPE Test is measuring different constructs.

## 7. Weighting for each section / paper

### 7.1. The weight of English

Although the four sections of Vocabulary, Grammar, Reading, and Cloze were retained from the GE Test, they were not retained in the same way. The number of items in each of those sections was increased, but grammar items were increased less than the other sections. Table 2 shows the number of items in the retained /common sections as well as the number of items of the new sections.

As Table 2 shows, the SPE Test has 70 items and the GE Test has 25 items. The table shows that the increase in the number of items of the retained sections as well as the

addition of two new sections not only resulted in the increase in the weight of each section but also in the weight of English/SPE Test as a whole.

**Table 2: Number of items in the SPE and GE tests**

| | Skills/Test Sections | GE | SPE |
|---|---|---|---|
| **Retained/common sections** | Vocabulary | 10 | 20 |
| | Reading | 5 | 15 |
| | Cloze | 5 | 15 |
| | Grammar | 5 | 10 |
| **Newly added sections** | Language Functions | 0 | 5 |
| | Sentence Structure | 0 | 5 |
| **Total** | | **25** | **70** |

## 7.2. The weight of Non-English subjects versus English

As part of the University Entrance Exam, some Non-English courses have always been tested along with the SPE and GE tests (Persian literature, Arabic, and theology). In 2002, while the weight of English increased in the SPE Test, the weight of the non-English subjects remained as before. Under the SPE Test, knowledge of English, as opposed to the Non-English subjects, is now a stronger determinant for admission to university than before. Table 3 shows the weighting of English and Non-English subjects in the old and new admission systems.

**Table 3: Weighting of English and Non-English subjects for English and non-English-major applicants in the old and current admission systems**

| | | English | | Non-English | | Total weighting | |
|---|---|---|---|---|---|---|---|
| | | G | S | G | S | English | Non-English |
| **SPE Group (English Major)** | **Old admission system** | 100 | | 225 | | **100** | **225** |
| | **Current admission system** | 50 | 280 | 225 | | **330** | **225** |
| **GE Group (Non-English Majors)** | **Current admission system** | 50 | | 225 | 443.2 | **50** | **668** |

G= General Section, S= Specialised Section

As Table 3 shows, in the current admission systems for the SPE students the weight of English is more than Non-English subjects (330 versus 225 respectively) which includes three general subjects, while for the GE students the weight of Non-English subjects is more than English (668 versus 50). This is because for the GE students Non-English subjects include three general subjects as well as all the specialised

subjects, while for the SPE group Non-English subjects only include the three general courses.

In the old system, 4 points was assigned to GE, but in the new system 2 points are assigned to the GE and 4 points to the SPE. However, despite the fact that the weight of the GE is reduced in the new system, considering that the number of items in SPE is more than that of the GE, the overall weight of English in the new system (330) is more than the old system (100).

**8. Time for each section / paper**

105 minutes is allotted to the SPE and 20 minutes to the GE (Persian literature takes 18 minutes, Arabic 20 minutes, and theology 17 minutes).

**9. Description of suitable language courses or textbooks**

The SPE Test is based on a broader curriculum than the school curriculum i.e. the test is based on the Pre-university textbook, as well as *Bridging the Gap* authored by the NOET and some other books recommended by the Organisation (9.1.3 below). I describe these books briefly in this section.

**9.1. The Pre-university textbook**

The Pre-university textbook is called *Learning to Read for Pre-University Students*. The main aim of the book, as stated in the introduction to this book and as the name of the book suggests, is developing reading skills. However, the book also includes some sections on vocabulary and grammar. Some of the reading tips in the book include skimming for general understanding, paying attention to the topic of the passage, how to record meaning of words in a notebook (e.g. with examples), skipping some words, reading extensively, etc. The book also makes some recommendations about learning vocabulary, e.g. recording the meaning of words along with examples, learning groups of related words, using pictures, using an English-to-English dictionary, etc. However,

at the end of the book there is a word list which might contradict these recommendations, particularly the use of an English-to-English dictionary. A note before the list says to the students, 'You can change this list into a bilingual dictionary by providing Farsi equivalents for the words'. Next, I describe a typical lesson.

### 9.1.1. A typical lesson

Each lesson includes reading, grammar, and vocabulary sections, each of which includes some subsections.

**A. Reading**

1. Pre-reading activities (entitled 'Before You Read') including open questions, statements which the students are asked to rank order, agree, disagree, etc.
2. A reading passage with definitions or synonyms of new words beside the passage
3. Post-reading activities (entitled 'After You Read') including:
   a. True-false questions
   b. Multiple-choice questions involving selection of a summary statement
   c. Sentence completion requiring single-word responses, half a sentence, or a complete sentence
   d. Open questions requiring short answers
   e. Open questions requiring finding particular sentences in the reading passage and then copying them from the passage
   f. Discussion questions
4. 'Sentence Functions' introduces some sentences which express 'reasons', 'advice', etc and then asks the students to provide similar sentences for similar situations based on the reading passage.
5. Tips in Farsi on reading skills such as paragraph headings, explicit-implicit information, guessing the meaning of unfamiliar words, etc along with some examples and exercises

**B. Vocabulary**

The vocabulary section entitled 'Vocabulary Review' includes about five multiple-choice questions for practicing the new words in the reading passage.

**C. Grammar**

1. Introduction of the structure to be taught e.g. conjunctions in adverbial clauses, verb + object + bare infinitive

2. Examples of grammatical structures presented from the reading passage followed by some general questions on the examples e.g. on various meanings of 'as' in different sentences
3. Grammar exercises which may include the following types of items.
   a. Matching items (e.g. matching two parts of a sentence based on 'conjunctions' just introduced)
   b. Fill-in-the-blank items (e.g. with 'when', 'as', 'since', 'whether')
   c. A cloze passage to be completed with the correct forms of certain given words
   d. Sentence combination (e.g. combining two sentences with conjunctions)
   e. Rewriting sentences with the verbs in parentheses
   f. Transforming sentences using the new structures
4. Explanation of grammar with more examples

After each exercise (whether reading, grammar, or vocabulary exercises), students are asked to compare their answers with their partners'.

### 9.1.2. The book, *Bridging the Gap*

*Bridging the Gap* includes fourteen units of reading passages followed by comprehension questions. The purpose of the book, as the introduction to the book states, is to familiarise the English-major applicants with a sample of the books which can bridge the gap between the school English textbook and the level of knowledge expected by the SPE Test and the English departments at universities. The book emphasises that it is not the only book to be studied but only to show the expected level of knowledge. Therefore, the books on which *Bridging the Gap* is based are also recommended to the students. These books are listed in Section C below. At the end of the book there is a list of words for each unit with equivalents and definitions in English as well as in Farsi. The types of items which follow the reading passages are as follows:

1. True false items
2. Multiple-choice items
3. Matching items (vocabulary)
4. Fill-in-the-blank items (vocabulary)
5. Open questions

6. Discussion questions
7. A multiple-choice cloze passage related to the topic of the main passage (vocabulary and grammar)
8. A passage similar to a multiple-choice cloze passage and related to the topic of the main passage where students are asked to choose the correct choice out of three choices in parentheses (vocabulary and grammar)

*Bridging the Gap* also recommends the following methods of learning:

1. Improving English by reading, listening, and speaking
2. Learning words in contexts
3. Guessing the meaning of words based on the general understanding of the passage.

### 9.1.3. The books on which *Bridging the Gap* is based and which are recommended by the NOET


Fowler, W. S. (1997). *First Certificate: Course Book for the Revised Exam.* London: Longman.

Haines, S., and Stewart, B. (1999). *New First Certificate Masterclass: Student's Book.* Oxford: Oxford University Press.

Haines, S., and Stewart, B. (1999). *New First Certificate Masterclass: Workbook.* Oxford: Oxford University Press.

Harrison, M., and Rosalie, R. (1996). *First Certificate Practice Tests: Five tests for the new Cambridge First Certificate in English.* Oxford: Oxford University Press.

Lukey-Coutsocostas, K., and Dalmaris, D. (1999). *Candidate First Certificate Practice Tests* (new edition for the revised FCE). New York: Prentice Hall.

Morris, S., and Stanton, A. (1996). *The Nelson First Certificate: Course Book for the Revised Exam.* New York: Longman.

Morris, S., and Stanton, A. (1999). *The Nelson First Certificate for the Revised Exam: Workbook.* New York: Longman.

Nolasco, R. (1993). *Streetwise: Intermediate Student's Book.* Oxford: Oxford University Press.

Nolasco, R. (1993). *Streetwise: Intermediate Workbook.* Oxford: Oxford University Press.

Nolasco, R. (1993). *Streetwise: Upper-Intermediate Student's Book.* Oxford: Oxford University Press.

Nolasco, R. (1993). *Streetwise: Upper-Intermediate Workbook* Oxford: Oxford University Press.

Page, J., and Naunton, J. (1998). *Think Ahead. - First Certificate Workbook* (new edition). London: Longman.

Powell-Davis, P., and Walker, C. (1995). *Active Comprehension 3.* Oxford: Oxford University Press.

Richards, J. C., Hull, J. and Proctor, S. (1999). *New Interchange 2: English for International Communication, Student's Book.* Cambridge: Cambridge University Press.

Richards, J. C., Hull, J. and Proctor, S. *(1999). New Interchange 2: English for*

*International Communication, Workbook.* Cambridge: Cambridge University Press.
Richards, J. C., Hull, J. and Proctor, S. *(1999). New Interchange 3: English for International Communication, Student's Book.* Cambridge: Cambridge University Press.
Richards, J. C., Hull, J. and Proctor, S. *(1999). New Interchange 3: English for International Communication, Workbook.* Cambridge: Cambridge University Press.
Richards, J.C., and Sandy, C. (2000). *Passages: An Upper-level Multi-Skills Course, Student's Book* 2. Cambridge: Cambridge University Press.
Soars, L., and Soars, J. *(1998). Headway English Course: Intermediate, Student's Book.* Oxford: Oxford University Press.
Soars, L., and Soars, J. *(1998). Headway English Course: Intermediate, Workbook.* Oxford: Oxford University Press.
Soars, L., and Soars, J. *(1998). Headway English Course: Upper-Intermediate, Student's Book.* Oxford: Oxford University Press.
Soars, L., and Soars, J. *(1998). Headway English Course: Upper-Intermediate, Workbook.* Oxford: Oxford University Press.
Soars, L., and Soars, J. *(1998). New Headway English Course: Upper-Intermediate, Student's Book.* Oxford: Oxford University Press.
Soars, L., and Soars, J. *(1998). New Headway English Course: Upper Intermediate, Workbook.* Oxford: Oxford University Press.
Soars, L., and Soars, J. *(1999). New Headway English Course: Intermediate, Student's Book.* Oxford: Oxford University Press.
Soars, L., and Soars, J. *(1999). New Headway English Course: Intermediate, Workbook.* Oxford: Oxford University Press.
Soars, L., and Soars, J. (2000). *New Headway English Course: Pre-Intermediate, Student's Book.* Oxford: Oxford University Press.
Soars, L., and Soars, J. (2000). *New Headway English Course: Pre-Intermediate, Workbook.* Oxford: Oxford University Press.

**10. Rubrics**

The directions in both SPE and GE tests are in English. They are written directions.

**11. Test methods**

All the items in GE and SPE are in multiple-choice forms.

**12. Criteria for marking**

The method of scoring is objective. Due to the limited **capacity of universities, it is** the "norms" that determine admission not "criteria".

**13. Sample papers**

See Appendices 1 and 2

## Appendix 4: A summary of the pilot study results

| | Mean | | Sig |
|---|---|---|---|
| | GE | SPE | |
| **Questionnaire 1** | | | |
| **Learning beliefs on agreement scale** | | | |
| 1. It is easier for children than for adults to learn a foreign language. | 4.18 | 4.02 | .074 |
| 2. Some people have a special ability for learning foreign languages. | 4.11 | 4.10 | .971 |
| 3. Some languages are easier to learn than others. | 4.01 | 4.08 | .396 |
| 4. English is: a) a very difficult language b) a difficult language c) a language of medium difficulty e) an easy language f) a very easy language | 2.90 | 3.21 | .000 |
| 5. I believe that I will learn to speak English very well. | 3.25 | 4.03 | .000 |
| 6. People from my country are good at learning foreign languages. | 3.11 | 3.07 | .127 |
| 7. It is important to speak English with an excellent pronunciation. | 3.81 | 4.21 | .000 |
| 8. It is necessary to know about English-speaking cultures in order to speak English. | 3.53 | 3.76 | .026 |
| 9. You shouldn't say anything in English until you can say it correctly. | 3.97 | 3.98 | .788 |
| 10. It is easier for someone who already speaks a foreign language to learn another one. | 3.59 | 3.50 | .718 |
| 11. People who are good at mathematics or science are not good at learning foreign languages. | 2.46 | 2.43 | .199 |
| 12. It is best to learn English in an English-speaking country. | 4.05 | 3.84 | .127 |
| 13. I enjoy practicing English with the English speakers I meet. | 3.83 | 4.54 | .000 |
| 14. It's o.k. to guess if you don't know a word in English. | 3.94 | 4.11 | .252 |
| 15. If someone spent one hour a day learning a language, how long would it take them to speak the language very well: a) less than a year b) 1-2 years c) 3-5 years e) 5-10 years f) You can't learn a language in 1 hour a day | 2.56 | 2.77 | .094 |
| 16. I have a special ability for learning foreign languages. | 3.03 | 3.80 | .000 |
| 17. The most important part of learning a foreign language is learning vocabulary words. | 1.72 | 1.55 | .013 |
| 18. It is important to repeat and practice a lot. | 4.70 | 4.80 | .155 |
| 19. Women are better than men at learning foreign languages. | 2.98 | 3.23 | .083 |
| 20. People in my country feel that it is important to speak English. | 3.55 | 3.74 | .030 |
| 21. I feel timid speaking English with other people. | 2.97 | 2.69 | .049 |
| 22. If beginning students are permitted to make errors in English, it will be difficult for them to speak correctly later on. | 4.17 | 4.31 | .090 |
| 23. The most important part of learning a foreign language is learning the grammar. | 2.96 | 2.54 | .002 |
| 24. I would like to learn English so that I can get to know English speakers better. | 3.02 | 3.28 | .019 |
| 25. It is easier to speak than understand a foreign language. | 2.97 | 2.81 | .430 |
| 26. It is important to practice with cassettes or tapes. | 3.82 | 4.26 | .000 |
| 27. Learning a foreign language is different than learning other academic subjects. | 3.55 | 3.88 | .008 |
| 28. The most important part of learning English is learning how to translate from my native language. | 2.48 | 2.24 | .047 |
| 29. If I learn English very well, I will have better opportunities for a good job. | 4.29 | 4.43 | .000 |
| 30. People who speak more than one language are very intelligent. | 3.59 | 4.00 | .002 |
| 31. I want to learn to speak English well. | 4.21 | 4.67 | .000 |
| 32. I would like to have English speaking friends. | 3.66 | 4.29 | .000 |
| 33. Everyone can learn to speak a foreign language. | 3.88 | 4.04 | .193 |
| 34. It is easier to read and write English than to speak and understand it. | 3.59 | 3.63 | .872 |
| 35. Language learning involves a lot of memorization. | 3.34 | 3.42 | .080 |
| **Knowledge about the test (choices: Yes, No, I'm not sure)** | | | |
| 1. A section in the SPE includes conversations between two people. | 1.58 | 2.09 | .000 |
| 2. There are equal numbers of items in the different sections. | 1.73 | 1.90 | .069 |
| 3. The Reading Comprehension section includes 2 passages. | 1.64 | 2.11 | .000 |
| 4. Grammar has the same weight as other sections. | 1.84 | 1.82 | .983 |
| 5. The total number of items in the SPE is 70. | 1.23 | 2.01 | .000 |
| 6. Grammar, Vocabulary, and Word Order together carry half of the marks. | 1.52 | 1.65 | .341 |
| 7. The SPE takes 85 minutes to complete | 1.43 | 1.79 | .000 |
| 8. Every item is equally weighted regardless of which section they appear in. | 1.58 | 1.84 | .006 |
| **Attendance in English language institutes and preparation classes** | | | |
| 3. Have you attended an English language institute so far? If so, how many terms? | 6.11 | 10.76 | .000 |
| 9. Attendance in preparation class (only 'yes' answers reported) | 72.1% | 27.6% | .000 |
| **Questionnaire 2 (various opinions about the test on agreement scale)** | | | |
| 1. Preparation classes are useful for the SPE Exam. | 3.99 | 4.46 | .000 |
| 2. The SPE test items are appropriate. | 3.29 | 3.55 | .026 |
| 3. Mock SPE exams can predict the results of the real SPE exam. | 3.45 | 3.54 | .760 |
| 4. Mock SPE exams are an appropriate means of evaluating our learning. | 4.20 | 4.31 | .365 |
| 5. The difficulty level of the SPE exams is the same each year. | 2.81 | 2.88 | .790 |
| 6. Success in the SPE Exam means success in English courses at university. | 3.41 | 3.59 | .207 |
| 7. Preparation for General Courses like Persian is useful for success in the SPE Exam. | 4.13 | 4.24 | .231 |
| 8. Going to classes such as those in an English language institute is useful. | 3.87 | 4.40 | .000 |
| 9. Studying the specified sources for the SPE Exam is enough for preparation. | 3.07 | 3.11 | .877 |
| 10. Background in English is useful for success in the SPE Exam. | 4.04 | 4.16 | .087 |

| | | | |
|---|---|---|---|
| 11. The difficulty levels of the SPE and GE are different. | 3.73 | 4.18 | .000 |
| 12. The SPE Exam is an appropriate means of evaluating our ability in English. | 3.34 | 3.96 | .000 |
| 13. I try to find appropriate ways of test preparation myself. | 3.80 | 3.91 | .239 |
| 14. I'm aware of the purpose of each section of the SPE Exam. | 3.12 | 3.33 | .113 |
| 15. I prefer to use materials with explanation in English rather than those with explanation in Farsi. | 2.62 | 3.37 | .000 |
| 16. Success in the SPE Exam is important for me | 3.68 | 4.80 | .000 |
| 17. Test-taking skills are useful for test preparation. | 4.43 | 4.77 | .000 |
| 18. Effective use of time during the SPE Exam is essential. | 4.61 | 4.80 | .013 |
| 19. I try to reduce the stress that I encounter during test preparation. | 4.35 | 4.50 | .118 |
| 20. When I'm encouraged, I try harder. | 4.32 | 4.60 | .007 |
| 21. The SPE test items are based on pre-specified sources. | 3.40 | 3.37 | .876 |
| 22. The changes which occur in the textbooks are compatible with the purpose of the SPE Exam. | 3.26 | 3.49 | .040 |
| 23. I try to improve my self-confidence. | 4.41 | 4.59 | .020 |
| 24. The time allocated to the SPE Exam is enough. | 3.14 | 3.44 | .027 |
| 25. The difficulty levels of the different sections of the SPE Exam are the same. | 2.83 | 2.80 | .969 |
| 26. I try to learn from my mistakes. | 4.28 | 4.57 | .000 |
| 27. The teaching methods of the teacher is an effective factor in the preparation. | 4.33 | 4.49 | .248 |
| 28. The SPE Exam makes me try. | 3.90 | 4.51 | .000 |
| 29. I study a particular part because I am weak in that part regardless of the marks allocated to it. | 3.73 | 4.27 | .000 |
| 30. I enjoy learning for the SPE Exam. | 3.25 | 4.47 | .000 |
| 31. When I come across items from outside the textbook, I feel anxious. | 3.25 | 3.17 | .002 |
| 32. I try to prepare for the SPE Exam as far as I can. | 4.17 | 4.61 | .000 |
| 33. I don't like to study for the SPE Exam. | 2.75 | 1.99 | .000 |
| 34. I have bad results in English tests because I get too nervous. | | | |
| 35. If I do badly in the SPE I may stop my studies. | 2.10 | 1.81 | .054 |
| 36. I think studying for the SPE Test has improved my English. | 3.66 | 4.36 | .000 |
| 37. I always feel nervous before a test. | 3.24 | 3.18 | .755 |
| 38. I study for those sections of the SPE Exam that have more weight than other sections. | 3.87 | 4.00 | .239 |
| 39. I prefer the teacher to teach in English. | 3.10 | 4.01 | .000 |
| 40. I'm sure that I will succeed on the SPE. | 3.09 | 4.00 | .000 |
| 41. I try to seek teacher's advice on the appropriate ways of test preparation. | 3.90 | 4.36 | .000 |
| 42. I'm afraid of failing the SPE Exam. | 3.42 | 3.32 | .022 |
| 43. When I fail an exam, I work harder. | 3.89 | 4.05 | .078 |
| 44. The different sections of the SPE Exam require equal amount of time. | 2.66 | 2.83 | .072 |
| 45. I study for those sections of the SPE Exam that other students are weak at. | 3.59 | 3.75 | .342 |
| 46. I find many excuses for not studying for the SPE Exam. | 2.75 | 2.01 | .000 |
| 47. I always evaluate my learning. | 3.55 | 4.12 | .000 |
| 48. I encounter a lot of stress during the preparation. | 3.09 | 3.07 | .059 |
| 49. The SPE Exam is a fair Exam. | 2.94 | 3.28 | .001 |
| 50. School classes have a more important role in preparing us for the SPE Exam than private preparation classes. | 3.15 | 2.66 | .000 |
| 51. I always monitor the way I'm preparing. | 3.54 | 3.94 | .000 |
| 52. Ranking item: REPORTED IN A SEPARATE TABLE | | | |
| 53. Open question: REPORTED IN A SEPARATE TABLE | | | |
| 54. I think the SPE Test is(a)very difficult(b)difficult(c)of medium difficulty(d)easy(e)very easy | 3.50 | 3.24 | .000 |
| 55. Open question: REPORTED IN A SEPARATE TABLE | | | |
| 56. Open question: REPORTED IN A SEPARATE TABLE | | | |
| 57. Ranking item: REPORTED IN A SEPARATE TABLE | | | |
| 59. Ranking item: REPORTED IN A SEPARATE TABLE | | | |
| 61. What do you think is your level of motivation? Please tick the appropriate box. | 3.06 | 4.14 | .000 |
| **Questionnaire III (reported activities on frequency scale)** | | | |
| 1. Using instructional CD's | 1.64 | 1.94 | .022 |
| 2. Writing things like diaries in English | 1.83 | 2.26 | .000 |
| 3. Speaking with native speakers | 1.41 | 1.62 | .010 |
| 4 Watching English films | 2.24 | 2.59 | .000 |
| 5. Reading English newspapers and magazines | 1.58 | 2.16 | .000 |
| 6. listening to English programs on the radio | 1.55 | 2.10 | .000 |
| 7. listening to English songs | 2.45 | 2.93 | .000 |
| 8. Using the Internet | 2.39 | 2.42 | .149 |
| 9. learning through games | 2.07 | 2.21 | .008 |
| 10. Using opportunities in class to speak in English | 2.09 | 2.56 | .000 |
| 11. Seeking other people's advice on appropriate ways of learning | 2.62 | 2.94 | .003 |
| 12. Doing homework | 2.97 | 3.31 | .002 |
| 13. Using emails (in English) | 1.96 | 2.24 | .057 |
| 14. Chatting with native speakers on the Internet | 1.89 | 2.03 | .136 |
| 15. Practicing with exam papers of previous years | 2.67 | 2.97 | .015 |

| | | | |
|---|---|---|---|
| 16. Learning examination skills | 2.83 | 3.21 | .000 |
| 17. Attending preparation class | 2.15 | 3.11 | .000 |
| 18. Using materials with explanations completely in English, not in Farsi | 1.68 | 2.51 | .000 |
| 19. Studying *Bridging the Gap* | 1.32 | 2.63 | .000 |
| 20. Watching English programmes on TV | 2.16 | 2.69 | .000 |
| 21. Reading aloud to improve pronunciation and intonation | 2.55 | 3.12 | .000 |
| 22. Using materials that are available in the market | 2.12 | 2.41 | .000 |
| 23. Memorising | 2.85 | 3.02 | .234 |
| 24. Studying high school textbooks of lower levels | 2.54 | 2.81 | .014 |
| 25. Taking mock SPE exams | 2.71 | 2.93 | .018 |
| 26. Getting feedback from teacher | 2.73 | 3.05 | .000 |
| 27. Getting feedback from others | 2.64 | 3.02 | .000 |
| 28. Studying grammar | 3.11 | 3.42 | .002 |
| 29. Studying vocabulary | 3.44 | 3.73 | .000 |
| 30. Studying reading | 3.06 | 3.46 | .000 |
| 31. Studying "general courses" | 3.17 | 3.47 | .006 |
| 32. Going to English institutes | 1.98 | 2.88 | .000 |
| 33. Using sources with explanations in Farsi | 2.65 | 2.87 | .057 |
| 34. Answering sample exam papers under timed conditions similar to the real exam | 2.65 | 3.02 | .001 |
| 35. Thinking about appropriate ways of learning | 2.78 | 3.10 | .006 |
| 36. Studying the high school textbook | 3.44 | 3.57 | .244 |
| 37. Talking to classmates or friends in English | 1.69 | 2.14 | .000 |
| 39. Translating reading texts into Farsi | 3.08 | 3.31 | .030 |
| 40. Increasing my vocabulary | 3.09 | 3.41 | .000 |
| 41. Learning grammar | 2.96 | 3.32 | .000 |
| 42. Trying to guess the meaning of the new words from context | 2.86 | 3.23 | .000 |
| 43. Practicing speed reading | 2.98 | 3.35 | .000 |
| 44. Skimming the text first to get the main idea and then going back to read it more carefully | 3.04 | 3.30 | .007 |
| 45. Thinking in English while reading | 2.22 | 2.92 | .000 |
| 46. Practicing with multiple-choice reading passages | 2.67 | 3.08 | .000 |
| 47. Reading English story books | 1.86 | 2.75 | .000 |
| 48. Trying to improve concentration | 2.65 | 2.90 | .012 |
| 49. Using comprehension questions to get a general idea of the passage | 2.68 | 3.02 | .001 |
| 50. Taking notes of major points while reading | 2.63 | 3.09 | .000 |
| 51. Reading for pleasure | 2.06 | 2.59 | .000 |
| 52. Reading without looking up every single word | 2.28 | 2.66 | .002 |
| 53. Predicting what will come next in a reading text | 2.41 | 2.66 | .012 |
| 54. Reading as much as possible in English | 2.32 | 2.80 | .000 |
| 55. Keeping in mind what has been read while reading | 3.03 | 3.33 | .003 |
| 56. Paying attention to the topic of the text | 3.20 | 3.43 | .007 |
| 57. Practicing with cloze passages | 2.44 | 2.97 | .000 |
| 58. Paying attention to and analyzing the structure and organization of the text | 2.53 | 2.91 | .001 |
| 60. Using Leitner Box. | 1.59 | 1.89 | .024 |
| 61. Memorising meanings | 3.32 | 3.47 | .051 |
| 62. Learning vocabulary through listening | | | |
| 63. Learning meanings through examples in which the new word is used | 2.94 | 3.20 | .001 |
| 64. Learning vocabulary through watching films | | | |
| 65. Reading passages | 2.83 | 3.20 | .000 |
| 66. Using words in sentences that I make | 2.21 | 2.75 | .000 |
| 67. Reading story books | 1.82 | 2.36 | .000 |
| 68. Paying attention to the parts of speech (noun, verb, adjective, adverbs). | 2.32 | 2.82 | .000 |
| 69. Paying attention to suffixes and prefixes. | 2.38 | 2.79 | .000 |
| 70. Grouping words according to relation in meaning | 2.54 | 3.08 | .000 |
| 71.Practising | 3.20 | 3.50 | .000 |
| 72. Relating new words with previously known words | 2.77 | 3.19 | .000 |
| 73. Making a mental image of the situation in which the new word is used | 2.53 | 3.08 | .000 |
| 74. Associating the sound of the new word with the image of the word | 2.35 | 2.65 | .001 |
| 75. Using rhyming to remember new words (e.g. rice and ice) | 1.79 | 2.08 | .001 |
| 76. Writing the new words several times | 2.83 | 3.04 | .113 |
| 78. Comparing the grammar rules in English with grammar rules in Farsi | 2.50 | 2.47 | .268 |
| 79. Learning and using grammar rules in new situations | 2.65 | 2.81 | .005 |
| 80. Memorizing grammatical rules | 3.01 | 3.03 | .606 |
| 81. Using grammar rules even if I don't know them exactly | 2.17 | 2.26 | .566 |
| 82. Analysing grammar rules in texts | 2.62 | 3.11 | .000 |
| 83. Answering sample questions | 2.97 | 3.35 | .001 |
| 84. Summarising grammar rules | 2.88 | 3.11 | .068 |
| 85. Learning grammatical clues | 2.89 | 3.19 | .001 |

**Results of open-ended questions and ranking questions:**

**Most important skill**

52. Which of the following language skills are more essential to your success than other skills? Please rank 5 of them according to importance to your success (1=most important, 5=least important).

| | SPE | GE | Sig |
|---|---|---|---|
| reading | 3.27 | 3.46 | .639 |
| listening | 4.58 | 4.47 | .890 |
| speaking | 5.06 | 4.80 | .452 |
| writing | 4.96 | 5.12 | .697 |
| grammar | 3.08 | 3.40 | .264 |
| vocabulary | 1.50 | 1.51 | .988 |
| pronunciation | 4.34 | 4.24 | .424 |

| ranks | SPE | GE |
|---|---|---|
| 1 | vocabulary | Vocabulary |
| 2 | grammar | Grammar |
| 3 | reading | Reading |
| 4 | pronunciation | pronunciation |
| 5 | listening | listening |
| 6 | writing | Speaking |
| 7 | speaking | Writing |

**The most difficult test sections**

53. Which section of the SPE Exam is the most difficult? Please write your answer in the space provided.

**Percentage of the SPE and GE students ranking test sections from most difficult to least difficult**

| | % | |
|---|---|---|
| **Test sections** | **SPE** | **GE** |
| grammar | 18.9 | 14.9 |
| Vocabulary | 5.4 | 11.8 |
| Reading | 27.0 | 25.9 |
| Cloze | 2.7 | 2.6 |
| Speaking | 6.3 | 6.0 |
| Pronunciation | .9 | .6 |
| Writing | .9 | |
| Listening | .9 | 1.1 |
| All | | 2.9 |
| Grammar & vocabulary | .9 | |
| Grammar & reading | | 1.1 |
| Grammar & cloze | 3.6 | |
| Grammar & speaking | | .3 |
| Vocabulary & reading | .9 | .6 |
| Vocabulary & pronunciation | | .3 |
| Reading & grammar | 2.7 | |
| Reading & cloze | .9 | |
| Reading & speaking | .9 | |
| Cloze & reading | | 1.1 |
| Speaking & vocabulary | .9 | |
| Speaking & reading | | .3 |

| | | |
|---|---|---|
| Speaking & pronunciation | | .3 |
| Pronunciation & writing | | .3 |
| **Valid** | 73.9 | 70.1 |
| **Missing** | 26.1 | 29.9 |
| **Total** | 100.0 | 100.0 |
| **Significance** | .072 | |

## Materials

55. Do you learn some sections of the SPE Exam ONLY through test papers?

| Yes | No | If your answer was "yes": | |
|---|---|---|---|
| | | Which section(s)? | |

**Sections for which test papers are used as materials**

| | % | |
|---|---|---|
| **Test sections** | **SPE** | **GE** |
| No | 91.9 | 87.4 |
| grammar | .9 | 2.9 |
| Vocabulary | | 1.1 |
| Reading | .9 | 2.0 |
| Cloze | | .6 |
| pronunciation | | |
| All | .9 | 1.1 |
| Grammar & vocabulary | .9 | 1.4 |
| Grammar & reading | | .3 |
| Vocabulary & reading | | .3 |
| Reading &grammar | 2.7 | .3 |
| Reading & vocabulary | | .3 |
| Cloze & reading | | .3 |
| **Valid** | 98.2 | 98.0 |
| **Missing** | 1.8 | 2.0 |
| **Total** | 100.0 | 100.0 |
| **Significance** | .345 | |

## Time spent on test sections

56. On which section of the SPE Test do you spend the most time **in the test preparation**? Please write your answer in the space provided.

**Time spent on each section in preparation from most to least**

| | % | |
|---|---|---|
| **Test sections** | **SPE** | **GE** |
| grammar | 13.5 | 7.5 |
| Vocabulary | 13.5 | 30.5 |
| Reading | 17.1 | 13.5 |
| Cloze | | |
| Speaking | .9 | .3 |
| Writing | | |
| All | 3.6 | 2.9 |
| Grammar & vocabulary | 6.3 | 4.9 |
| Grammar & reading | | .9 |
| Grammar & cloze | 1.8 | .3 |
| Vocabulary & grammar | .9 | |
| Vocabulary & reading | .9 | 3.7 |

| | | |
|---|---|---|
| Vocabulary & pronunciation | .9 | .3 |
| Reading & grammar | 3.6 | |
| Reading & vocabulary | 1.8 | |
| Cloze & grammar | | .3 |
| Cloze & reading | | .6 |
| Speaking & grammar | | .3 |
| Speaking & reading | | .3 |
| Pronunciation & vocabulary | | |
| **Valid** | 64.9 | 66.1 |
| **Missing** | 35.1 | 33.9 |
| **Total** | 100.0 | 100.0 |
| **Significance** | .338 | |

## Reasons for learning English

57. Please rank 5 of your reasons for learning English in terms of the degree of importance: (1=the most important, 5=the least important).

**Reasons for learning English ranked in terms of importance**

| | SPE | GE | Sig |
|---|---|---|---|
| English sources | 2.59 | 2.41 | .568 |
| Job | 2.70 | 3.21 | .009 |
| Exam | 2.28 | 2.67 | .005 |
| Communication | 2.92 | 3.06 | .652 |
| Parents | 4.25 | 4.19 | .934 |
| Cultural products | 3.80 | 3.26 | .039 |
| prestige | 3.90 | 3.52 | .254 |

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| **SPE** | Exam | English sources | Job | Communication | Cultural products | prestige | Parents |
| **GE** | English sources | Exam | Communication | Job | Cultural products | prestige | Parents |

## Factors for success

59. Please rank 5 of the following factors for success in the SPE Exam in terms of the degree of importance: (1=the most important, 5=the least important).

**Average rankings of factors for success**

| | SPE | GE | Sig |
|---|---|---|---|
| motivation | 1.64 | 1.75 | .560 |
| effort | 2.17 | 2.31 | .435 |
| self-confidence | 2.59 | 2.76 | .149 |
| teaching methods | 3.68 | 3.27 | .245 |
| how to learn | 3.34 | 3.40 | .918 |
| Luck | 5.64 | 4.80 | .225 |
| background in English | 3.98 | 3.84 | .049 |
| difficulty | 4.96 | 4.76 | .814 |
| preparation class | 4.08 | 4.02 | .376 |
| school class | 5.64 | 4.03 | .028 |
| general courses | 4.66 | 5.08 | .727 |
| regional quota | 6.63 | 5.02 | .113 |

| ranks | SPE | GE |
|---|---|---|
| 1 | motivation | motivation |
| 2 | effort | effort |
| 3 | self-confidence | self-confidence |
| 4 | how to learn | teaching methods |
| 5 | teaching methods | how to learn |
| 6 | background in English | background in English |
| 7 | preparation class | preparation class |
| 8 | general courses | school class |
| 9 | difficulty | difficulty |
| 10 | school class | luck |
| 11 | luck | regional quota |
| 12 | regional quota | general courses |

## Appendix 5: Test-retest reliability of the pilot study questionnaires

| Items | GE | SPE | Sig GE | Sig SPE |
|---|---|---|---|---|
| **Questionnaire 1: Learning beliefs & knowledge of the test** | | | | |
| 1. It is easier for children than for adults to learn a foreign language. | .901** | .509** | .000 | .001 |
| 2. Some people have a special ability for learning foreign languages. | .691** | .741** | .000 | .000 |
| 3. Some languages are easier to learn than others. | .671** | .710** | .000 | .000 |
| 4. English is:<br>a) a very difficult language b) a difficult language c) a language of medium difficulty e) an easy language f) a very easy language | .878** | .729** | .000 | .000 |
| 5. I believe that I will learn to speak English very well. | .735** | .494** | .000 | .002 |
| 6. People from my country are good at learning foreign languages. | .671** | .693** | .000 | .000 |
| 7. It is important to speak English with an excellent pronunciation. | .895** | .542** | .000 | .000 |
| 8. It is necessary to know about English-speaking cultures in order to speak English. | .783** | .874** | .000 | .000 |
| 9. You shouldn't say anything in English until you can say it correctly. | .839** | .694** | .000 | .000 |
| 10. It is easier for someone who already speaks a foreign language to learn another one. | .733** | .671** | .000 | .000 |
| 11. People who are good at mathematics or science are not good at learning foreign languages. | .736** | .515** | .000 | .001 |
| 12. It is best to learn English in an English-speaking country. | .770** | .746** | .000 | .000 |
| 13. I enjoy practicing English with the English speakers I meet. | .816** | .625** | .000 | .000 |
| 14. It's o.k. to guess if you don't know a word in English. | .737** | .423** | .000 | .007 |
| 15. If someone spent one hour a day learning a language, how long would it take them to speak the language very well: a) less than a year b) 1-2 years c) 3-5 years e) 5-10 years f) You can't learn a language in 1 hour a day | .996** | .893** | .000 | .000 |
| 16. I have a special ability for learning foreign languages. | .853** | .628** | .000 | .000 |
| 17. The most important part of learning a foreign language is learning vocabulary words. | .807** | .504** | .000 | .001 |
| 18. It is important to repeat and practice a lot. | .787** | .677** | .000 | .000 |
| 19. Women are better than men at learning foreign languages. | .874** | .655** | .000 | .000 |
| 20. People in my country feel that it is important to speak English. | .871** | .654** | .000 | .000 |
| 21. I feel timid speaking English with other people. | .881** | .835** | .000 | .000 |
| 22. If beginning students are permitted to make errors in English, it will be difficult for them to speak correctly later on. | .697** | .572** | .000 | .000 |
| 23. The most important part of learning a foreign language is learning the grammar. | .843** | .532** | .000 | .001 |
| 24. I would like to learn English so that I can get to know English speakers better. | .745** | .697** | .000 | .000 |
| 25. It is easier to speak than understand a foreign language. | .880** | .710** | .000 | .000 |
| 26. It is important to practice with cassettes or tapes. | .893** | .632** | .000 | .000 |
| 27. Learning a foreign language is different than learning other academic subjects. | .457** | .645** | .007 | .000 |
| 28. The most important part of learning English is learning how to translate from my native language. | .858** | .648** | .000 | .000 |
| 29. If I learn English very well, I will have better opportunities for a good job. | .738** | .611** | .000 | .000 |
| 30. People who speak more than one language are very intelligent. | .790** | .820** | .000 | .000 |
| 31. I want to learn to speak English well. | .394* | .468** | .021 | .003 |
| 32. I would like to have English speaking friends. | .697** | .783** | .000 | .000 |
| 33. Everyone can learn to speak a foreign language. | .592** | .348* | .000 | .030 |
| 34. It is easier to read and write English than to speak and understand it. | .732** | .513** | .000 | .001 |
| 35. Language learning involves a lot of memorization. | .705** | .668** | .000 | .000 |
| *Knowledge about the SPE Test* | | | | |
| 1. A section in the SPE includes conversations between two people. | .919** | .800** | .000 | .000 |
| 2. There are equal numbers of items in the different sections. | .727** | .690** | .000 | .000 |
| 3. The Reading Comprehension section includes 2 passages. | .734** | .794** | .000 | .000 |
| 4. Grammar has the same weight as other sections. | .871** | .485** | .000 | .003 |
| 5. The total number of items in the SPE is 70. | .978** | .804** | .000 | .000 |
| 6. Grammar, Vocabulary, and Word Order together carry half of the marks. | .819** | .578** | .000 | .000 |
| 7. The SPE takes 85 minutes to complete | .882** | .648** | .000 | .000 |
| 8. Every item is equally weighted regardless of which section they appear in. | .768** | .692** | .000 | .000 |
| | | | | |
| **Questionnaire 2: Attitudes about the Test** | | | | |
| 1. Preparation classes are useful for the SPE Exam. | .780** | .624** | .000 | .000 |
| 2. The SPE test items are appropriate. | .877** | .759** | .000 | .000 |
| 3. Mock SPE exams can predict the results of the real SPE exam. | .572** | .771** | .000 | .000 |
| 4. Mock SPE exams are an appropriate means of evaluating our learning. | .819** | .677** | .000 | .000 |
| 5. The difficulty level of the SPE exams is the same each year. | .701** | .634** | .000 | .000 |
| 6. Success in the SPE Exam means success in English courses at university. | .811** | .467** | .000 | .003 |
| 7. Preparation for General Courses like Persian is useful for success in the SPE Exam. | .859** | .391* | .000 | .014 |
| 8. Going to classes such as those in an English language institute is useful. | .863** | .573** | .000 | .000 |
| 9. Studying the specified sources for the SPE Exam is enough for preparation. | .517** | .676** | .001 | .000 |
| 10. Background in English is useful for success in the SPE Exam. | .643** | .458** | .000 | .003 |
| 11. The difficulty levels of the SPE and GE are different. | .713** | .512** | .000 | .001 |

| Items | GE | SPE | Sig GE | Sig SPE |
|---|---|---|---|---|
| 12. The SPE Exam is an appropriate means of evaluating our ability in English. | .840** | .512** | .000 | .001 |
| 13. I try to find appropriate ways of test preparation myself. | .779** | .639** | .000 | .000 |
| 14. I'm aware of the purpose of each section of the SPE Exam. | .534** | .642** | .001 | .000 |
| 15. I prefer to use materials with explanation in English rather than those with explanation in Farsi. | .841** | .769** | .000 | .000 |
| 16. Success in the SPE Exam is important for me | .704** | .300 | .000 | .064 |
| 17. Test- taking skills is useful for test preparation. | .758** | .599** | .000 | .000 |
| 18. Effective use of time during the SPE Exam is essential. | .430** | .410** | .010 | .009 |
| 19. I try to reduce the stress that I encounter during test preparation. | .482** | .452** | .003 | .004 |
| 20. When I'm encouraged, I try harder. | .688** | .442** | .000 | .005 |
| 21. The SPE test items are based on pre-specified sources. | .731** | .639** | .000 | .000 |
| 22. The changes which occur in the textbooks are compatible with the purpose of the SPE Exam. | .894** | .700** | .000 | .000 |
| 23. I try to improve my self-confidence. | .709** | .447** | .000 | .004 |
| 24. The time allocated to the SPE Exam is enough. | .876** | .656** | .000 | .000 |
| 25. The difficulty levels of the different sections of the SPE Exam are the same. | .808** | .651** | .000 | .000 |
| 26. I try to learn from my mistakes. | .687** | .540** | .000 | .001 |
| 27. The teaching methods of the teacher is an effective factor in the preparation. | .675** | .421** | .000 | .008 |
| 28. The SPE Exam makes me try. | .720** | .569** | .000 | .000 |
| 29. I study a particular part because I am weak in that part regardless of the marks allocated to it. | .865** | .650** | .000 | .000 |
| 30. I enjoy learning for the SPE Exam. | .943** | .539** | .000 | .000 |
| 31. When I come across items from outside the textbook, I feel anxious. | .747** | .530** | .000 | .001 |
| 32. I try to prepare for the SPE Exam as far as I can. | .636** | .354* | .000 | .027 |
| 33. I don't like to study for the SPE Exam. | .833** | .218 | .000 | .194 |
| 34. I have bad results in English tests because I get too nervous. | .774** | .643** | .000 | .000 |
| 35. If I do badly in the SPE, I may stop my studies. | .728** | .324* | .000 | .044 |
| 36. I think studying for the SPE Test has improved my English. | .804** | .496** | .000 | .002 |
| 37. I always feel nervous before a test. | .814** | .866** | .000 | .000 |
| 38. I study for those sections of the SPE Exam that have more weight than other sections. | .722** | .582** | .000 | .000 |
| 39. I prefer the teacher to teach in English. | .867** | .550** | .000 | .000 |
| 40. I'm sure that I will succeed on the SPE. | .838** | .739** | .000 | .000 |
| 41. I try to seek teacher's advice on the appropriate ways of test preparation. | .814** | .595** | .000 | .000 |
| 42. I'm afraid of failing the SPE Exam. | .811** | .514** | .000 | .001 |
| 43. When I fail an exam, I work harder. | .654** | .321* | .000 | .049 |
| 44. The different sections of the SPE Exam require equal amount of time. | .745** | .432** | .000 | .007 |
| 45. I study for those sections of the SPE Exam that other students are weak at. | .827** | .629** | .000 | .000 |
| 46. I find many excuses for not studying for the SPE Exam. | .853** | .476** | .000 | .003 |
| 47. I always evaluate my learning. | .718** | .379* | .000 | .019 |
| 48. I encounter a lot of stress during the preparation. | .872** | .571** | .000 | .000 |
| 49. The SPE Exam is a fair Exam. | .873** | .842** | .000 | .000 |
| 50. School classes have a more important role in preparing us for the SPE Exam than private preparation classes. | .839** | .648** | .000 | .000 |
| 51. I always monitor the way I'm preparing. | .748** | .473** | .000 | .002 |
| 52. Which of the following language skills are more essential to your success than other skills? Please rank 5 of them according to importance to your success (1=most important, 5=least important). | | | . | . |
| II52. Reading | .614** | .686** | .000 | .000 |
| II52. Listening | .494** | .561** | .005 | .002 |
| II52. Speaking | .538** | .753** | .003 | .000 |
| II52. Writing | .710** | .268 | .000 | .144 |
| II52. Grammar | .821** | .515** | .000 | .001 |
| II52. Vocabulary | .746** | .432** | .000 | .007 |
| II52. Pronunciation | .775** | .615** | .000 | .000 |
| 54. I think the SPE Test is (a) very difficult (b) difficult (c) of medium difficulty (d) easy (e) very easy | .932** | .622** | .000 | .000 |
| 57. Please rank 5 of your reasons for learning English in terms of the degree of importance: (1= the most important, 5= the least important). | | | | |
| II57. Using English resources | .688** | .371* | .000 | .034 |
| II57. Getting a job | .685** | .557** | .000 | .001 |
| II57. Passing the exam | .656** | .427* | .001 | .017 |
| II57. Communication | .673** | .791** | .000 | .000 |
| II57. Fulfilling parents' wishes | .951** | .974** | .000 | .000 |
| II57. Understanding cultural products such as films, arts, literature, etc | .848** | .755** | .000 | .000 |
| II57. For social status | .780** | .557** | .000 | .003 |
| 59. Please rank 5 of the following factors for success in the SPE Exam in terms of the degree of importance: (1= the most important, 5= the least important). | | | | |

| Items | GE | SPE | Sig GE | Sig SPE |
|---|---|---|---|---|
| II59. Motivation | .834** | .521** | .000 | .001 |
| II59. Effort | .735** | .487** | .000 | .003 |
| II59. Self-confidence | .768** | .560** | .000 | .001 |
| II59. Teaching methods | .816** | .657** | .000 | .006 |
| II59. Knowing how to learn | .698** | .769** | .000 | .000 |
| II59. Luck | .929** | .752** | .002 | .012 |
| II59. English background | .726** | .588** | .001 | .002 |
| II59. Difficulty level of the test | .746 | .958** | .088 | .000 |
| II59. Preparation class | .710 | .869** | .074 | .001 |
| II59. School classes | .929** | .708* | .000 | .010 |
| II59. General subjects | .746 | .695* | .088 | .018 |
| II59. Regional quota | .974** | 1.000** | .000 | . |
| 61. What do you think is your level of motivation? Please tick the appropriate box. | .817** | .810** | .000 | .000 |
| 62. Which of the following General subjects do you try to learn more than other subjects in order to be admitted to the university of your choice? Please rank them according to importance to your success (1=most important, 4=least important). | | | | |
| II62. English | .985** | .684** | .000 | .000 |
| II62. Persian literature | .827** | .657** | .000 | .000 |
| II62. Arabic | .783** | .935** | .000 | .000 |
| II62. Theology | .853** | .721** | .000 | .000 |
| | | | | |
| **Questionnaire 3: Reported activities** | | | | |
| 1. Using instructional CD's | .814** | .609** | .000 | .000 |
| 2. Writing things like diaries in English | .805** | .554** | .000 | .000 |
| 3. Speaking with native speakers | .770** | .640** | .000 | .000 |
| 4. Watching English films | .901** | .669** | .000 | .000 |
| 5. Reading English newspapers and magazines | .784** | .823** | .000 | .000 |
| 6. listening to English programs on the radio | .850** | .731** | .000 | .000 |
| 7. listening to English songs | .899** | .734** | .000 | .000 |
| 8. Using the Internet | .968** | .837** | .000 | .000 |
| 9. Learning through games | .802** | .509** | .000 | .001 |
| 10. Using opportunities in class to speak in English | .852** | .727** | .000 | .000 |
| 11. Seeking other people's advice on appropriate ways of learning | .726** | .741** | .000 | .000 |
| 12. Doing homework | .959** | .740** | .000 | .000 |
| 13. Using emails | .841** | .733** | .000 | .000 |
| 14. Using chat rooms | .867** | .724** | .000 | .000 |
| 15. Practicing with exam papers of previous years | .750** | .590** | .000 | .000 |
| 16. Learning examination skills | .865** | .593** | .000 | .000 |
| 17. Attending preparation class | .886** | .867** | .000 | .000 |
| 18. Using materials with explanations completely in English, not in Farsi | .836** | .541** | .000 | .000 |
| 19. Studying *Bridging the Gap* | .649** | .926** | .000 | .000 |
| 20. Watching English programmes on TV | .886** | .639** | .000 | .000 |
| 21. Reading aloud to improve pronunciation and intonation | .799** | .736** | .000 | .000 |
| 22. Using materials that are available in the market | .637** | .440** | .000 | .006 |
| 23. Memorizing | .707** | .657** | .000 | .000 |
| 24. Studying high school textbooks of lower levels | .863** | .823** | .000 | .000 |
| 25. Taking mock SPE exams | .756** | .803** | .000 | .000 |
| 26. Getting feedback from teacher | .846** | .560** | .000 | .000 |
| 27. Getting feedback from others | .649** | .608** | .000 | .000 |
| 28. Studying grammar | .750** | .681** | .000 | .000 |
| 29. Studying vocabulary | .826** | .537** | .000 | .000 |
| 30. Studying reading | .736** | .677** | .000 | .000 |
| 31. Studying "general courses" | .497** | .627** | .003 | .000 |
| 32. Going to English institutes | .935** | .621** | .000 | .000 |
| 33. Using sources with explanations in Farsi | .741** | .865** | .000 | .000 |
| 34. Answering sample exam papers under timed conditions similar to the real exam | .936** | .642** | .000 | .000 |
| 35. Thinking about appropriate ways of learning | .757** | .748** | .000 | .000 |
| 36. Studying the high school textbook | .606** | .758** | .000 | .000 |
| 37. Talking to classmates or friends in English | .861** | .594** | .000 | .000 |
| 39. Translating reading texts into Farsi | .682** | .558** | .000 | .000 |
| 40. Increasing my vocabulary | .606** | .366* | .000 | .000 |
| 41. Learning grammar | .770** | .575** | .000 | .000 |
| 42. Trying to guess the meaning of the new words from context | .720** | .492** | .000 | .001 |
| 43. Practicing speed reading | .682** | .641** | .000 | .000 |
| 44. Skimming the text first to get the main idea and then going back to read it more carefully | .524** | .644** | .001 | .000 |
| 45. Thinking in English while reading | .888** | .746** | .000 | .000 |
| 46. Practicing with multiple-choice reading passages | .760** | .690** | .000 | .000 |

| Items | GE | SPE | Sig GE | Sig SPE |
|---|---|---|---|---|
| 47. Reading English story books | .974** | .840** | .000 | .000 |
| 48. Trying to improve concentration | .738** | .702** | .000 | .000 |
| 49. Using comprehension questions to get a general idea of the passage | .706** | .416** | .000 | .009 |
| 50. Taking notes of major points while reading | .688** | .546** | .000 | .000 |
| 51. Reading for pleasure | .847** | .755** | .000 | .000 |
| 52. Reading without looking up all the words I don't know | .849** | .524** | .000 | .001 |
| 53. Predicting and guessing what will come next | .701** | .657** | .000 | .000 |
| 54. Reading as much as possible in English | .917** | .444** | .000 | .005 |
| 55. Keeping in mind what has been read while reading | .774** | .310* | .000 | .054 |
| 56. Paying attention to the topic of the text | .400* | .443** | .019 | .005 |
| 57. Practicing with cloze passages | .891** | .790** | .000 | .000 |
| 58. Paying attention to the organization of the text | .712** | .213 | .000 | .200 |
| 60. Using Leitner Box | .873** | .576** | .000 | .000 |
| 61. Memorising meanings | .712** | .510** | .000 | .001 |
| 62. Listening | .585** | .382* | .000 | .023 |
| 63. Learning meanings through examples in which the new word is used | .749** | .637** | .000 | .000 |
| 64. Watching films | .845** | .871** | .000 | .000 |
| 65. Reading passages | .767** | .329* | .000 | .044 |
| 66. Using words in sentences that I make | .897** | .637** | .000 | .000 |
| 67. Reading story books | .957** | .735** | .000 | .000 |
| 68. Paying attention to the parts of speech (noun, verb, adjective, adverbs). | .903** | .661** | .000 | .000 |
| 69. Paying attention to suffixes and prefixes. | .744** | .505** | .000 | .001 |
| 70. Grouping words according to relation in meaning | .674** | .575** | .000 | .000 |
| 71. Practice | .823** | .271 | .000 | .104 |
| 72. Relating new words with previously known words | .622** | .575** | .000 | .000 |
| 73. Making a mental image of the situation in which the new word is used | .630** | .681** | .000 | .000 |
| 74. Associating the sound of the new word with the image of the word | .821** | .568** | .000 | .000 |
| 75. Using rhyming to remember new words (e.g. rice and ice) | .826** | .761** | .000 | .000 |
| 76. Writing the new words several times | .924** | .750** | .000 | .000 |
| 78. Comparing the grammar rules in English with grammar rules in Farsi | .886** | .647** | .000 | .000 |
| 79. Learning and using grammar rules in new situations | .816** | .442** | .000 | .008 |
| 80. Memorizing grammatical rules | .670** | .674** | .000 | .000 |
| 81. Using grammar rules even if I don't know them exactly | .761** | .515** | .000 | .002 |
| 82. Analysing grammar rules in texts | .803** | .787** | .000 | .000 |
| 83. Answering sample questions | .877** | .641** | .000 | .000 |
| 84. Summarising grammar rules | .592** | .775** | .000 | .000 |
| 85. Learning grammatical clues | .669** | .479** | .000 | .004 |
| **Average correlation of each group** | **.77** | **.62** | | |

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

## Appendix 6: Internal consistency between overlapping items of the pilot study questionnaires

**Note:** *= Correlation is significant at the 0.05 level (2-tailed), **= Correlation is significant at the 0.01 level (2-tailed).

**English background/Institute**

3. Have you attended an English language institute so far? If so, how many terms?
8. Going to classes such as those in an English language institute is useful.
10. Background in English is useful for success in the SPE Exam.
32. Going to English institutes

**Correlations**

| | | 3 | 8 | 10 | 32 |
|---|---|---|---|---|---|
| 3 | Correlation Coefficient | 1.000 | .224(*) | .193 | .303(**) |
| | Sig. (2-tailed) | . | .023 | .051 | .004 |
| | N | 104 | 103 | 103 | 89 |
| 8 | Correlation Coefficient | .224(*) | 1.000 | .257(**) | .270(**) |
| | Sig. (2-tailed) | .023 | . | .007 | .008 |
| | N | 103 | 110 | 110 | 95 |
| 10 | Correlation Coefficient | .193 | .257(**) | 1.000 | .020 |
| | Sig. (2-tailed) | .051 | .007 | . | .846 |
| | N | 103 | 110 | 110 | 95 |
| 32 | Correlation Coefficient | .303(**) | .270(**) | .020 | 1.000 |
| | Sig. (2-tailed) | .004 | .008 | .846 | . |
| | N | 89 | 95 | 95 | 95 |

Group = 1

**Correlations**

| | | 3 | 8 | 10 | 32 |
|---|---|---|---|---|---|
| 3 | Correlation Coefficient | 1.000 | .297(**) | .257(**) | .502(**) |
| | Sig. (2-tailed) | . | .000 | .000 | .000 |
| | N | 332 | 323 | 328 | 276 |
| 8 | Correlation Coefficient | .297(**) | 1.000 | .242(**) | .299(**) |
| | Sig. (2-tailed) | .000 | . | .000 | .000 |
| | N | 323 | 338 | 336 | 280 |
| 10 | Correlation Coefficient | .257(**) | .242(**) | 1.000 | .103 |
| | Sig. (2-tailed) | .000 | .000 | . | .083 |
| | N | 328 | 336 | 344 | 286 |
| 32 | Correlation Coefficient | .502(**) | .299(**) | .103 | 1.000 |
| | Sig. (2-tailed) | .000 | .000 | .083 | . |
| | N | 276 | 280 | 286 | 287 |

Group = 2

**Preparation class**

9. If you are attending preparation classes, please mark the type of the class(es).

II.1. Preparation classes are useful for the SPE Exam.

III.17. Attending preparation class

**Correlations**

| | | 9 | 1 | 17 |
|---|---|---|---|---|
| 9 | Correlation Coefficient | 1.000 | -.019 | .373(**) |
| | Sig. (2-tailed) | . | .846 | .000 |
| | N | 106 | 106 | 96 |
| 1 | Correlation Coefficient | -.019 | 1.000 | .217(*) |
| | Sig. (2-tailed) | .846 | . | .030 |
| | N | 106 | 110 | 100 |
| 17 | Correlation Coefficient | .373(**) | .217(*) | 1.000 |
| | Sig. (2-tailed) | .000 | .030 | . |
| | N | 96 | 100 | 100 |

Group = 1

**Correlations**

| | | 9 | 1 | 17 |
|---|---|---|---|---|
| 9 | Correlation Coefficient | 1.000 | .256(**) | .720(**) |
| | Sig. (2-tailed) | . | .000 | .000 |
| | N | 317 | 314 | 270 |
| 1 | Correlation Coefficient | .256(**) | 1.000 | .393(**) |
| | Sig. (2-tailed) | .000 | . | .000 |
| | N | 314 | 344 | 292 |
| 17 | Correlation Coefficient | .720(**) | .393(**) | 1.000 |
| | Sig. (2-tailed) | .000 | .000 | . |
| | N | 270 | 292 | 294 |

Group = 2

**Anxiety**

21. I feel timid speaking English with other people.
34. I have bad results in English tests because I get too nervous.
37. I always feel nervous before a test.
42. I'm afraid of failing the SPE Exam.
48. I encounter a lot of stress during the preparation.

**Correlations**

| | | 21 | 34 | 37 | 42 | 48 |
|---|---|---|---|---|---|---|
| 21 | Correlation Coefficient | 1.000 | .312(**) | .216(*) | .291(**) | .315(**) |
| | Sig. (2-tailed) | . | .001 | .026 | .003 | .001 |
| | N | 107 | 106 | 106 | 105 | 106 |
| 34 | Correlation Coefficient | .312(**) | 1.000 | .462(**) | .493(**) | .533(**) |
| | Sig. (2-tailed) | .001 | . | .000 | .000 | .000 |
| | N | 106 | 108 | 107 | 106 | 107 |
| 37 | Correlation Coefficient | .216(*) | .462(**) | 1.000 | .484(**) | .641(**) |
| | Sig. (2-tailed) | .026 | .000 | . | .000 | .000 |
| | N | 106 | 107 | 109 | 107 | 109 |
| 42 | Correlation Coefficient | .291(**) | .493(**) | .484(**) | 1.000 | .501(**) |
| | Sig. (2-tailed) | .003 | .000 | .000 | . | .000 |
| | N | 105 | 106 | 107 | 107 | 107 |
| 48 | Correlation Coefficient | .315(**) | .533(**) | .641(**) | .501(**) | 1.000 |
| | Sig. (2-tailed) | .001 | .000 | .000 | .000 | . |
| | N | 106 | 107 | 109 | 107 | 110 |

Group = 1

**Correlations**

| | | 21 | 34 | 37 | 42 | 48 |
|---|---|---|---|---|---|---|
| 21 | Correlation Coefficient | 1.000 | .311(**) | .326(**) | .302(**) | .291(**) |
| | Sig. (2-tailed) | . | .000 | .000 | .000 | .000 |
| | N | 340 | 336 | 331 | 335 | 335 |
| 34 | Correlation Coefficient | .311(**) | 1.000 | .277(**) | .282(**) | .433(**) |
| | Sig. (2-tailed) | .000 | . | .000 | .000 | .000 |
| | N | 336 | 343 | 336 | 340 | 339 |
| 37 | Correlation Coefficient | .326(**) | .277(**) | 1.000 | .322(**) | .552(**) |
| | Sig. (2-tailed) | .000 | .000 | . | .000 | .000 |
| | N | 331 | 336 | 338 | 335 | 334 |
| 42 | Correlation Coefficient | .302(**) | .282(**) | .322(**) | 1.000 | .364(**) |
| | Sig. (2-tailed) | .000 | .000 | .000 | . | .000 |
| | N | 335 | 340 | 335 | 341 | 338 |
| 48 | Correlation Coefficient | .291(**) | .433(**) | .552(**) | .364(**) | 1.000 |
| | Sig. (2-tailed) | .000 | .000 | .000 | .000 | . |
| | N | 335 | 339 | 334 | 338 | 341 |

Group = 2

**Mock exams**

II.3. Mock exams can predict the results of the real exam.
II.4. Mock exams are an appropriate means of evaluating our learning.
III.25. Taking mock exams

**Correlations**

| | | 3 | 4 | 25 |
|---|---|---|---|---|
| 3 | Correlation Coefficient | 1.000 | .326(**) | -.086 |
| | Sig. (2-tailed) | . | .001 | .395 |
| | N | 109 | 106 | 101 |
| 4 | Correlation Coefficient | .326(**) | 1.000 | .099 |
| | Sig. (2-tailed) | .001 | . | .330 |
| | N | 106 | 107 | 99 |
| 25 | Correlation Coefficient | -.086 | .099 | 1.000 |
| | Sig. (2-tailed) | .395 | .330 | . |
| | N | 101 | 99 | 101 |

Group = 1

**Correlations**

| | | 3 | 4 | 25 |
|---|---|---|---|---|
| 3 | Correlation Coefficient | 1.000 | .274(**) | .196(**) |
| | Sig. (2-tailed) | . | .000 | .000 |
| | N | 344 | 342 | 316 |
| 4 | Correlation Coefficient | .274(**) | 1.000 | .165(**) |
| | Sig. (2-tailed) | .000 | . | .003 |
| | N | 342 | 344 | 316 |
| 25 | Correlation Coefficient | .196(**) | .165(**) | 1.000 |
| | Sig. (2-tailed) | .000 | .003 | . |
| | N | 316 | 316 | 318 |

Group = 2

**Exam papers**

15. Practicing with exam papers of previous years
34. Answering sample exam papers under timed conditions similar to the real exam
46. Practicing with multiple-choice reading passages
83. Answering sample questions

**Correlations**

| | | 15 | 34 | 46 | 83 |
|---|---|---|---|---|---|
| 15 | Correlation Coefficient | 1.000 | .338(**) | .191 | .439(**) |
| | Sig. (2-tailed) | . | .001 | .064 | .000 |
| | N | 100 | 94 | 95 | 91 |
| 34 | Correlation Coefficient | .338(**) | 1.000 | .438(**) | .492(**) |
| | Sig. (2-tailed) | .001 | . | .000 | .000 |
| | N | 94 | 103 | 98 | 93 |
| 46 | Correlation Coefficient | .191 | .438(**) | 1.000 | .465(**) |
| | Sig. (2-tailed) | .064 | .000 | . | .000 |
| | N | 95 | 98 | 103 | 93 |
| 83 | Correlation Coefficient | .439(**) | .492(**) | .465(**) | 1.000 |
| | Sig. (2-tailed) | .000 | .000 | .000 | . |
| | N | 91 | 93 | 93 | 100 |

Group = 1

**Correlations**

| | | 15 | 34 | 46 | 83 |
|---|---|---|---|---|---|
| | Correlation Coefficient | 1.000 | .409(**) | .409(**) | .443(**) |
| 15 | Sig. (2-tailed) | . | .000 | .000 | .000 |
| | N | 310 | 284 | 278 | 276 |
| | Correlation Coefficient | .409(**) | 1.000 | .419(**) | .506(**) |
| 34 | Sig. (2-tailed) | .000 | . | .000 | .000 |
| | N | 284 | 308 | 282 | 278 |
| | Correlation Coefficient | .409(**) | .419(**) | 1.000 | .572(**) |
| 46 | Sig. (2-tailed) | .000 | .000 | . | .000 |
| | N | 278 | 282 | 305 | 280 |
| | Correlation Coefficient | .443(**) | .506(**) | .572(**) | 1.000 |
| 83 | Sig. (2-tailed) | .000 | .000 | .000 | . |
| | N | 276 | 278 | 280 | 304 |
| | | | | | |

Group = 2

**Exam skills**

15. Practicing with exam papers of previous years
25. Taking mock exams
34. Answering sample exam papers under timed conditions similar to the real exam
46. Practicing with multiple-choice reading passages
49. Using comprehension questions to get a general idea of the passage
83. Answering sample questions

**Correlations**

| | | 15 | 25 | 34 | 46 | 49 | 83 |
|---|---|---|---|---|---|---|---|
| 15 | Correlation Coefficient | 1.000 | .406(**) | .338(**) | .191 | .218(*) | .439(**) |
| | Sig. (2-tailed) | . | .000 | .001 | .064 | .034 | .000 |
| | N | 100 | 94 | 94 | 95 | 94 | 91 |
| 25 | Correlation Coefficient | .406(**) | 1.000 | .409(**) | .350(**) | .228(*) | .521(**) |
| | Sig. (2-tailed) | .000 | . | .000 | .000 | .025 | .000 |
| | N | 94 | 101 | 97 | 96 | 97 | 91 |
| 34 | Correlation Coefficient | .338(**) | .409(**) | 1.000 | .438(**) | .346(**) | .492(**) |
| | Sig. (2-tailed) | .001 | .000 | . | .000 | .000 | .000 |
| | N | 94 | 97 | 103 | 98 | 99 | 93 |
| 46 | Correlation Coefficient | .191 | .350(**) | .438(**) | 1.000 | .527(**) | .465(**) |
| | Sig. (2-tailed) | .064 | .000 | .000 | . | .000 | .000 |
| | N | 95 | 96 | 98 | 103 | 100 | 93 |
| 49 | Correlation Coefficient | .218(*) | .228(*) | .346(**) | .527(**) | 1.000 | .349(**) |
| | Sig. (2-tailed) | .034 | .025 | .000 | .000 | . | .001 |
| | N | 94 | 97 | 99 | 100 | 103 | 93 |
| 83 | Correlation Coefficient | .439(**) | .521(**) | .492(**) | .465(**) | .349(**) | 1.000 |
| | Sig. (2-tailed) | .000 | .000 | .000 | .000 | .001 | . |
| | N | 91 | 91 | 93 | 93 | 93 | 100 |

Group = 1

**Correlations**

| | | 15 | 25 | 34 | 46 | 49 | 83 |
|---|---|---|---|---|---|---|---|
| 15 | Correlation Coefficient | 1.000 | .377(**) | .409(**) | .409(**) | .315(**) | .443(**) |
| | Sig. (2-tailed) | . | .000 | .000 | .000 | .000 | .000 |
| | N | 310 | 293 | 284 | 278 | 280 | 276 |
| 25 | Correlation Coefficient | .377(**) | 1.000 | .467(**) | .411(**) | .203(**) | .373(**) |
| | Sig. (2-tailed) | .000 | . | .000 | .000 | .001 | .000 |
| | N | 293 | 318 | 294 | 289 | 281 | 284 |
| 34 | Correlation Coefficient | .409(**) | .467(**) | 1.000 | .419(**) | .342(**) | .506(**) |
| | Sig. (2-tailed) | .000 | .000 | . | .000 | .000 | .000 |
| | N | 284 | 294 | 308 | 282 | 277 | 278 |
| 46 | Correlation Coefficient | .409(**) | .411(**) | .419(**) | 1.000 | .421(**) | .572(**) |
| | Sig. (2-tailed) | .000 | .000 | .000 | . | .000 | .000 |
| | N | 278 | 289 | 282 | 305 | 279 | 280 |
| 49 | Correlation Coefficient | .315(**) | .203(**) | .342(**) | .421(**) | 1.000 | .338(**) |
| | Sig. (2-tailed) | .000 | .001 | .000 | .000 | . | .000 |
| | N | 280 | 281 | 277 | 279 | 301 | 277 |
| 83 | Correlation Coefficient | .443(**) | .373(**) | .506(**) | .572(**) | .338(**) | 1.000 |
| | Sig. (2-tailed) | .000 | .000 | .000 | .000 | .000 | . |
| | N | 276 | 284 | 278 | 280 | 277 | 304 |

Group = 2

**Evaluation**

3. Mock exams can predict the results of the real exam.
4. Mock exams are an appropriate means of evaluating our learning.
47. I always evaluate my learning.
15. Practicing with exam papers of previous years
25. Taking mock exams
34. Answering sample exam papers under timed conditions similar to the real exam
46. Practicing with multiple-choice reading passages
83. Answering sample questions

**Correlations**

| | | 3 | 4 | 47 | 15 | 25 | 34 | 46 | 83 |
|---|---|---|---|---|---|---|---|---|---|
| 3 | Correlation Coefficient | 1.000 | .326(**) | .235(*) | .175 | -.086 | .048 | .136 | .065 |
| | Sig. (2-tailed) | . | .001 | .017 | .083 | .395 | .635 | .172 | .521 |
| | N | 109 | 106 | 102 | 99 | 101 | 102 | 103 | 99 |
| 4 | Correlation Coefficient | .326(**) | 1.000 | .289(**) | .077 | .099 | .056 | .087 | .256(*) |
| | Sig. (2-tailed) | .001 | . | .003 | .454 | .330 | .583 | .390 | .011 |
| | N | 106 | 107 | 101 | 97 | 99 | 99 | 100 | 97 |
| 47 | Correlation Coefficient | .235(*) | .289(**) | 1.000 | .254(*) | .341(**) | .456(**) | .455(**) | .356(**) |
| | Sig. (2-tailed) | .017 | .003 | . | .014 | .001 | .000 | .000 | .000 |
| | N | 102 | 101 | 103 | 94 | 94 | 96 | 97 | 93 |
| 15 | Correlation Coefficient | .175 | .077 | .254(*) | 1.000 | .406(**) | .338(**) | .191 | .439(**) |
| | Sig. (2-tailed) | .083 | .454 | .014 | . | .000 | .001 | .064 | .000 |
| | N | 99 | 97 | 94 | 100 | 94 | 94 | 95 | 91 |
| 25 | Correlation Coefficient | -.086 | .099 | .341(**) | .406(**) | 1.000 | .409(**) | .350(**) | .521(**) |
| | Sig. (2-tailed) | .395 | .330 | .001 | .000 | . | .000 | .000 | .000 |
| | N | 101 | 99 | 94 | 94 | 101 | 97 | 96 | 91 |
| 34 | Correlation Coefficient | .048 | .056 | .456(**) | .338(**) | .409(**) | 1.000 | .438(**) | .492(**) |
| | Sig. (2-tailed) | .635 | .583 | .000 | .001 | .000 | . | .000 | .000 |
| | N | 102 | 99 | 96 | 94 | 97 | 103 | 98 | 93 |
| 46 | Correlation Coefficient | .136 | .087 | .455(**) | .191 | .350(**) | .438(**) | 1.000 | .465(**) |
| | Sig. (2-tailed) | .172 | .390 | .000 | .064 | .000 | .000 | . | .000 |
| | N | 103 | 100 | 97 | 95 | 96 | 98 | 103 | 93 |
| 83 | Correlation Coefficient | .065 | .256(*) | .356(**) | .439(**) | .521(**) | .492(**) | .465(**) | 1.000 |
| | Sig. (2-tailed) | .521 | .011 | .000 | .000 | .000 | .000 | .000 | . |
| | N | 99 | 97 | 93 | 91 | 91 | 93 | 93 | 100 |

Group = 1

**Correlations**

| | | 3 | 4 | 47 | 15 | 25 | 34 | 46 | 83 |
|---|---|---|---|---|---|---|---|---|---|
| 3 | Correlation Coefficient | 1.000 | .274(**) | .171(**) | .089 | .196(**) | .227(**) | .144(*) | .063 |
| | Sig. (2-tailed) | . | .000 | .002 | .119 | .000 | .000 | .012 | .277 |
| | N | 344 | 342 | 334 | 308 | 316 | 305 | 302 | 303 |
| 4 | Correlation Coefficient | .274(**) | 1.000 | .205(**) | .127(*) | .165(**) | .171(**) | .253(**) | .181(**) |
| | Sig. (2-tailed) | .000 | . | .000 | .025 | .003 | .003 | .000 | .002 |
| | N | 342 | 344 | 334 | 309 | 316 | 307 | 302 | 303 |
| 47 | Correlation Coefficient | .171(**) | .205(**) | 1.000 | .295(**) | .284(**) | .404(**) | .368(**) | .360(**) |
| | Sig. (2-tailed) | .002 | .000 | . | .000 | .000 | .000 | .000 | .000 |
| | N | 334 | 334 | 338 | 301 | 310 | 300 | 299 | 296 |
| 15 | Correlation Coefficient | .089 | .127(*) | .295(**) | 1.000 | .377(**) | .409(**) | .409(**) | .443(**) |
| | Sig. (2-tailed) | .119 | .025 | .000 | . | .000 | .000 | .000 | .000 |
| | N | 308 | 309 | 301 | 310 | 293 | 284 | 278 | 276 |
| 25 | Correlation Coefficient | .196(**) | .165(**) | .284(**) | .377(**) | 1.000 | .467(**) | .411(**) | .373(**) |
| | Sig. (2-tailed) | .000 | .003 | .000 | .000 | . | .000 | .000 | .000 |
| | N | 316 | 316 | 310 | 293 | 318 | 294 | 289 | 284 |
| 34 | Correlation Coefficient | .227(**) | .171(**) | .404(**) | .409(**) | .467(**) | 1.000 | .419(**) | .506(**) |
| | Sig. (2-tailed) | .000 | .003 | .000 | .000 | .000 | . | .000 | .000 |
| | N | 305 | 307 | 300 | 284 | 294 | 308 | 282 | 278 |
| 46 | Correlation Coefficient | .144(*) | .253(**) | .368(**) | .409(**) | .411(**) | .419(**) | 1.000 | .572(**) |
| | Sig. (2-tailed) | .012 | .000 | .000 | .000 | .000 | .000 | . | .000 |
| | N | 302 | 302 | 299 | 278 | 289 | 282 | 305 | 280 |
| 83 | Correlation Coefficient | .063 | .181(**) | .360(**) | .443(**) | .373(**) | .506(**) | .572(**) | 1.000 |
| | Sig. (2-tailed) | .277 | .002 | .000 | .000 | .000 | .000 | .000 | . |
| | N | 303 | 303 | 296 | 276 | 284 | 278 | 280 | 304 |

Group = 2

**Dependence on self**

13. I try to find appropriate ways of test preparation myself.
51. I always monitor the way I'm preparing.
35. Thinking about appropriate ways of learning

**Correlations**

| | | 13 | 51 | 35 |
|---|---|---|---|---|
| 13 | Correlation Coefficient | 1.000 | .391(**) | .150 |
| | Sig. (2-tailed) | . | .000 | .135 |
| | N | 108 | 107 | 101 |
| 51 | Correlation Coefficient | .391(**) | 1.000 | .308(**) |
| | Sig. (2-tailed) | .000 | . | .002 |
| | N | 107 | 109 | 102 |
| 35 | Correlation Coefficient | .150 | .308(**) | 1.000 |
| | Sig. (2-tailed) | .135 | .002 | . |
| | N | 101 | 102 | 103 |

Group = 1

**Correlations**

| | | 13 | 51 | 35 |
|---|---|---|---|---|
| 13 | Correlation Coefficient | 1.000 | .179(**) | .213(**) |
| | Sig. (2-tailed) | . | .001 | .000 |
| | N | 342 | 339 | 301 |
| 51 | Correlation Coefficient | .179(**) | 1.000 | .269(**) |
| | Sig. (2-tailed) | .001 | . | .000 |
| | N | 339 | 344 | 302 |
| 35 | Correlation Coefficient | .213(**) | .269(**) | 1.000 |
| | Sig. (2-tailed) | .000 | .000 | . |
| | N | 301 | 302 | 305 |

Group = 2

**Dependence on teacher**

27. The teaching methods of the teacher is an effective factor in the preparation.
41. I try to seek teacher's advice on appropriate ways of test preparation.
26. Getting feedback from teacher

**Correlations**

| | | 27 | 41 | 26 |
|---|---|---|---|---|
| 27 | Correlation Coefficient | 1.000 | .244(*) | .135 |
| | Sig. (2-tailed) | . | .011 | .168 |
| | N | 110 | 108 | 106 |
| 41 | Correlation Coefficient | .244(*) | 1.000 | .334(**) |
| | Sig. (2-tailed) | .011 | . | .001 |
| | N | 108 | 108 | 104 |
| 26 | Correlation Coefficient | .135 | .334(**) | 1.000 |
| | Sig. (2-tailed) | .168 | .001 | . |
| | N | 106 | 104 | 106 |

Group = 1

**Correlations**

| | | 27 | 41 | 26 |
|---|---|---|---|---|
| 27 | Correlation Coefficient | 1.000 | .346(**) | .207(**) |
| | Sig. (2-tailed) | . | .000 | .000 |
| | N | 344 | 342 | 322 |
| 41 | Correlation Coefficient | .346(**) | 1.000 | .326(**) |
| | Sig. (2-tailed) | .000 | . | .000 |
| | N | 342 | 344 | 322 |
| 26 | Correlation Coefficient | .207(**) | .326(**) | 1.000 |
| | Sig. (2-tailed) | .000 | .000 | . |
| | N | 322 | 322 | 326 |

Group = 2

**Dependence on others**

11. Seeking other people's advice on appropriate ways of learning
27. Getting feedback from others

**Correlations**

| | | 11 | 27 |
|---|---|---|---|
| 11 | Correlation Coefficient | 1.000 | .388(**) |
| | Sig. (2-tailed) | . | .000 |
| | N | 100 | 98 |
| 27 | Correlation Coefficient | .388(**) | 1.000 |
| | Sig. (2-tailed) | .000 | . |
| | N | 98 | 106 |

Group = 1

**Correlations**

| | | 11 | 27 |
|---|---|---|---|
| 11 | Correlation Coefficient | 1.000 | .465(**) |
| | Sig. (2-tailed) | . | .000 |
| | N | 305 | 291 |
| 27 | Correlation Coefficient | .465(**) | 1.000 |
| | Sig. (2-tailed) | .000 | . |
| | N | 291 | 319 |

Group = 2

**Motivations**

20. People in my country feel that it is important to speak English.
24. I would like to learn English so that I can get to know English speakers better.
29. If I learn English very well, I will have better opportunities for a good job.
31. I want to learn to speak English well.
32. I would like to have English speaking friends.
6. Success in the SPE Exam means success in English courses at university.
14. I'm aware of the purpose of each section of the SPE Exam.
16. Success in the SPE Exam is important for me
28. The SPE Exam makes me try.
30. I enjoy learning for the SPE Exam.
32. I try to prepare for the SPE Exam as far as I can.
36. I think studying for the SPE Test has improved my English.
40. I'm sure that I will succeed on the SPE.
43. When I fail an exam, I work harder.
61. What do you think is your level of motivation? Please tick the appropriate box.
13. I enjoy practicing English with the English speakers I meet.

**Correlations**

| | | 20 | 24 | 29 | 31 | 32 | 6 | 14 | 16 | 28 | 30 | 32 | 36 | 40 | 43 | 61 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 20 | rho | 1.000 | .096 | .206(*) | .162 | .168 | .026 | .068 | .023 | .038 | -.038 | -.017 | .065 | -.049 | .015 | -.055 | .081 |
| | Sig. (2t) | . | .321 | .031 | .095 | .083 | .789 | .489 | .816 | .698 | .696 | .862 | .507 | .614 | .880 | .578 | .405 |
| | N | 109 | 108 | 109 | 107 | 107 | 109 | 106 | 108 | 109 | 109 | 107 | 108 | 107 | 107 | 104 | 109 |
| 24 | rho | .096 | 1.000 | -.041 | -.034 | .162 | .003 | .033 | -.046 | .166 | .198(*) | -.009 | .113 | -.030 | .164 | -.102 | .114 |
| | Sig. (2t) | .321 | . | .670 | .729 | .097 | .977 | .737 | .639 | .085 | .040 | .923 | .246 | .762 | .092 | .304 | .239 |
| | N | 108 | 108 | 108 | 106 | 106 | 108 | 105 | 107 | 108 | 108 | 106 | 107 | 106 | 106 | 103 | 108 |
| 29 | rho | .206(*) | -.041 | 1.000 | .333(**) | .334(**) | .066 | .079 | .213(*) | .187 | .320(**) | .240(*) | .163 | .073 | .047 | .312(**) | .364(**) |
| | Sig. (2t) | .031 | .670 | . | .000 | .000 | .496 | .419 | .027 | .051 | .001 | .013 | .091 | .454 | .628 | .001 | .000 |
| | N | 109 | 108 | 109 | 107 | 107 | 109 | 106 | 108 | 109 | 109 | 107 | 108 | 107 | 107 | 104 | 109 |
| 31 | rho | .162 | -.034 | .333(**) | 1.000 | .436(**) | -.081 | -.040 | .440(**) | .171 | .255(**) | .259(**) | .143 | .171 | .138 | .215(*) | .266(**) |
| | Sig. (2t) | .095 | .729 | .000 | . | .000 | .404 | .688 | .000 | .077 | .008 | .008 | .143 | .081 | .160 | .030 | .006 |
| | N | 107 | 106 | 107 | 107 | 107 | 107 | 104 | 106 | 107 | 107 | 105 | 106 | 105 | 105 | 102 | 107 |
| 32 | rho | .168 | .162 | .334(**) | .436(**) | 1.000 | -.029 | .101 | .364(**) | .072 | .262(**) | .253(**) | .145 | .280(**) | -.041 | .294(**) | .317(**) |
| | Sig. (2t) | .083 | .097 | .000 | .000 | . | .767 | .307 | .000 | .463 | .006 | .009 | .139 | .004 | .679 | .003 | .001 |
| | N | 107 | 106 | 107 | 107 | 107 | 107 | 104 | 106 | 107 | 107 | 105 | 106 | 105 | 105 | 102 | 107 |
| 6 | rho | .026 | .003 | .066 | -.081 | -.029 | 1.000 | .062 | .077 | .289(**) | .040 | .103 | .203(*) | -.024 | .121 | -.092 | -.102 |
| | Sig. (2t) | .789 | .977 | .496 | .404 | .767 | . | .530 | .424 | .002 | .680 | .290 | .034 | .805 | .211 | .353 | .290 |
| | N | 109 | 108 | 109 | 107 | 107 | 110 | 106 | 109 | 110 | 110 | 108 | 109 | 108 | 108 | 105 | 110 |
| 14 | rho | .068 | .033 | .079 | -.040 | .101 | .062 | 1.000 | .037 | .256(**) | .207(*) | .172 | .169 | .244(*) | .188 | .032 | .133 |
| | Sig. (2t) | .489 | .737 | .419 | .688 | .307 | .530 | . | .711 | .008 | .034 | .080 | .084 | .013 | .056 | .752 | .174 |
| | N | 106 | 105 | 106 | 104 | 104 | 106 | 106 | 105 | 106 | 106 | 104 | 105 | 104 | 104 | 102 | 106 |
| 16 | rho | .023 | -.046 | .213(*) | .440(**) | .364(**) | .077 | .037 | 1.000 | .165 | .272(**) | .237(*) | .225(*) | .209(*) | .162 | .296(**) | .206(*) |
| | Sig. (2t) | .816 | .639 | .027 | .000 | .000 | .424 | .711 | . | .086 | .004 | .014 | .019 | .030 | .095 | .002 | .032 |
| | N | 108 | 107 | 108 | 106 | 106 | 109 | 105 | 109 | 109 | 109 | 107 | 108 | 107 | 107 | 104 | 109 |
| 28 | rho | .038 | .166 | .187 | .171 | .072 | .289(**) | .256(**) | .165 | 1.000 | .348(**) | .305(**) | .272(**) | .236(*) | .389(**) | .160 | .028 |
| | Sig. (2t) | .698 | .085 | .051 | .077 | .463 | .002 | .008 | .086 | . | .000 | .001 | .004 | .014 | .000 | .103 | .771 |
| | N | 109 | 108 | 109 | 107 | 107 | 110 | 106 | 109 | 110 | 110 | 108 | 109 | 108 | 108 | 105 | 110 |
| 30 | rho | -.038 | .198(*) | .320(**) | .255(**) | .262(**) | .040 | .207(*) | .272(**) | .348(**) | 1.000 | .525(**) | .414(**) | .307(**) | .285(**) | .221(*) | .353(**) |
| | Sig. (2t) | .696 | .040 | .001 | .008 | .006 | .680 | .034 | .004 | .000 | . | .000 | .000 | .001 | .003 | .024 | .000 |
| | N | 109 | 108 | 109 | 107 | 107 | 110 | 106 | 109 | 110 | 110 | 108 | 109 | 108 | 108 | 105 | 110 |
| 32 | rho | -.017 | -.009 | .240(*) | .259(**) | .253(**) | .103 | .172 | .237(*) | .305(**) | .525(**) | 1.000 | .406(**) | .293(**) | .291(**) | .175 | .263(**) |
| | Sig. (2t) | .862 | .923 | .013 | .008 | .009 | .290 | .080 | .014 | .001 | .000 | . | .000 | .002 | .002 | .077 | .006 |
| | N | 107 | 106 | 107 | 105 | 105 | 108 | 104 | 107 | 108 | 108 | 108 | 107 | 106 | 106 | 103 | 108 |
| 36 | rho | .065 | .113 | .163 | .143 | .145 | .203(*) | .169 | .225(*) | .272(**) | .414(**) | .406(**) | 1.000 | .092 | .150 | .177 | .147 |
| | Sig. (2t) | .507 | .246 | .091 | .143 | .139 | .034 | .084 | .019 | .004 | .000 | .000 | . | .347 | .123 | .072 | .128 |
| | N | 108 | 107 | 108 | 106 | 106 | 109 | 105 | 108 | 109 | 109 | 107 | 109 | 107 | 107 | 104 | 109 |
| 40 | rho | -.049 | -.030 | .073 | .171 | .280(**) | -.024 | .244(*) | .209(*) | .236(*) | .307(**) | .293(**) | .092 | 1.000 | .164 | .351(**) | .110 |
| | Sig. (2t) | .614 | .762 | .454 | .081 | .004 | .805 | .013 | .030 | .014 | .001 | .002 | .347 | . | .092 | .000 | .258 |
| | N | 107 | 106 | 107 | 105 | 105 | 108 | 104 | 107 | 108 | 108 | 106 | 107 | 108 | 107 | 104 | 108 |
| 43 | rho | .015 | .164 | .047 | .138 | -.041 | .121 | .188 | .162 | .389(**) | .285(**) | .291(**) | .150 | .164 | 1.000 | .018 | .193(*) |
| | Sig. (2t) | .880 | .092 | .628 | .160 | .679 | .211 | .056 | .095 | .000 | .003 | .002 | .123 | .092 | . | .852 | .046 |
| | N | 107 | 106 | 107 | 105 | 105 | 108 | 104 | 107 | 108 | 108 | 106 | 107 | 107 | 108 | 104 | 108 |
| 61 | rho | -.055 | -.102 | .312(**) | .215(*) | .294(**) | -.092 | .032 | .296(**) | .160 | .221(*) | .175 | .177 | .351(**) | .018 | 1.000 | .119 |
| | Sig. (2t) | .578 | .304 | .001 | .030 | .003 | .353 | .752 | .002 | .103 | .024 | .077 | .072 | .000 | .852 | . | .229 |
| | N | 104 | 103 | 104 | 102 | 102 | 105 | 102 | 104 | 105 | 105 | 103 | 104 | 104 | 104 | 106 | 105 |
| 13 | rho | .081 | .114 | .364(**) | .266(**) | .317(**) | -.102 | .133 | .206(*) | .028 | .353(**) | .263(**) | .147 | .110 | .193(*) | .119 | 1.000 |
| | Sig. (2t) | .405 | .239 | .000 | .006 | .001 | .290 | .174 | .032 | .771 | .000 | .006 | .128 | .258 | .046 | .229 | . |
| | N | 109 | 108 | 109 | 107 | 107 | 110 | 106 | 109 | 110 | 110 | 108 | 109 | 108 | 108 | 105 | 110 |

Group = 1

**Motivations (Continued)**

**Correlations**

| | | 20 | 24 | 29 | 31 | 32 | 6 | 14 | 16 | 28 | 30 | 32 | 36 | 40 | 43 | 61 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 20 | rho | 1.000 | -.013 | .281(**) | .267(**) | .163(**) | .118(*) | -.002 | .230(**) | .158(**) | .178(**) | .124(*) | .108 | .069 | .153(**) | .151(**) | .170(**) |
| | Sig (2t) | . | .815 | .000 | .000 | .003 | .030 | .974 | .000 | .004 | .001 | .024 | .050 | .208 | .005 | .006 | .002 |
| | N | 340 | 336 | 338 | 337 | 337 | 336 | 330 | 334 | 336 | 334 | 333 | 327 | 337 | 333 | 322 | 337 |
| 24 | Rho | -.013 | 1.000 | .156(**) | .142(**) | .271(**) | .201(**) | .137(*) | .040 | .057 | .081 | .121(*) | .085 | .083 | .067 | .026 | .176(**) |
| | Sig. (2t) | .815 | . | .004 | .009 | .000 | .000 | .013 | .467 | .296 | .141 | .027 | .126 | .127 | .221 | .636 | .001 |
| | N | 336 | 340 | 338 | 337 | 336 | 336 | 331 | 335 | 336 | 333 | 332 | 327 | 337 | 332 | 322 | 337 |
| 29 | Rho | .281(**) | .156(**) | 1.000 | .310(**) | .247(**) | .145(**) | -.021 | .158(**) | .152(**) | .091 | .162(**) | .153(**) | .018 | .082 | .102 | .201(**) |
| | Sig. (2t) | .000 | .004 | . | .000 | .000 | .007 | .701 | .004 | .005 | .095 | .003 | .005 | .738 | .137 | .067 | .000 |
| | N | 338 | 338 | 343 | 339 | 338 | 338 | 332 | 337 | 339 | 337 | 335 | 330 | 339 | 334 | 325 | 339 |
| 31 | Rho | .267(**) | .142(**) | .310(**) | 1.000 | .430(**) | .031 | -.004 | .169(**) | .164(**) | .175(**) | .156(**) | .127(*) | .127(*) | .120(*) | .284(**) | .354(**) |
| | Sig. (2t) | .000 | .009 | .000 | . | .000 | .567 | .949 | .002 | .002 | .001 | .004 | .021 | .019 | .029 | .000 | .000 |
| | N | 337 | 337 | 339 | 344 | 342 | 340 | 334 | 337 | 340 | 337 | 336 | 331 | 340 | 335 | 325 | 341 |
| 32 | Rho | .163(**) | .271(**) | .247(**) | .430(**) | 1.000 | .029 | .089 | .114(*) | .103 | .125(*) | .084 | .140(*) | .078 | .014 | .292(**) | .400(**) |
| | Sig. (2t) | .003 | .000 | .000 | .000 | . | .596 | .106 | .037 | .058 | .021 | .124 | .011 | .149 | .797 | .000 | .000 |
| | N | 337 | 336 | 338 | 342 | 343 | 338 | 332 | 336 | 339 | 337 | 335 | 330 | 339 | 334 | 324 | 340 |
| 6 | Rho | .118(*) | .201(**) | .145(**) | .031 | .029 | 1.000 | .023 | .070 | .215(**) | .071 | .106 | .205(**) | .020 | .105 | .012 | .100 |
| | Sig. (2t) | .030 | .000 | .007 | .567 | .596 | . | .672 | .197 | .000 | .195 | .052 | .000 | .707 | .056 | .830 | .066 |
| | N | 336 | 336 | 338 | 340 | 338 | 343 | 335 | 337 | 340 | 337 | 337 | 332 | 339 | 335 | 325 | 339 |
| 14 | Rho | -.002 | .137(*) | -.021 | -.004 | .089 | .023 | 1.000 | .151(**) | .132(*) | .195(**) | .224(**) | .029 | .236(**) | .213(**) | .118(*) | .108(*) |
| | Sig. (2t) | .974 | .013 | .701 | .949 | .106 | .672 | . | .006 | .016 | .000 | .000 | .604 | .000 | .000 | .035 | .050 |
| | N | 330 | 331 | 332 | 334 | 332 | 335 | 336 | 331 | 333 | 330 | 329 | 325 | 334 | 329 | 319 | 333 |
| 16 | Rho | .230(**) | .040 | .158(**) | .169(**) | .114(*) | .070 | .151(**) | 1.000 | .340(**) | .391(**) | .194(**) | .307(**) | .367(**) | .150(**) | .350(**) | .278(**) |
| | Sig. (2t) | .000 | .467 | .004 | .002 | .037 | .197 | .006 | . | .000 | .000 | .000 | .000 | .000 | .006 | .000 | .000 |
| | N | 334 | 335 | 337 | 337 | 336 | 337 | 331 | 341 | 337 | 334 | 333 | 328 | 338 | 332 | 323 | 337 |
| 28 | Rho | .158(**) | .057 | .152(**) | .164(**) | .103 | .215(**) | .132(*) | .340(**) | 1.000 | .305(**) | .242(**) | .378(**) | .213(**) | .222(**) | .175(**) | .247(**) |
| | Sig. (2t) | .004 | .296 | .005 | .002 | .058 | .000 | .016 | .000 | . | .000 | .000 | .000 | .000 | .000 | .001 | .000 |
| | N | 336 | 336 | 339 | 340 | 339 | 340 | 333 | 337 | 344 | 339 | 338 | 333 | 341 | 336 | 328 | 340 |
| 30 | Rho | .178(**) | .081 | .091 | .175(**) | .125(*) | .071 | .195(**) | .391(**) | .305(**) | 1.000 | .183(**) | .331(**) | .382(**) | .274(**) | .412(**) | .230(**) |
| | Sig. (2t) | .001 | .141 | .095 | .001 | .021 | .195 | .000 | .000 | .000 | . | .001 | .000 | .000 | .000 | .000 | .000 |
| | N | 334 | 333 | 337 | 337 | 337 | 337 | 330 | 334 | 339 | 341 | 335 | 330 | 338 | 333 | 325 | 337 |
| 32 | Rho | .124(*) | .121(*) | .162(**) | .156(**) | .084 | .106 | .224(**) | .194(**) | .242(**) | .183(**) | 1.000 | .180(**) | .064 | .347(**) | .086 | .214(**) |
| | Sig. (2t) | .024 | .027 | .003 | .004 | .124 | .052 | .000 | .000 | .000 | .001 | . | .001 | .239 | .000 | .122 | .000 |
| | N | 333 | 332 | 335 | 336 | 335 | 337 | 329 | 333 | 338 | 335 | 340 | 331 | 337 | 334 | 324 | 336 |
| 36 | Rho | .108 | .085 | .153(**) | .127(*) | .140(*) | .205(**) | .029 | .307(**) | .378(**) | .331(**) | .180(**) | 1.000 | .231(**) | .237(**) | .260(**) | .255(**) |
| | Sig. (2t) | .050 | .126 | .005 | .021 | .011 | .000 | .604 | .000 | .000 | .000 | .001 | . | .000 | .000 | .000 | .000 |
| | N | 327 | 327 | 330 | 331 | 330 | 332 | 325 | 328 | 333 | 330 | 331 | 335 | 332 | 328 | 319 | 331 |
| 40 | Rho | .069 | .083 | .018 | .127(*) | .078 | .020 | .236(**) | .367(**) | .213(**) | .382(**) | .064 | .231(**) | 1.000 | .228(**) | .425(**) | .170(**) |
| | Sig. (2t) | .208 | .127 | .738 | .019 | .149 | .707 | .000 | .000 | .000 | .000 | .239 | .000 | . | .000 | .000 | .002 |
| | N | 337 | 337 | 339 | 340 | 339 | 339 | 334 | 338 | 341 | 338 | 337 | 332 | 343 | 337 | 327 | 340 |
| 43 | Rho | .153(**) | .067 | .082 | .120(*) | .014 | .105 | .213(**) | .150(**) | .222(**) | .274(**) | .347(**) | .237(**) | .228(**) | 1.000 | .149(**) | .157(**) |
| | Sig. (2t) | .005 | .221 | .137 | .029 | .797 | .056 | .000 | .006 | .000 | .000 | .000 | .000 | .000 | . | .007 | .004 |
| | N | 333 | 332 | 334 | 335 | 334 | 335 | 329 | 332 | 336 | 333 | 334 | 328 | 337 | 338 | 323 | 335 |
| 61 | Rho | .151(**) | .026 | .102 | .284(**) | .292(**) | .012 | .118(*) | .350(**) | .175(**) | .412(**) | .086 | .260(**) | .425(**) | .149(**) | 1.000 | .242(**) |
| | Sig. (2t) | .006 | .636 | .067 | .000 | .000 | .830 | .035 | .000 | .001 | .000 | .122 | .000 | .000 | .007 | . | .000 |
| | N | 322 | 322 | 325 | 325 | 324 | 325 | 319 | 323 | 328 | 325 | 324 | 319 | 327 | 323 | 329 | 325 |
| 13 | Rho | .170(**) | .176(**) | .201(**) | .354(**) | .400(**) | .100 | .108(*) | .278(**) | .247(**) | .230(**) | .214(**) | .255(**) | .170(**) | .157(**) | .242(**) | 1.000 |
| | Sig. (2t) | .002 | .001 | .000 | .000 | .000 | .066 | .050 | .000 | .000 | .000 | .000 | .000 | .002 | .004 | .000 | . |
| | N | 337 | 337 | 339 | 341 | 340 | 339 | 333 | 337 | 340 | 337 | 336 | 331 | 340 | 335 | 325 | 344 |

Group = 2

**De-motivation**

33. I don't like to study for the SPE Exam.
46. I find many excuses for not studying for the SPE Exam.

**Correlations**

| | | 33 | 46 |
|---|---|---|---|
| 33 | Correlation Coefficient | 1.000 | .589(**) |
| | Sig. (2-tailed) | . | .000 |
| | N | 107 | 104 |
| 46 | Correlation Coefficient | .589(**) | 1.000 |
| | Sig. (2-tailed) | .000 | . |
| | N | 104 | 107 |

Group = 1

**Correlations**

| | | 33 | 46 |
|---|---|---|---|
| 33 | Correlation Coefficient | 1.000 | .432(**) |
| | Sig. (2-tailed) | . | .000 |
| | N | 341 | 333 |
| 46 | Correlation Coefficient | .432(**) | 1.000 |
| | Sig. (2-tailed) | .000 | . |
| | N | 333 | 338 |

Group = 2

**Difficulty of English and SPE/GE**

4. English is:

a) a very difficult language b) a difficult language c) a language of medium difficulty
e) an easy language f) a very easy language

5. I believe that I will learn to speak English very well.

15. If someone spent one hour a day learning a language, how long would it take them to speak the language very well:

a) less than a year b) 1-2 years c) 3-5 years
e) 5-10 years f) You can't learn a language in 1 hour a day

54. I think the SPE/GE Test is.....

(a) very difficult (b) difficult (c) of medium difficulty (d) easy (e) very easy

**Correlations**

| | | 4 | 5 | 15 | 54 |
|---|---|---|---|---|---|
| 4 | Correlation Coefficient | 1.000 | .268(**) | -.182 | -.431(**) |
| | Sig. (2-tailed) | . | .005 | .065 | .000 |
| | N | 110 | 109 | 104 | 109 |
| 5 | Correlation Coefficient | .268(**) | 1.000 | .028 | -.038 |
| | Sig. (2-tailed) | .005 | . | .775 | .698 |
| | N | 109 | 110 | 104 | 109 |
| 15 | Correlation Coefficient | -.182 | .028 | 1.000 | .045 |
| | Sig. (2-tailed) | .065 | .775 | . | .651 |
| | N | 104 | 104 | 105 | 104 |
| 54 | Correlation Coefficient | -.431(**) | -.038 | .045 | 1.000 |
| | Sig. (2-tailed) | .000 | .698 | .651 | . |
| | N | 109 | 109 | 104 | 110 |

Group = 1

**Correlations**

| | | 4 | 5 | 15 | 54 |
|---|---|---|---|---|---|
| 4 | Correlation Coefficient | 1.000 | .484(**) | -.111(*) | -.524(**) |
| | Sig. (2-tailed) | . | .000 | .042 | .000 |
| | N | 344 | 340 | 335 | 328 |
| 5 | Correlation Coefficient | .484(**) | 1.000 | -.042 | -.406(**) |
| | Sig. (2-tailed) | .000 | . | .439 | .000 |
| | N | 340 | 344 | 335 | 327 |
| 15 | Correlation Coefficient | -.111(*) | -.042 | 1.000 | .122(*) |
| | Sig. (2-tailed) | .042 | .439 | . | .028 |
| | N | 335 | 335 | 338 | 325 |
| 54 | Correlation Coefficient | -.524(**) | -.406(**) | .122(*) | 1.000 |
| | Sig. (2-tailed) | .000 | .000 | .028 | . |
| | N | 328 | 327 | 325 | 331 |

Group = 2

**Self-confidence**

5. I believe that I will learn to speak English very well.
16. I have a special ability for learning foreign languages.
19. I try to reduce the stress that I encounter during test preparation.
23. I try to improve my self-confidence.

**Correlations**

| | | 5 | 16 | 19 | 23 |
|---|---|---|---|---|---|
| 5 | Correlation Coefficient | 1.000 | .254(**) | .033 | .051 |
| | Sig. (2-tailed) | . | .008 | .731 | .599 |
| | N | 110 | 109 | 109 | 108 |
| 16 | Correlation Coefficient | .254(**) | 1.000 | .006 | .100 |
| | Sig. (2-tailed) | .008 | . | .947 | .301 |
| | N | 109 | 109 | 109 | 108 |
| 19 | Correlation Coefficient | .033 | .006 | 1.000 | -.120 |
| | Sig. (2-tailed) | .731 | .947 | . | .218 |
| | N | 109 | 109 | 109 | 108 |
| 23 | Correlation Coefficient | .051 | .100 | -.120 | 1.000 |
| | Sig. (2-tailed) | .599 | .301 | .218 | . |
| | N | 108 | 108 | 108 | 108 |

Group = 1

**Correlations**

| | | 5 | 16 | 19 | 23 |
|---|---|---|---|---|---|
| | Correlation Coefficient | 1.000 | .554(**) | -.008 | -.058 |
| 5 | Sig. (2-tailed) | . | .000 | .883 | .285 |
| | N | 344 | 340 | 339 | 340 |
| | Correlation Coefficient | .554(**) | 1.000 | .046 | -.080 |
| 16 | Sig. (2-tailed) | .000 | . | .400 | .142 |
| | N | 340 | 343 | 341 | 342 |
| | Correlation Coefficient | -.008 | .046 | 1.000 | .225(**) |
| 19 | Sig. (2-tailed) | .883 | .400 | . | .000 |
| | N | 339 | 341 | 342 | 341 |
| | Correlation Coefficient | -.058 | -.080 | .225(**) | 1.000 |
| 23 | Sig. (2-tailed) | .285 | .142 | .000 | . |
| | N | 340 | 342 | 341 | 344 |

Group = 2

**Lack of self-confidence**

21. I feel timid speaking English with other people.
20. When I'm encouraged, I try harder.
35. If I do badly in the SPE, I may stop my studies.

**Correlations**

| | | 21 | 20 | 35 |
|---|---|---|---|---|
| 21 | Correlation Coefficient | 1.000 | -.030 | .139 |
| | Sig. (2-tailed) | . | .760 | .153 |
| | N | 107 | 107 | 107 |
| 20 | Correlation Coefficient | -.030 | 1.000 | -.256(**) |
| | Sig. (2-tailed) | .760 | . | .007 |
| | N | 107 | 110 | 110 |
| 35 | Correlation Coefficient | .139 | -.256(**) | 1.000 |
| | Sig. (2-tailed) | .153 | .007 | . |
| | N | 107 | 110 | 110 |

Group = 1

**Correlations**

| | | 21 | 20 | 35 |
|---|---|---|---|---|
| 21 | Correlation Coefficient | 1.000 | .073 | .240(**) |
| | Sig. (2-tailed) | . | .186 | .000 |
| | N | 340 | 332 | 332 |
| 20 | Correlation Coefficient | .073 | 1.000 | -.111(*) |
| | Sig. (2-tailed) | .186 | . | .042 |
| | N | 332 | 340 | 334 |
| 35 | Correlation Coefficient | .240(**) | -.111(*) | 1.000 |
| | Sig. (2-tailed) | .000 | .042 | . |
| | N | 332 | 334 | 340 |

Group = 2

**Use of English**

39. I prefer the teacher to teach in English.
15. I prefer to use materials with explanation in English rather than those with explanation in Farsi.
5. Reading English newspapers and magazines
47. Reading English story books
51. Reading for pleasure
54. Reading as much as possible in English
45. Thinking in English while reading
67. Reading story books
66. Using words in sentences that I make
79. Learning and using grammar rules in new situations
1. Using instructional CD's
4. Watching English films
9. learning through games
20. watching English programmes on TV
6. listening to English programs on the radio
7. listening to English songs
2. Writing things like diaries in English
3. Speaking with native speakers
10. Using opportunities in class to speak in English
37. Talking to classmates or friends in English
18. Using materials with explanations completely in English, not in Farsi

**Correlations**

| | | 39 | 15 | 5 | 47 | 51 | 54 | 45 | 67 | 66 | 79 | 1 | 4 | 9 | 20 | 6 | 7 | 2 | 3 | 10 | 37 | 18 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 39 | Rho | 1.000 | .416(**) | .090 | .115 | .147 | .198 | .077 | .256(*) | .149 | .093 | -.007 | .158 | .030 | .173 | .206 | .011 | .138 | -.042 | .382(**) | .133 | .166 |
| | Sig.(2t) | . | .000 | .402 | .269 | .136 | .056 | .449 | .012 | .141 | .375 | .951 | .122 | .775 | .091 | .057 | .914 | .226 | .718 | .000 | .205 | .116 |
| | N | 109 | 107 | 88 | 95 | 104 | 94 | 99 | 96 | 99 | 93 | 70 | 97 | 95 | 97 | 86 | 97 | 79 | 78 | 92 | 93 | 91 |
| 15 | Rho | .416(**) | 1.000 | .206 | .351(**) | .301(**) | .354(**) | .311(**) | .265(**) | .169 | .101 | .176 | .320(**) | .118 | .303(**) | .289(**) | .166 | .264(*) | .183 | .251(*) | .227(*) | .340(**) |
| | Sig. (2t) | .000 | . | .054 | .000 | .002 | .001 | .002 | .009 | .095 | .336 | .151 | .001 | .256 | .003 | .007 | .106 | .019 | .114 | .016 | .029 | .001 |
| | N | 107 | 108 | 88 | 95 | 104 | 93 | 98 | 96 | 98 | 93 | 68 | 96 | 95 | 96 | 85 | 96 | 79 | 76 | 91 | 93 | 91 |
| 5 | Rho | .090 | .206 | 1.000 | .452(**) | .307(**) | .193 | .354(**) | .509(**) | .212 | .267(*) | .490(**) | .578(**) | .394(**) | .390(**) | .594(**) | .564(**) | .232 | .413(**) | .228(*) | .464(**) | .341(**) |
| | Sig. (2t) | .402 | .054 | . | .000 | .004 | .088 | .001 | .000 | .056 | .018 | .000 | .000 | .000 | .000 | .000 | .000 | .055 | .000 | .043 | .000 | .002 |
| | N | 88 | 88 | 89 | 81 | 87 | 79 | 83 | 82 | 82 | 78 | 62 | 86 | 80 | 86 | 79 | 86 | 69 | 69 | 79 | 79 | 83 |
| 47 | Rho | .115 | .351(**) | .452(**) | 1.000 | .422(**) | .341(**) | .494(**) | .617(**) | .252(*) | .095 | .177 | .482(**) | .169 | .401(**) | .405(**) | .350(**) | .326(**) | .231 | .273(*) | .518(**) | .288(**) |
| | Sig.(2t) | .269 | .000 | .000 | . | .000 | .001 | .000 | .000 | .015 | .393 | .181 | .000 | .120 | .000 | .000 | .001 | .005 | .055 | .010 | .000 | .007 |
| | N | 95 | 95 | 81 | 96 | 95 | 87 | 94 | 91 | 92 | 83 | 59 | 86 | 86 | 87 | 77 | 87 | 72 | 70 | 87 | 88 | 86 |
| 51 | Rho | .147 | .301(**) | .307(**) | .422(**) | 1.000 | .404(**) | .302(**) | .351(**) | .342(**) | .230(*) | .240 | .295(**) | .327(**) | .238(*) | .272(*) | .177 | .286(*) | .322(**) | .320(**) | .427(**) | .151 |
| | Sig.(2t) | .136 | .002 | .004 | .000 | . | .000 | .002 | .000 | .001 | .028 | .053 | .004 | .001 | .021 | .013 | .087 | .011 | .005 | .002 | .000 | .155 |
| | N | 104 | 104 | 87 | 95 | 105 | 93 | 98 | 95 | 98 | 92 | 66 | 94 | 95 | 94 | 83 | 95 | 78 | 76 | 91 | 92 | 90 |
| 54 | Rho | .198 | .354(**) | .193 | .341(**) | .404(**) | 1.000 | .389(**) | .414(**) | .092 | .341(**) | .076 | .246(*) | .216(*) | .377(**) | .143 | .146 | .265(*) | .151 | .249(*) | .165 | .106 |
| | Sig. (2t) | .056 | .001 | .088 | .001 | .000 | . | .000 | .000 | .393 | .002 | .556 | .023 | .047 | .000 | .223 | .178 | .024 | .217 | .021 | .135 | .339 |
| | N | 94 | 93 | 79 | 87 | 93 | 95 | 90 | 86 | 89 | 82 | 63 | 86 | 85 | 86 | 75 | 87 | 72 | 69 | 86 | 84 | 83 |
| 45 | Rho | .077 | .311(**) | .354(**) | .494(**) | .302(**) | .389(**) | 1.000 | .481(**) | .261(*) | .185 | .218 | .276(**) | .108 | .352(**) | .305(**) | .309(**) | .329(**) | .234(*) | .315(**) | .409(**) | .294(**) |
| | Sig.(2t) | .449 | .002 | .001 | .000 | .002 | .000 | . | .000 | .011 | .086 | .085 | .009 | .316 | .001 | .006 | .003 | .004 | .046 | .003 | .000 | .006 |
| | N | 99 | 98 | 83 | 94 | 98 | 90 | 100 | 93 | 95 | 87 | 63 | 89 | 89 | 90 | 79 | 90 | 76 | 73 | 88 | 90 | 86 |
| 67 | Rho | .256(*) | .265(**) | .509(**) | .617(**) | .351(**) | .414(**) | .481(**) | 1.000 | .439(**) | .073 | .162 | .436(**) | .147 | .388(**) | .559(**) | .420(**) | .279(*) | .327(**) | .279(*) | .495(**) | .307(**) |
| | Sig.(2t) | .012 | .009 | .000 | .000 | .000 | .000 | .000 | . | .000 | .505 | .215 | .000 | .175 | .000 | .000 | .000 | .018 | .006 | .011 | .000 | .005 |
| | N | 96 | 96 | 82 | 91 | 95 | 86 | 93 | 97 | 92 | 86 | 60 | 86 | 87 | 87 | 78 | 87 | 72 | 69 | 83 | 86 | 84 |
| 66 | Rho | .149 | .169 | .212 | .252(*) | .342(**) | .092 | .261(*) | .439(**) | 1.000 | .120 | .029 | .214(*) | .206 | .062 | .231(*) | .190 | .275(*) | .297(*) | .241(*) | .551(**) | .107 |
| | Sig. (2t) | .141 | .095 | .056 | .015 | .001 | .393 | .011 | .000 | . | .262 | .822 | .046 | .051 | .562 | .040 | .073 | .018 | .011 | .023 | .000 | .329 |
| | N | 99 | 98 | 82 | 92 | 98 | 89 | 95 | 92 | 100 | 89 | 64 | 88 | 90 | 89 | 79 | 90 | 74 | 72 | 89 | 89 | 85 |
| 79 | Rho | .093 | .101 | .267(*) | .095 | .230(*) | .341(**) | .185 | .073 | .120 | 1.000 | -.017 | .268(*) | .214(*) | .325(**) | .131 | .285(**) | .241(*) | -.031 | .118 | .082 | .163 |
| | Sig.(2t) | .375 | .336 | .018 | .393 | .028 | .002 | .086 | .505 | .262 | . | .898 | .014 | .049 | .003 | .264 | .009 | .045 | .801 | .300 | .463 | .154 |
| | N | 93 | 93 | 78 | 83 | 92 | 82 | 87 | 86 | 89 | 94 | 59 | 84 | 85 | 84 | 74 | 84 | 70 | 67 | 79 | 82 | 78 |
| 1 | Rho | -.007 | .176 | .490(**) | .177 | .240 | .076 | .218 | .162 | .029 | -.017 | 1.000 | .276(*) | .141 | .215 | .408(**) | .305(*) | .384(**) | .564(**) | .231 | .189 | .397(**) |
| | Sig.(2t) | .951 | .151 | .000 | .181 | .053 | .556 | .085 | .215 | .822 | .898 | . | .024 | .266 | .076 | .001 | .011 | .003 | .000 | .073 | .144 | .002 |
| | N | 70 | 68 | 62 | 59 | 66 | 63 | 63 | 60 | 64 | 59 | 70 | 67 | 64 | 69 | 63 | 68 | 58 | 53 | 61 | 61 | 61 |
| 4 | rho | .158 | .320(**) | .578(**) | .482(**) | .295(**) | .246(*) | .276(**) | .436(**) | .214(*) | .268(*) | .276(*) | 1.000 | .228(*) | .406(**) | .460(**) | .563(**) | .359(**) | .369(**) | .323(**) | .336(**) | .180 |
| | Sig. (2t) | .122 | .001 | .000 | .000 | .004 | .023 | .009 | .000 | .046 | .014 | .024 | . | .032 | .000 | .000 | .000 | .002 | .001 | .002 | .002 | .096 |
| | N | 97 | 96 | 86 | 86 | 94 | 86 | 89 | 86 | 88 | 84 | 67 | 97 | 88 | 93 | 84 | 93 | 74 | 76 | 86 | 86 | 86 |
| 9 | Rho | .030 | .118 | .394(**) | .169 | .327(**) | .216(*) | .108 | .147 | .206 | .214(*) | .141 | .228(*) | 1.000 | .175 | .307(**) | .168 | .187 | .300(*) | .167 | .373(**) | .095 |
| | Sig.(2t) | .775 | .256 | .000 | .120 | .001 | .047 | .316 | .175 | .051 | .049 | .266 | .032 | . | .103 | .006 | .114 | .111 | .010 | .124 | .000 | .389 |
| | N | 95 | 95 | 80 | 86 | 95 | 85 | 89 | 87 | 90 | 85 | 64 | 88 | 96 | 88 | 79 | 90 | 74 | 72 | 86 | 85 | 84 |
| 20 | Rho | .173 | .303(**) | .390(**) | .401(**) | .238(*) | .377(**) | .352(**) | .388(**) | .062 | .325(**) | .215 | .406(**) | .175 | 1.000 | .447(**) | .475(**) | .185 | .209 | .283(**) | .303(**) | .301(**) |
| | Sig.(2t) | .091 | .003 | .000 | .000 | .021 | .000 | .001 | .000 | .562 | .003 | .076 | .000 | .103 | . | .000 | .000 | .116 | .073 | .009 | .005 | .005 |
| | N | 97 | 96 | 86 | 87 | 94 | 86 | 90 | 87 | 89 | 84 | 69 | 93 | 88 | 97 | 83 | 93 | 73 | 74 | 85 | 86 | 87 |
| 6 | Rho | .206 | .289(**) | .594(**) | .405(**) | .272(*) | .143 | .305(**) | .559(**) | .231(*) | .131 | .408(**) | .460(**) | .307(**) | .447(**) | 1.000 | .449(**) | .309(*) | .368(**) | .303(**) | .431(**) | .222 |
| | Sig. (2t) | .057 | .007 | .000 | .000 | .013 | .223 | .006 | .000 | .040 | .264 | .001 | .000 | .006 | .000 | . | .000 | .010 | .002 | .007 | .000 | .050 |
| | N | 86 | 85 | 79 | 77 | 83 | 75 | 79 | 78 | 79 | 74 | 63 | 84 | 79 | 83 | 86 | 83 | 68 | 68 | 77 | 78 | 78 |
| 7 | Rho | .011 | .166 | .564(**) | .350(**) | .177 | .146 | .309(**) | .420(**) | .190 | .285(**) | .305(*) | .563(**) | .168 | .475(**) | .449(**) | 1.000 | .344(**) | .332(**) | .218(*) | .231(*) | .249(*) |
| | Sig.(2t) | .914 | .106 | .000 | .001 | .087 | .178 | .003 | .000 | .073 | .009 | .011 | .000 | .114 | .000 | .000 | . | .003 | .004 | .041 | .032 | .019 |
| | N | 97 | 96 | 86 | 87 | 95 | 87 | 90 | 87 | 90 | 84 | 68 | 93 | 90 | 93 | 83 | 98 | 74 | 75 | 88 | 87 | 89 |
| 2 | Rho | .138 | .264(*) | .232 | .326(**) | .286(*) | .265(*) | .329(**) | .279(*) | .275(*) | .241(*) | .384(**) | .359(**) | .187 | .185 | .309(*) | .344(**) | 1.000 | .438(**) | .387(**) | .374(**) | .243(*) |
| | Sig.(2t) | .226 | .019 | .055 | .005 | .011 | .024 | .004 | .018 | .018 | .045 | .003 | .002 | .111 | .116 | .010 | .003 | . | .000 | .001 | .001 | .041 |
| | N | 79 | 79 | 69 | 72 | 78 | 72 | 76 | 72 | 74 | 70 | 58 | 74 | 74 | 73 | 68 | 74 | 80 | 63 | 73 | 72 | 71 |
| 3 | Rho | -.042 | .183 | .413(**) | .231 | .322(**) | .151 | .234(*) | .327(**) | .297(*) | -.031 | .564(**) | .369(**) | .300(*) | .209 | .368(**) | .332(**) | .438(**) | 1.000 | .306(**) | .358(**) | .261(*) |
| | Sig. (2t) | .718 | .114 | .000 | .055 | .005 | .217 | .046 | .006 | .011 | .801 | .000 | .001 | .010 | .073 | .002 | .004 | .000 | . | .009 | .002 | .029 |
| | N | 78 | 76 | 69 | 70 | 76 | 69 | 73 | 69 | 72 | 67 | 53 | 76 | 72 | 74 | 68 | 75 | 63 | 78 | 72 | 73 | 70 |
| 10 | Rho | .382(**) | .251(*) | .228(*) | .273(*) | .320(**) | .249(*) | .315(**) | .279(*) | .241(*) | .118 | .231 | .323(**) | .167 | .283(**) | .303(**) | .218(*) | .387(**) | .306(**) | 1.000 | .346(**) | .246(*) |
| | Sig. (2t) | .000 | .016 | .043 | .010 | .002 | .021 | .003 | .011 | .023 | .300 | .073 | .002 | .124 | .009 | .007 | .041 | .001 | .009 | . | .001 | .021 |
| | N | 92 | 91 | 79 | 87 | 91 | 86 | 88 | 83 | 89 | 79 | 61 | 86 | 86 | 85 | 77 | 88 | 73 | 72 | 93 | 85 | 87 |
| 37 | Rho | .133 | .227(*) | .464(**) | .518(**) | .427(**) | .165 | .409(**) | .495(**) | .551(**) | .082 | .189 | .336(**) | .373(**) | .303(**) | .431(**) | .231(*) | .374(**) | .358(**) | .346(**) | 1.000 | .204 |
| | Sig.(2t) | .205 | .029 | .000 | .000 | .000 | .135 | .000 | .000 | .000 | .463 | .144 | .002 | .000 | .005 | .000 | .032 | .001 | .002 | .001 | . | .062 |
| | N | 93 | 93 | 79 | 88 | 92 | 84 | 90 | 86 | 89 | 82 | 61 | 86 | 85 | 86 | 78 | 87 | 72 | 73 | 85 | 94 | 85 |
| 18 | Rho | .166 | .340(**) | .341(**) | .288(**) | .151 | .106 | .294(**) | .307(**) | .107 | .163 | .397(**) | .180 | .095 | .301(**) | .222 | .249(*) | .243(*) | .261(*) | .246(*) | .204 | 1.000 |
| | Sig.(2t) | .116 | .001 | .002 | .007 | .155 | .339 | .006 | .005 | .329 | .154 | .002 | .096 | .389 | .005 | .050 | .019 | .041 | .029 | .021 | .062 | . |
| | N | 91 | 91 | 83 | 86 | 90 | 83 | 86 | 84 | 85 | 78 | 61 | 86 | 84 | 87 | 78 | 89 | 71 | 70 | 87 | 85 | 92 |

Group = 1

**Use of English (Continued)**

**Correlations**

| | | 39 | 15 | 5 | 47 | 51 | 54 | 45 | 67 | 66 | 79 | 1 | 4 | 9 | 20 | 6 | 7 | 2 | 3 | 10 | 37 | 18 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 39 | Rho | 1.000 | .386(**) | .221(**) | .240(**) | .199(**) | .168(**) | .210(**) | .312(**) | .268(**) | .122(*) | .148(*) | .120(*) | .134(*) | .102 | .180(**) | .093 | .078 | .182(**) | .216(**) | .258(**) | .261(**) |
| | Sig.(2t) | . | .000 | .000 | .000 | .001 | .004 | .000 | .000 | .000 | .035 | .029 | .042 | .025 | .086 | .003 | .108 | .224 | .004 | .000 | .000 | .000 |
| | N | 345 | 338 | 260 | 279 | 300 | 292 | 302 | 279 | 289 | 301 | 218 | 287 | 280 | 285 | 267 | 297 | 245 | 247 | 287 | 276 | 286 |
| 15 | Rho | .386(**) | 1.000 | .357(**) | .200(**) | .221(**) | .242(**) | .158(**) | .255(**) | .246(**) | .050 | .211(**) | .262(**) | .095 | .125(*) | .235(**) | .157(**) | .168(**) | .227(**) | .152(*) | .192(**) | .392(**) |
| | Sig. (2t) | .000 | . | .000 | .001 | .000 | .000 | .006 | .000 | .000 | .390 | .002 | .000 | .115 | .037 | .000 | .007 | .009 | .000 | .011 | .001 | .000 |
| | N | 338 | 339 | 255 | 272 | 294 | 287 | 297 | 273 | 283 | 295 | 215 | 283 | 275 | 280 | 262 | 292 | 239 | 242 | 282 | 271 | 281 |
| 5 | Rho | .221(**) | .357(**) | 1.000 | .441(**) | .349(**) | .238(**) | .272(**) | .541(**) | .319(**) | .107 | .424(**) | .515(**) | .287(**) | .388(**) | .484(**) | .361(**) | .413(**) | .445(**) | .270(**) | .314(**) | .540(**) |
| | Sig. (2t) | .000 | .000 | . | .000 | .000 | .000 | .000 | .000 | .000 | .108 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 |
| | N | 260 | 255 | 262 | 226 | 236 | 223 | 238 | 222 | 226 | 228 | 190 | 249 | 234 | 238 | 242 | 247 | 212 | 216 | 232 | 225 | 234 |
| 47 | Rho | .240(**) | .200(**) | .441(**) | 1.000 | .429(**) | .287(**) | .353(**) | .656(**) | .363(**) | .147(*) | .216(**) | .255(**) | .256(**) | .307(**) | .277(**) | .273(**) | .287(**) | .281(**) | .328(**) | .350(**) | .431(**) |
| | Sig.(2t) | .000 | .001 | .000 | . | .000 | .000 | .000 | .000 | .000 | .018 | .003 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 |
| | N | 279 | 272 | 226 | 280 | 262 | 251 | 260 | 251 | 250 | 258 | 186 | 243 | 238 | 243 | 230 | 253 | 210 | 213 | 241 | 240 | 244 |
| 51 | Rho | .199(**) | .221(**) | .349(**) | .429(**) | 1.000 | .374(**) | .336(**) | .420(**) | .352(**) | .211(**) | .159(*) | .253(**) | .292(**) | .245(**) | .269(**) | .244(**) | .209(**) | .331(**) | .300(**) | .398(**) | .283(**) |
| | Sig.(2t) | .001 | .000 | .000 | .000 | . | .000 | .000 | .000 | .000 | .000 | .027 | .000 | .000 | .000 | .000 | .000 | .002 | .000 | .000 | .000 | .000 |
| | N | 300 | 294 | 236 | 262 | 301 | 273 | 281 | 259 | 271 | 275 | 195 | 257 | 254 | 260 | 240 | 267 | 217 | 222 | 262 | 253 | 268 |
| 54 | Rho | .168(**) | .242(**) | .238(**) | .287(**) | .374(**) | 1.000 | .284(**) | .260(**) | .344(**) | .270(**) | .183(*) | .239(**) | .285(**) | .306(**) | .245(**) | .285(**) | .170(*) | .257(**) | .301(**) | .429(**) | .328(**) |
| | Sig. (2t) | .004 | .000 | .000 | .000 | .000 | . | .000 | .000 | .000 | .000 | .013 | .000 | .000 | .000 | .000 | .000 | .013 | .000 | .000 | .000 | .000 |
| | N | 292 | 287 | 223 | 251 | 273 | 293 | 270 | 254 | 258 | 269 | 184 | 248 | 245 | 250 | 231 | 259 | 212 | 217 | 252 | 243 | 257 |
| 45 | Rho | .210(**) | .158(**) | .272(**) | .353(**) | .336(**) | .284(**) | 1.000 | .291(**) | .335(**) | .245(**) | .158(*) | .160(*) | .244(**) | .262(**) | .290(**) | .126(*) | .252(**) | .179(**) | .391(**) | .315(**) | .254(**) |
| | Sig.(2t) | .000 | .006 | .000 | .000 | .000 | .000 | . | .000 | .000 | .000 | .026 | .010 | .000 | .000 | .000 | .038 | .000 | .007 | .000 | .000 | .000 |
| | N | 302 | 297 | 238 | 260 | 281 | 270 | 303 | 261 | 269 | 278 | 199 | 258 | 257 | 259 | 242 | 270 | 222 | 225 | 262 | 257 | 259 |
| 67 | Rho | .312(**) | .255(**) | .541(**) | .656(**) | .420(**) | .260(**) | .291(**) | 1.000 | .465(**) | .049 | .274(**) | .314(**) | .299(**) | .375(**) | .379(**) | .287(**) | .336(**) | .375(**) | .322(**) | .353(**) | .500(**) |
| | Sig.(2t) | .000 | .000 | .000 | .000 | .000 | .000 | .000 | . | .000 | .434 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 |
| | N | 279 | 273 | 222 | 251 | 259 | 254 | 261 | 279 | 259 | 258 | 181 | 242 | 235 | 243 | 229 | 247 | 206 | 211 | 240 | 239 | 242 |
| 66 | Rho | .268(**) | .246(**) | .319(**) | .363(**) | .352(**) | .344(**) | .335(**) | .465(**) | 1.000 | .256(**) | .189(*) | .284(**) | .329(**) | .245(**) | .346(**) | .290(**) | .353(**) | .266(**) | .417(**) | .445(**) | .307(**) |
| | Sig. (2t) | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | . | .000 | .010 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 |
| | N | 289 | 283 | 226 | 250 | 271 | 258 | 269 | 259 | 289 | 264 | 185 | 247 | 243 | 249 | 233 | 258 | 209 | 212 | 245 | 245 | 246 |
| 79 | Rho | .122(*) | .050 | .107 | .147(*) | .211(**) | .270(**) | .245(**) | .049 | .256(**) | 1.000 | .129 | .134(*) | .231(**) | .180(**) | .282(**) | .049 | .181(**) | .129 | .244(**) | .303(**) | .173(**) |
| | Sig.(2t) | .035 | .390 | .108 | .018 | .000 | .000 | .000 | .434 | .000 | . | .073 | .032 | .000 | .004 | .000 | .428 | .008 | .055 | .000 | .000 | .005 |
| | N | 301 | 295 | 228 | 258 | 275 | 269 | 278 | 258 | 264 | 302 | 193 | 254 | 254 | 254 | 238 | 266 | 217 | 221 | 262 | 249 | 258 |
| 1 | Rho | .148(*) | .211(**) | .424(**) | .216(**) | .159(*) | .183(*) | .158(*) | .274(**) | .189(*) | .129 | 1.000 | .372(**) | .308(**) | .296(**) | .485(**) | .325(**) | .491(**) | .380(**) | .243(**) | .299(**) | .385(**) |
| | Sig.(2t) | .029 | .002 | .000 | .003 | .027 | .013 | .026 | .000 | .010 | .073 | . | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .001 | .000 |
| | N | 218 | 215 | 190 | 186 | 195 | 184 | 199 | 181 | 185 | 193 | 219 | 200 | 195 | 194 | 189 | 206 | 192 | 187 | 195 | 188 | 189 |
| 4 | rho | .120(*) | .262(**) | .515(**) | .255(**) | .253(**) | .239(**) | .160(*) | .314(**) | .284(**) | .134(*) | .372(**) | 1.000 | .374(**) | .466(**) | .360(**) | .561(**) | .207(**) | .357(**) | .217(**) | .256(**) | .388(**) |
| | Sig. (2t) | .042 | .000 | .000 | .000 | .000 | .000 | .010 | .000 | .000 | .032 | .000 | . | .000 | .000 | .000 | .000 | .002 | .000 | .000 | .000 | .000 |
| | N | 287 | 283 | 249 | 243 | 257 | 248 | 258 | 242 | 247 | 254 | 200 | 289 | 251 | 263 | 252 | 272 | 224 | 231 | 255 | 244 | 249 |
| 9 | Rho | .134(*) | .095 | .287(**) | .256(**) | .292(**) | .285(**) | .244(**) | .299(**) | .329(**) | .231(**) | .308(**) | .374(**) | 1.000 | .386(**) | .308(**) | .340(**) | .265(**) | .337(**) | .385(**) | .345(**) | .222(**) |
| | Sig.(2t) | .025 | .115 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | . | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 |
| | N | 280 | 275 | 234 | 238 | 254 | 245 | 257 | 235 | 243 | 254 | 195 | 251 | 283 | 252 | 238 | 263 | 217 | 221 | 250 | 237 | 248 |
| 20 | Rho | .102 | .125(*) | .388(**) | .307(**) | .245(**) | .306(**) | .262(**) | .375(**) | .245(**) | .180(**) | .296(**) | .466(**) | .386(**) | 1.000 | .367(**) | .442(**) | .134(*) | .219(**) | .313(**) | .333(**) | .359(**) |
| | Sig.(2t) | .086 | .037 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .004 | .000 | .000 | .000 | . | .000 | .000 | .049 | .001 | .000 | .000 | .000 |
| | N | 285 | 280 | 238 | 243 | 260 | 250 | 259 | 243 | 249 | 254 | 194 | 263 | 252 | 287 | 246 | 269 | 217 | 221 | 250 | 245 | 260 |
| 6 | Rho | .180(**) | .235(**) | .484(**) | .277(**) | .269(**) | .245(**) | .290(**) | .379(**) | .346(**) | .282(**) | .485(**) | .360(**) | .308(**) | .367(**) | 1.000 | .295(**) | .304(**) | .363(**) | .352(**) | .414(**) | .376(**) |
| | Sig. (2t) | .003 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | . | .000 | .000 | .000 | .000 | .000 | .000 |
| | N | 267 | 262 | 242 | 230 | 240 | 231 | 242 | 229 | 233 | 238 | 189 | 252 | 238 | 246 | 269 | 258 | 216 | 223 | 239 | 230 | 241 |
| 7 | Rho | .093 | .157(**) | .361(**) | .273(**) | .244(**) | .285(**) | .126(*) | .287(**) | .290(**) | .049 | .325(**) | .561(**) | .340(**) | .442(**) | .295(**) | 1.000 | .191(**) | .255(**) | .155(*) | .268(**) | .304(**) |
| | Sig.(2t) | .108 | .007 | .000 | .000 | .000 | .000 | .038 | .000 | .000 | .428 | .000 | .000 | .000 | .000 | .000 | . | .003 | .000 | .012 | .000 | .000 |
| | N | 297 | 292 | 247 | 253 | 267 | 259 | 270 | 247 | 258 | 266 | 206 | 272 | 263 | 269 | 258 | 299 | 232 | 237 | 260 | 249 | 261 |
| 2 | Rho | .078 | .168(**) | .413(**) | .287(**) | .209(**) | .170(*) | .252(**) | .336(**) | .353(**) | .181(**) | .491(**) | .207(**) | .265(**) | .134(*) | .304(**) | .191(**) | 1.000 | .393(**) | .266(**) | .300(**) | .285(**) |
| | Sig.(2t) | .224 | .009 | .000 | .000 | .002 | .013 | .000 | .000 | .000 | .008 | .000 | .002 | .000 | .049 | .000 | .003 | . | .000 | .000 | .000 | .000 |
| | N | 245 | 239 | 212 | 210 | 217 | 212 | 222 | 206 | 209 | 217 | 192 | 224 | 217 | 217 | 216 | 232 | 246 | 215 | 218 | 209 | 213 |
| 3 | Rho | .182(**) | .227(**) | .445(**) | .281(**) | .331(**) | .257(**) | .179(**) | .375(**) | .266(**) | .129 | .380(**) | .357(**) | .337(**) | .219(**) | .363(**) | .255(**) | .393(**) | 1.000 | .380(**) | .303(**) | .315(**) |
| | Sig. (2t) | .004 | .000 | .000 | .000 | .000 | .000 | .007 | .000 | .000 | .055 | .000 | .000 | .000 | .001 | .000 | .000 | .000 | . | .000 | .000 | .000 |
| | N | 247 | 242 | 216 | 213 | 222 | 217 | 225 | 211 | 212 | 221 | 187 | 231 | 221 | 221 | 223 | 237 | 215 | 248 | 227 | 212 | 220 |
| 10 | Rho | .216(**) | .152(*) | .270(**) | .328(**) | .300(**) | .301(**) | .391(**) | .322(**) | .417(**) | .244(**) | .243(**) | .217(**) | .385(**) | .313(**) | .352(**) | .155(*) | .266(**) | .380(**) | 1.000 | .459(**) | .262(**) |
| | Sig. (2t) | .000 | .011 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .001 | .000 | .000 | .000 | .000 | .012 | .000 | .000 | . | .000 | .000 |
| | N | 287 | 282 | 232 | 241 | 262 | 252 | 262 | 240 | 245 | 262 | 195 | 255 | 250 | 250 | 239 | 260 | 218 | 227 | 289 | 243 | 253 |
| 37 | Rho | .258(**) | .192(**) | .314(**) | .350(**) | .398(**) | .429(**) | .315(**) | .353(**) | .445(**) | .303(**) | .299(**) | .256(**) | .345(**) | .333(**) | .414(**) | .268(**) | .300(**) | .303(**) | .459(**) | 1.000 | .367(**) |
| | Sig.(2t) | .000 | .001 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | . | .000 |
| | N | 276 | 271 | 225 | 240 | 253 | 243 | 257 | 239 | 245 | 249 | 188 | 244 | 237 | 245 | 230 | 249 | 209 | 212 | 243 | 278 | 244 |
| 18 | Rho | .261(**) | .392(**) | .540(**) | .431(**) | .283(**) | .328(**) | .254(**) | .500(**) | .307(**) | .173(**) | .385(**) | .388(**) | .222(**) | .359(**) | .376(**) | .304(**) | .285(**) | .315(**) | .262(**) | .367(**) | 1.000 |
| | Sig.(2t) | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .005 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | . |
| | N | 286 | 281 | 234 | 244 | 268 | 257 | 259 | 242 | 246 | 258 | 189 | 249 | 248 | 260 | 241 | 261 | 213 | 220 | 253 | 244 | 288 |

Group = 2

**Time management in test-taking**

18. Effective use of time during the SPE Exam is essential.
34. Answering sample exam papers under timed conditions similar to the real exam

**Correlations**

| | | 18 | 34 |
|---|---|---|---|
| 18 | Correlation Coefficient | 1.000 | .133 |
| | Sig. (2-tailed) | . | .183 |
| | N | 110 | 102 |
| 34 | Correlation Coefficient | .133 | 1.000 |
| | Sig. (2-tailed) | .183 | . |
| | N | 102 | 103 |

Group = 1

**Correlations**

| | | 18 | 34 |
|---|---|---|---|
| 18 | Correlation Coefficient | 1.000 | .197(**) |
| | Sig. (2-tailed) | . | .001 |
| | N | 346 | 307 |
| 34 | Correlation Coefficient | .197(**) | 1.000 |
| | Sig. (2-tailed) | .001 | . |
| | N | 307 | 308 |

Group = 2

**General courses**

7. Preparation for General Courses like Persian is useful for success in the SPE Exam.
31. Studying "general courses"
62. Which of the following General subjects do you try to learn more than other subjects in order to be admitted to the university of your choice? Please rank them according to importance to your success (1=most important, 4=least important).

**Correlations**

| | | 7 | 31 | 62 |
|---|---|---|---|---|
| 7 | Correlation Coefficient | 1.000 | .140 | . |
| | Sig. (2-tailed) | . | .155 | . |
| | N | 109 | 105 | 0 |
| 31 | Correlation Coefficient | .140 | 1.000 | . |
| | Sig. (2-tailed) | .155 | . | . |
| | N | 105 | 106 | 0 |
| 62 | Correlation Coefficient | . | . | . |
| | Sig. (2-tailed) | . | . | . |
| | N | 0 | 0 | 0 |

Group = 1

**Correlations**

| | | 7 | 31 | 62 |
|---|---|---|---|---|
| 7 | Correlation Coefficient | 1.000 | .159(**) | . |
| | Sig. (2-tailed) | . | .005 | . |
| | N | 339 | 316 | 0 |
| 31 | Correlation Coefficient | .159(**) | 1.000 | . |
| | Sig. (2-tailed) | .005 | . | . |
| | N | 316 | 323 | 0 |
| 62 | Correlation Coefficient | . | . | . |
| | Sig. (2-tailed) | . | . | . |
| | N | 0 | 0 | 0 |

Group = 2

**Appropriacy of the SPE**

2. The SPE test items are appropriate.
12. The SPE Exam is an appropriate means of evaluating our ability in English.
49. The SPE Exam is a fair Exam.

**Correlations**

| | | 2 | 12 | 49 |
|---|---|---|---|---|
| 2 | Correlation Coefficient | 1.000 | .299(**) | .341(**) |
| | Sig. (2-tailed) | . | .002 | .000 |
| | N | 109 | 106 | 105 |
| 12 | Correlation Coefficient | .299(**) | 1.000 | .128 |
| | Sig. (2-tailed) | .002 | . | .192 |
| | N | 106 | 108 | 105 |
| 49 | Correlation Coefficient | .341(**) | .128 | 1.000 |
| | Sig. (2-tailed) | .000 | .192 | . |
| | N | 105 | 105 | 107 |

Group = 1

**Correlations**

| | | 2 | 12 | 49 |
|---|---|---|---|---|
| 2 | Correlation Coefficient | 1.000 | .204(**) | .275(**) |
| | Sig. (2-tailed) | . | .000 | .000 |
| | N | 342 | 337 | 338 |
| 12 | Correlation Coefficient | .204(**) | 1.000 | .188(**) |
| | Sig. (2-tailed) | .000 | . | .001 |
| | N | 337 | 341 | 338 |
| 49 | Correlation Coefficient | .275(**) | .188(**) | 1.000 |
| | Sig. (2-tailed) | .000 | .001 | . |
| | N | 338 | 338 | 343 |

Group = 2

**Test-taking skills**
17. Test- taking skills are useful for test preparation.
16. Learning examination skills

**Correlations**

| | | 17 | 16 |
|---|---|---|---|
| 17 | Correlation Coefficient | 1.000 | .038 |
| | Sig. (2-tailed) | . | .704 |
| | N | 110 | 103 |
| 16 | Correlation Coefficient | .038 | 1.000 |
| | Sig. (2-tailed) | .704 | . |
| | N | 103 | 103 |

Group = 1

**Correlations**

| | | 17 | 16 |
|---|---|---|---|
| 17 | Correlation Coefficient | 1.000 | .245(**) |
| | Sig. (2-tailed) | . | .000 |
| | N | 344 | 316 |
| 16 | Correlation Coefficient | .245(**) | 1.000 |
| | Sig. (2-tailed) | .000 | . |
| | N | 316 | 319 |

Group = 2

**Treating mistakes/errors**
26. I try to learn from my mistakes.
81. Using grammar rules even if I don't know them exactly

**Correlations**

| | | 26 | 81 |
|---|---|---|---|
| 26 | Correlation Coefficient | 1.000 | -.196 |
| | Sig. (2-tailed) | . | .064 |
| | N | 107 | 90 |
| 81 | Correlation Coefficient | -.196 | 1.000 |
| | Sig. (2-tailed) | .064 | . |
| | N | 90 | 93 |

Group = 1

**Correlations**

| | | 26 | 81 |
|---|---|---|---|
| 26 | Correlation Coefficient | 1.000 | .054 |
| | Sig. (2-tailed) | . | .354 |
| | N | 342 | 297 |
| 81 | Correlation Coefficient | .054 | 1.000 |
| | Sig. (2-tailed) | .354 | . |
| | N | 297 | 302 |

Group = 2

**Communication vs. accuracy/errors**

7. It is important to speak English with an excellent pronunciation.
9. You shouldn't say anything in English until you can say it correctly.
22. If beginning students are permitted to make errors in English, it will be difficult for them to speak correctly later on.

**Correlations**

| | | 7 | 9 | 22 |
|---|---|---|---|---|
| 7 | Correlation Coefficient | 1.000 | .225(*) | .179 |
| | Sig. (2-tailed) | . | .018 | .063 |
| | N | 110 | 110 | 109 |
| 9 | Correlation Coefficient | .225(*) | 1.000 | .230(*) |
| | Sig. (2-tailed) | .018 | . | .016 |
| | N | 110 | 110 | 109 |
| 22 | Correlation Coefficient | .179 | .230(*) | 1.000 |
| | Sig. (2-tailed) | .063 | .016 | . |
| | N | 109 | 109 | 109 |

Group = 1

**Correlations**

| | | 7 | 9 | 22 |
|---|---|---|---|---|
| 7 | Correlation Coefficient | 1.000 | .154(**) | .161(**) |
| | Sig. (2-tailed) | . | .005 | .003 |
| | N | 345 | 338 | 340 |
| 9 | Correlation Coefficient | .154(**) | 1.000 | .167(**) |
| | Sig. (2-tailed) | .005 | . | .002 |
| | N | 338 | 341 | 337 |
| 22 | Correlation Coefficient | .161(**) | .167(**) | 1.000 |
| | Sig. (2-tailed) | .003 | .002 | . |
| | N | 340 | 337 | 343 |

Group = 2

**Context**

14. It's o.k. to guess if you don't know a word in English.
42. Trying to guess the meaning of the new words from context
44. Skimming the text first to get the main idea and then going back to read it more carefully
56. Paying attention to the topic of the text
52. Reading without looking up all the words I don't know
49. Using comprehension questions to get a general idea of the passage
65. Reading passages
63. Learning meanings through examples in which the new word is used
82. Analysing grammar rules in texts

**Correlations**

| | | 14 | 42 | 44 | 56 | 52 | 49 | 65 | 63 | 82 |
|---|---|---|---|---|---|---|---|---|---|---|
| 14 | rho | 1.000 | .266(**) | .100 | .095 | .178 | .194(*) | -.038 | .038 | .016 |
| | Sig. (2-tailed) | . | .007 | .312 | .335 | .073 | .050 | .698 | .700 | .878 |
| | N | 110 | 103 | 104 | 105 | 103 | 103 | 106 | 106 | 95 |
| 42 | rho | .266(**) | 1.000 | .357(**) | .248(*) | .269(**) | .332(**) | .154 | .323(**) | .225(*) |
| | Sig. (2-tailed) | .007 | . | .000 | .012 | .007 | .001 | .126 | .001 | .033 |
| | N | 103 | 103 | 102 | 102 | 100 | 100 | 100 | 102 | 90 |
| 44 | rho | .100 | .357(**) | 1.000 | .332(**) | .194 | .297(**) | .176 | .213(*) | .119 |
| | Sig. (2-tailed) | .312 | .000 | . | .001 | .052 | .003 | .078 | .032 | .261 |
| | N | 104 | 102 | 105 | 102 | 101 | 101 | 101 | 102 | 91 |
| 56 | rho | .095 | .248(*) | .332(**) | 1.000 | .093 | .211(*) | .214(*) | .247(*) | .273(**) |
| | Sig. (2-tailed) | .335 | .012 | .001 | . | .355 | .034 | .031 | .012 | .009 |
| | N | 105 | 102 | 102 | 105 | 101 | 101 | 102 | 102 | 91 |
| 52 | rho | .178 | .269(**) | .194 | .093 | 1.000 | .104 | .222(*) | .158 | -.070 |
| | Sig. (2-tailed) | .073 | .007 | .052 | .355 | . | .299 | .027 | .115 | .514 |
| | N | 103 | 100 | 101 | 101 | 103 | 101 | 100 | 101 | 90 |
| 49 | rho | .194(*) | .332(**) | .297(**) | .211(*) | .104 | 1.000 | .297(**) | .346(**) | .439(**) |
| | Sig. (2-tailed) | .050 | .001 | .003 | .034 | .299 | . | .003 | .000 | .000 |
| | N | 103 | 100 | 101 | 101 | 101 | 103 | 100 | 101 | 90 |
| 65 | rho | -.038 | .154 | .176 | .214(*) | .222(*) | .297(**) | 1.000 | .405(**) | .337(**) |
| | Sig. (2-tailed) | .698 | .126 | .078 | .031 | .027 | .003 | . | .000 | .001 |
| | N | 106 | 100 | 101 | 102 | 100 | 100 | 106 | 104 | 94 |
| 63 | rho | .038 | .323(**) | .213(*) | .247(*) | .158 | .346(**) | .405(**) | 1.000 | .437(**) |
| | Sig. (2-tailed) | .700 | .001 | .032 | .012 | .115 | .000 | .000 | . | .000 |
| | N | 106 | 102 | 102 | 102 | 101 | 101 | 104 | 106 | 94 |
| 82 | rho | .016 | .225(*) | .119 | .273(**) | -.070 | .439(**) | .337(**) | .437(**) | 1.000 |
| | Sig. (2-tailed) | .878 | .033 | .261 | .009 | .514 | .000 | .001 | .000 | . |
| | N | 95 | 90 | 91 | 91 | 90 | 90 | 94 | 94 | 95 |

Group = 1

**Correlations**

| | | 14 | 42 | 44 | 56 | 52 | 49 | 65 | 63 | 82 |
|---|---|---|---|---|---|---|---|---|---|---|
| 14 | rho | 1.000 | .360(**) | .167(**) | .201(**) | .203(**) | .178(**) | .156(**) | .151(**) | -.008 |
| | Sig. (2-tailed) | . | .000 | .003 | .000 | .000 | .002 | .006 | .008 | .889 |
| | N | 345 | 320 | 308 | 318 | 308 | 298 | 310 | 308 | 294 |
| 42 | rho | .360(**) | 1.000 | .340(**) | .254(**) | .140(*) | .297(**) | .248(**) | .214(**) | .127(*) |
| | Sig. (2-tailed) | .000 | . | .000 | .000 | .015 | .000 | .000 | .000 | .031 |
| | N | 320 | 323 | 302 | 312 | 302 | 294 | 301 | 302 | 289 |
| 44 | rho | .167(**) | .340(**) | 1.000 | .256(**) | .167(**) | .251(**) | .239(**) | .303(**) | .232(**) |
| | Sig. (2-tailed) | .003 | .000 | . | .000 | .004 | .000 | .000 | .000 | .000 |
| | N | 308 | 302 | 310 | 302 | 293 | 287 | 292 | 290 | 276 |
| 56 | rho | .201(**) | .254(**) | .256(**) | 1.000 | .125(*) | .358(**) | .206(**) | .337(**) | .294(**) |
| | Sig. (2-tailed) | .000 | .000 | .000 | . | .029 | .000 | .000 | .000 | .000 |
| | N | 318 | 312 | 302 | 320 | 305 | 296 | 302 | 302 | 286 |
| 52 | rho | .203(**) | .140(*) | .167(**) | .125(*) | 1.000 | .119(*) | .142(*) | .071 | .008 |
| | Sig. (2-tailed) | .000 | .015 | .004 | .029 | . | .045 | .015 | .224 | .892 |
| | N | 308 | 302 | 293 | 305 | 310 | 286 | 290 | 292 | 276 |
| 49 | rho | .178(**) | .297(**) | .251(**) | .358(**) | .119(*) | 1.000 | .266(**) | .362(**) | .184(**) |
| | Sig. (2-tailed) | .002 | .000 | .000 | .000 | .045 | . | .000 | .000 | .002 |
| | N | 298 | 294 | 287 | 296 | 286 | 301 | 283 | 285 | 271 |
| 65 | rho | .156(**) | .248(**) | .239(**) | .206(**) | .142(*) | .266(**) | 1.000 | .353(**) | .249(**) |
| | Sig. (2-tailed) | .006 | .000 | .000 | .000 | .015 | .000 | . | .000 | .000 |
| | N | 310 | 301 | 292 | 302 | 290 | 283 | 312 | 300 | 279 |
| 63 | rho | .151(**) | .214(**) | .303(**) | .337(**) | .071 | .362(**) | .353(**) | 1.000 | .262(**) |
| | Sig. (2-tailed) | .008 | .000 | .000 | .000 | .224 | .000 | .000 | . | .000 |
| | N | 308 | 302 | 290 | 302 | 292 | 285 | 300 | 311 | 282 |
| 82 | rho | -.008 | .127(*) | .232(**) | .294(**) | .008 | .184(**) | .249(**) | .262(**) | 1.000 |
| | Sig. (2-tailed) | .889 | .031 | .000 | .000 | .892 | .002 | .000 | .000 | . |
| | N | 294 | 289 | 276 | 286 | 276 | 271 | 279 | 282 | 296 |

Group = 2

**Reading skills**
43. Practicing speed reading
48. Trying to improve concentration

**Correlation**

| | | 43 | 48 |
|---|---|---|---|
| 43 | Correlation Coefficient | 1.000 | .418(**) |
| | Sig. (2-tailed) | . | .000 |
| | N | 102 | 95 |
| 48 | Correlation Coefficient | .418(**) | 1.000 |
| | Sig. (2-tailed) | .000 | . |
| | N | 95 | 99 |

Group = 1

**Correlation**

| | | 43 | 48 |
|---|---|---|---|
| 43 | Correlation Coefficient | 1.000 | .301(**) |
| | Sig. (2-tailed) | . | .000 |
| | N | 315 | 293 |
| 48 | Correlation Coefficient | .301(**) | 1.000 |
| | Sig. (2-tailed) | .000 | . |
| | N | 293 | 305 |

Group = 2

**A. Use of mother tongue**
**A1. Translation**
28. The most important part of learning English is learning how to translate from my native language.
39. Translating reading texts into Farsi
**A2. Language of materials**
33. Using sources with explanations in Farsi

**Correlations**

| | | 28 | 39 | 33 |
|---|---|---|---|---|
| 28 | Correlation Coefficient | 1.000 | .285(**) | .013 |
| | Sig. (2-tailed) | . | .004 | .896 |
| | N | 108 | 101 | 98 |
| 39 | Correlation Coefficient | .285(**) | 1.000 | .236(*) |
| | Sig. (2-tailed) | .004 | . | .020 |
| | N | 101 | 102 | 96 |
| 33 | Correlation Coefficient | .013 | .236(*) | 1.000 |
| | Sig. (2-tailed) | .896 | .020 | . |
| | N | 98 | 96 | 100 |

Group = 1

**Correlations**

| | | 28 | 39 | 33 |
|---|---|---|---|---|
| 28 | Correlation Coefficient | 1.000 | .088 | .127(*) |
| | Sig. (2-tailed) | . | .119 | .026 |
| | N | 343 | 318 | 308 |
| 39 | Correlation Coefficient | .088 | 1.000 | .222(**) |
| | Sig. (2-tailed) | .119 | . | .000 |
| | N | 318 | 322 | 297 |
| 33 | Correlation Coefficient | .127(*) | .222(**) | 1.000 |
| | Sig. (2-tailed) | .026 | .000 | . |
| | N | 308 | 297 | 313 |

Group = 2

**B. Memorisation**

35. Language learning involves a lot of memorization.
61. Memorising word meanings
60. Using Lightner Box
80. Memorizing grammatical rules
85. Learning grammatical clues
23. Memorizing

**Correlations**

| | | 35 | 61 | 60 | 80 | 85 | 23 |
|---|---|---|---|---|---|---|---|
| 35 | Correlation Coefficient | 1.000 | .272(**) | .133 | .238(*) | .089 | .440(**) |
| | Sig. (2-tailed) | . | .005 | .222 | .021 | .403 | .000 |
| | N | 107 | 103 | 86 | 94 | 90 | 102 |
| 61 | Correlation Coefficient | .272(**) | 1.000 | -.116 | .289(**) | .348(**) | .372(**) |
| | Sig. (2-tailed) | .005 | . | .282 | .005 | .001 | .000 |
| | N | 103 | 106 | 88 | 95 | 91 | 102 |
| 60 | Correlation Coefficient | .133 | -.116 | 1.000 | -.033 | -.043 | .206 |
| | Sig. (2-tailed) | .222 | .282 | . | .768 | .711 | .059 |
| | N | 86 | 88 | 89 | 80 | 76 | 85 |
| 80 | Correlation Coefficient | .238(*) | .289(**) | -.033 | 1.000 | .230(*) | .328(**) |
| | Sig. (2-tailed) | .021 | .005 | .768 | . | .027 | .001 |
| | N | 94 | 95 | 80 | 97 | 92 | 92 |
| 85 | Correlation Coefficient | .089 | .348(**) | -.043 | .230(*) | 1.000 | .138 |
| | Sig. (2-tailed) | .403 | .001 | .711 | .027 | . | .200 |
| | N | 90 | 91 | 76 | 92 | 93 | 88 |
| 23 | Correlation Coefficient | .440(**) | .372(**) | .206 | .328(**) | .138 | 1.000 |
| | Sig. (2-tailed) | .000 | .000 | .059 | .001 | .200 | . |
| | N | 102 | 102 | 85 | 92 | 88 | 105 |

Group = 1

**Correlations**

| | | 35 | 61 | 60 | 80 | 85 | 23 |
|---|---|---|---|---|---|---|---|
| 35 | Correlation Coefficient | 1.000 | .211(**) | -.079 | .208(**) | .033 | .334(**) |
| | Sig. (2-tailed) | . | .000 | .203 | .000 | .575 | .000 |
| | N | 343 | 315 | 259 | 306 | 297 | 325 |
| 61 | Correlation Coefficient | .211(**) | 1.000 | -.051 | .395(**) | .142(*) | .386(**) |
| | Sig. (2-tailed) | .000 | . | .411 | .000 | .015 | .000 |
| | N | 315 | 320 | 259 | 298 | 290 | 308 |
| 60 | Correlation Coefficient | -.079 | -.051 | 1.000 | -.045 | .144(*) | .052 |
| | Sig. (2-tailed) | .203 | .411 | . | .482 | .025 | .410 |
| | N | 259 | 259 | 263 | 246 | 242 | 256 |
| 80 | Correlation Coefficient | .208(**) | .395(**) | -.045 | 1.000 | .177(**) | .311(**) |
| | Sig. (2-tailed) | .000 | .000 | .482 | . | .002 | .000 |
| | N | 306 | 298 | 246 | 310 | 296 | 299 |
| 85 | Correlation Coefficient | .033 | .142(*) | .144(*) | .177(**) | 1.000 | .045 |
| | Sig. (2-tailed) | .575 | .015 | .025 | .002 | . | .444 |
| | N | 297 | 290 | 242 | 296 | 301 | 291 |
| 23 | Correlation Coefficient | .334(**) | .386(**) | .052 | .311(**) | .045 | 1.000 |
| | Sig. (2-tailed) | .000 | .000 | .410 | .000 | .444 | . |
| | N | 325 | 308 | 256 | 299 | 291 | 329 |

Group = 2

**Vocabulary learning**

17. The most important part of learning a foreign language is learning vocabulary words.
40. Increasing my vocabulary
29. Studying vocabulary

**Correlations**

| | | 17 | 40 | 29 |
|---|---|---|---|---|
| 17 | Correlation Coefficient | 1.000 | .114 | .067 |
| | Sig. (2-tailed) | . | .254 | .493 |
| | N | 109 | 103 | 107 |
| 40 | Correlation Coefficient | .114 | 1.000 | .344(**) |
| | Sig. (2-tailed) | .254 | . | .000 |
| | N | 103 | 104 | 103 |
| 29 | Correlation Coefficient | .067 | .344(**) | 1.000 |
| | Sig. (2-tailed) | .493 | .000 | . |
| | N | 107 | 103 | 108 |

Group = 1

**Correlations**

| | | 17 | 40 | 29 |
|---|---|---|---|---|
| 17 | Correlation Coefficient | 1.000 | .095 | .248(**) |
| | Sig. (2-tailed) | . | .092 | .000 |
| | N | 343 | 317 | 327 |
| 40 | Correlation Coefficient | .095 | 1.000 | .349(**) |
| | Sig. (2-tailed) | .092 | . | .000 |
| | N | 317 | 322 | 313 |
| 29 | Correlation Coefficient | .248(**) | .349(**) | 1.000 |
| | Sig. (2-tailed) | .000 | .000 | . |
| | N | 327 | 313 | 332 |

Group = 2

**Grammar learning**

23. The most important part of learning a foreign language is learning the grammar.
41 Learning grammar
28. Studying grammar

**Correlations**

| | | 23 | 41 | 28 |
|---|---|---|---|---|
| 23 | Correlation Coefficient | 1.000 | .299(**) | .178 |
| | Sig. (2-tailed) | . | .002 | .069 |
| | N | 108 | 101 | 105 |
| 41 | Correlation Coefficient | .299(**) | 1.000 | .478(**) |
| | Sig. (2-tailed) | .002 | . | .000 |
| | N | 101 | 103 | 101 |
| 28 | Correlation Coefficient | .178 | .478(**) | 1.000 |
| | Sig. (2-tailed) | .069 | .000 | . |
| | N | 105 | 101 | 107 |

Group = 1

**Correlations**

| | | 23 | 41 | 28 |
|---|---|---|---|---|
| 23 | Correlation Coefficient | 1.000 | .188(**) | .192(**) |
| | Sig. (2-tailed) | . | .001 | .001 |
| | N | 344 | 318 | 324 |
| 41 | Correlation Coefficient | .188(**) | 1.000 | .464(**) |
| | Sig. (2-tailed) | .001 | . | .000 |
| | N | 318 | 322 | 312 |
| 28 | Correlation Coefficient | .192(**) | .464(**) | 1.000 |
| | Sig. (2-tailed) | .001 | .000 | . |
| | N | 324 | 312 | 328 |

Group = 2

## Appendix 7: Codebook for responses to the pilot study questionnaires

### Questionnaire 1

### 1.1. Background information

For Region 1, 2, and 3, I used codes 1, 2, and 3 in that order.

Item 2: For 'fields of study', I coded Mathematics 1, Natural Sciences 2, and Humanities 3.

I coded the five 'types of school' from 1 to 5. Smart School was coded 1, Nemooneh 2, Shahed 3, Nonprofit 4, and Public School 5.

For 'gender', I coded Male as 1 and Female as 2.

For 'attendance in English institutes', I marked "No" with 0 and length of attendance with the 'number of terms' attended. In data analysis, however, I divided the terms into four groups of 0, 1-10, 11-20 and 21-30.

Concerning attendance in preparation classes, I coded 'Yes' with 1 and 'No' with 0.

### 1.2. Responses to the BALLI Questionnaire

I coded items on the scale of agreement from 1-5 i.e. 'strongly agree' with 5, 'agree' with 4, 'no opinion' with 3, 'disagree' with 2 and 'strongly disagree' with 1. I also coded items 4 and 15 which had five choices of a-e from 1-5.

### 1.3. Knowledge about the test

I coded wrong response with 0 and correct response with 1.

### Questionnaire 2

I coded items 1- 51 which were on the scale of agreement in the same way I did for the BALLI items. I also coded items 54 and 61 from 1-5. In item 54 'very easy' was

coded 1 and 'very difficult' was coded 5, and in Item 61, 'without motivation' was coded 1 and 'very highly motivated' was coded 5. The coding of the ranking questions (items 52, 57, 59, 62) depended on the numbers the students assigned to each of the items to be ranked. Coding also allowed me to quantify responses to some of the open ended questions (53, 55, 56). Answering these questions required students to name sections of the SPE test. I coded grammar 1, vocabulary 2, reading 3, and cloze 4. However, a few students also named sections that did not exist on the test, in which case I coded speaking 5, pronunciation 6, writing 7, listening 8 and 'all' (all the sections or skills) 9. Sometimes, they named a combination of the sections e.g. grammar and reading, which I coded as 13, with 1 referring to grammar and 3 referring to reading.

**Questionnaire 3**

I used codes 1- 5 for Questionnaire 3 items which were based on frequency scale. I used Code 1 for 'never' and 'it's a good idea but I don't do it', 2 for 'rarely, 3 for 'sometimes', 4 for 'often, and 5 for 'always'.

Finally, I used 99 to code the missing data in all the three questionnaires.

## Appendix 8: Letters for piloting (translated into English)

**SPE:**

**Please write a letter to a friend who is going to take the SPE Test next year advising him/her on how to prepare for the test and give reasons for your advice.**

In order to prepare for the SPE Test, we should use materials from various sources for example the book called 'Vocabulary', which is useful for increasing our vocabulary, the book authored by 'Testing Organisation' and made available to students (*Bridging the Gap*) which is good for increasing vocabulary and improving reading comprehension. A student who is going to take part in the SPE exam should have a good English background. First of all the applicant should have interest, then motivation, then knowledge of and mastery over grammar and vocabulary in the 'pre-university textbook'. The applicant should also have expanded vocabulary and good comprehension ability in order to succeed.

**GE:**

**Please write a letter to a friend who is going to take the GE Test next year advising him/her on how to prepare for the test and give reasons for your advice.**

I used private preparation classes where during the week the teacher explained the points and we took notes. Weekends were spent on taking tests for which I prepared myself beforehand at home by studying the explanatory notes and the pre-university textbook followed by practice tests.

I advise those who want to prepare for the GE exam to use as many practice tests as they can. I recommend those who are weak in reading comprehension and cloze tests to use texts and close passages in previous exam papers compiled by the Testing Organisation so that they can overcome these weaknesses.

## Appendix 9: Test-retest reliability of the main study questionnaire

| | Question Items | SPE | Sig | GE | Sig |
|---|---|---|---|---|---|
| | **Learning beliefs** | | | | |
| 1 | I believe that I will learn to speak English very well. | .645** | .000 | .780** | .000 |
| 2 | It is important to speak English with an excellent pronunciation. | .759** | .000 | .754** | .000 |
| 3 | It is necessary to know about English-speaking cultures in order to speak English. | .876** | .000 | .326* | .012 |
| 4 | I enjoy practicing English with the English speakers I meet. | .685** | .000 | .691** | .000 |
| 5 | It's o.k. to guess if you don't know a word in English. | .780** | .000 | .790** | .000 |
| 6 | I have a special ability for learning foreign languages. | .900** | .000 | .707** | .000 |
| 7 | The most important part of learning a foreign language is learning vocabulary words. | .847** | .000 | .816** | .000 |
| 8 | People in my country feel that it is important to speak English. | .712** | .000 | .680** | .000 |
| 9 | I feel timid speaking English with other people. | .892** | .000 | .872** | .000 |
| 10 | The most important part of learning a foreign language is learning the grammar. | .778** | .000 | .739** | .000 |
| 11 | I would like to learn English so that I can get to know English speakers better. | .809** | .000 | .580** | .000 |
| 12 | It is important to practice with tapes and CDs. | .712** | .000 | .647** | .000 |
| 13 | Learning a foreign language is different than learning other academic subjects. | .768** | .000 | .732** | .000 |
| 14 | The most important part of learning English is learning how to translate into Farsi. | .767** | .000 | .659** | .000 |
| 15 | If I learn English very well, I will have better opportunities for a good job. | .729** | .000 | .736** | .000 |
| 16 | I want to learn to speak English well. | .779** | .000 | .660** | .000 |
| 17 | I would like to have English-speaking friends. | .838** | .000 | .817** | .000 |
| 18 | Language learning involves a lot of memorization. | .714** | .000 | .771** | .000 |
| 19 | Speaking and listening to English are more useful than reading and writing English. | .794** | .000 | .712** | .000 |
| 20 | Language learning takes a long time. | .662** | .000 | .785** | .000 |
| 21 | It is important to find as many ways as possible to use English. | .461** | .000 | .617** | .000 |
| 22. | The English language is a) very difficult b) difficult c) of medium difficulty d) easy e) very easy | .812** | .000 | .724** | .000 |
| | **Motivation, interest, purposes, and difficulty** | | | | |
| 1 | Success in the SPE Exam is important for me. | .802** | .000 | .286* | .028 |
| 2 | The SPE Exam makes me try. | .720** | .000 | .760** | .000 |
| 3 | I enjoy learning for the SPE Exam. | .831** | .000 | .771** | .000 |
| 4 | I encounter a lot of stress during the preparation. | .910** | .000 | .815** | .000 |
| 5 | I'm afraid of getting a bad mark on the SPE Exam. | .842** | .000 | .904** | .000 |
| 6 | I don't like to study for the SPE Exam. | .567** | .000 | .692** | .000 |
| 7 | What do you think is your motivation level for learning for the SPE? a) Highly motivated b) Well - motivated c) Motivated d) Slightly motivated e) Not at all motivated | .904** | .000 | .717** | .000 |
| 8 | I am preparing for the SPE Test because ..... | | | | |
| 8a | I will use English resources at university | .874** | .000 | .888** | .000 |
| 8b | I will have more and better opportunities for my job in the future | .600** | .000 | .623** | .000 |
| 8c | I want to pass the SPE Test | .972** | .000 | .865** | .000 |
| 8d | I will be able to communicate in English | .816** | .000 | .918** | .000 |
| 8e | I want to travel abroad | .706** | .000 | .731** | .000 |
| 8f | I will be able to appreciate cultural products such as films, art and literature | .788** | .000 | .883** | .000 |

| | Question Items | SPE | Sig | GE | Sig |
|---|---|---|---|---|---|
| 8g | I'm interested in English | .754** | .000 | .515** | .000 |
| 8h | I want to live in an English-speaking country | .887** | .000 | 1.000** | . |
| 8i | I will use English in many ways in future | .693** | .000 | .454** | .001 |
| 9 | I think the SPE Test is a) very difficult b) difficult c) of medium difficulty d) easy e) very easy | .917** | .000 | .806** | .000 |
| | **Reported activities** | | | | |
| 1 | I read English newspapers and magazines. | .850** | .000 | .625** | .000 |
| 2 | I use tapes or CDs to practice English. | .911** | .000 | .734** | .000 |
| 3 | I memorise word meanings. | .893** | .000 | .784** | .000 |
| 4 | I write things like diaries, notes, etc in English. | .865** | .000 | .497** | .000 |
| 5 | I practice grammar by making sentences. | .785** | .000 | .701** | .000 |
| 6 | I watch English films or programmes. | .945** | .000 | .880** | .000 |
| 7 | I translate texts into Farsi while reading. | .840** | .000 | .471** | .000 |
| 8 | I listen to English programs on the radio. | .815** | .000 | .846** | .000 |
| 9 | I use the Internet in English. | .908** | .000 | .912** | .000 |
| 10 | I learn English through computer games. | .820** | .000 | .733** | .000 |
| 11 | I memorise grammatical rules. | .806** | .000 | .585** | .000 |
| 12 | I look for people I can talk to in English. | .868** | .000 | .775** | .000 |
| 13 | I try to guess the meaning of new words from context. | .807** | .000 | .797** | .000 |
| 14 | I take mock SPE Exams. | .859** | .000 | .705** | .000 |
| 15 | I use texts to learn grammar. | .852** | .000 | .604** | .000 |
| 16 | I read English story books. | .863** | .000 | .677** | .000 |
| 17 | I pay attention to the topic of the text when I read. | .829** | .000 | .700** | .000 |
| 18 | I use new words in sentences that I make. | .905** | .000 | .559** | .000 |
| 19 | I read the text first to get a general idea and then go back to read it more carefully. | .788** | .000 | .584** | .000 |
| 20 | I review and practice what I learned. | .824** | .000 | .673** | .000 |
| 21 | I read various English texts as much as I can. | .882** | .000 | .779** | .000 |
| 22 | I practice with sample test questions. | .649** | .000 | .721** | .000 |
| 23 | I encourage myself to write or speak English even when I'm afraid of making mistakes. | .878** | .000 | .796** | .000 |
| 24 | I send and receive emails in English. | .867** | .000 | .743** | .000 |
| 25 | I make summaries of the information that I read in English. | .925** | .000 | .813** | .000 |
| 26 | I use English-to-English dictionary. | .904** | .000 | .802** | .000 |
| 27 | Are you attending a preparation class? (Yes, No) | .705** | .000 | 799** | .000 |
| 28 | Please number from one to six the sections of the SPE Test on which you spend most time in test preparation, where 1= most time and 6= least time | | | | |
| 28 | Reading | .586** | .000 | .566** | .000 |
| 12 | Cloze | .761** | .000 | .640** | .000 |
| 28 | Language Functions | .675** | .000 | .443** | .000 |
| 28 | Sentence Structure | .696** | .000 | .544** | .000 |
| 28 | Vocabulary | .779** | .000 | .838** | .000 |
| 28 | Grammar | .670** | .000 | .708** | .000 |
| 29 | Please number from one to five the following language skills on which you spend most time, where 1= most time and 5 = least time | | | | |
| 29 | Listening | .786** | .000 | .740** | .000 |
| 29 | Speaking | .804** | .000 | .829** | .000 |
| 29 | Reading | .551** | .000 | .622** | .000 |
| 9 | Grammar | .835** | .000 | .817** | .000 |
| 29 | Vocabulary | .765** | .000 | .546** | .000 |
| 30 | On which of the following subjects do you spend more time? (English, non-English subjects, I spend equal amount of time on them) | .849** | .000 | .384** | .003 |
| 31 | On which of the following types of materials do you spend | .782** | .000 | .701** | .000 |

| | Question Items | SPE | Sig | GE | Sig |
|---|---|---|---|---|---|
| | more time (Pre-university textbook, Extra materials, I spend equal amount of time on them) | | | | |
| 32 | Among extra materials, on which of the following do you spend more time (*Bridging the Gap*, Other extra materials, I spend equal amount of time on them) | .865** | .000 | .853** | .000 |
| | **Knowledge about the test** | | | | |
| 1 | A section in the SPE Test includes listening and speaking. | .963** | .000 | .615** | .000 |
| 2 | Each section of the SPE Test has the same number of questions. | .605** | .000 | .628** | .000 |
| 3 | SPE Test has six sections. | .900** | .000 | .814** | .000 |
| 4 | Vocabulary and Reading have the same weight. | .716** | .000 | .576** | .000 |
| 5 | The total number of questions in the SPE Test is 70. | 1.000** | . | .638** | .000 |
| 6 | The majority of the SPE questions come from the Pre-university book. | .917** | .000 | .486** | .000 |
| 7 | Marks allocated to English are more than non-English subjects. | .733** | .000 | .618** | .000 |
| 8 | Each question has the same mark as the other, regardless of which section it belongs to. | .782** | .000 | 547** | .000 |
| | **Average correlation of each group** | **.78** | | **.70** | |

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Appendix 10: A sample letter (by SPE 3, translated into English)

**Instruction: Please write a letter to friend who is going to take the SPE Test next year and advise him / her on how to prepare for the test.**

I strongly believe in preparation classes and in the tests given by the preparation class teacher and in institutes which conduct comprehensive mock exams (particularly those constructed by the National Organisation for Educational Testing). Since the books recommended by the Organisation are not always available and are not responsive to all the needs of the students, I believe that English institutes have a very positive effect on students' progress. You should also take tests at home provided all the conditions governing the real exam are present. It is also worth noting that it is very effective to use previous test papers one month before the exam, particularly for General Subjects.

## Appendix 11: An interview sample (with SPE 1, translated into English)

**Note:** The main types of questions in the interviews include questions asking for confirmation of responses to the questionnaire items and 'why' questions.

**1.1. Do you confirm your responses to this questionnaire item (please number from one to six the sections of the SPE Test for which you spend most time inside class and outside class where 1= most time and 6= least time)?**

Yes:

| Grammar | Vocabulary | Sentence Structure | Language Functions | Cloze Passage | Reading |
|---|---|---|---|---|---|
| 1st | 2nd | 5th | 6th | 4th | 3rd |

**1.1. Why did you rate the amount of time in the way you did?**

Out of 70 questions, about 50 questions are related to vocabulary and only 20 to grammar but at present I spend most time on grammar because grammar is my major problem. If don't know grammar I can't do anything. But in the language school which I go to they don't emphasise grammar but how much you know grammar or vocabulary is important there.

**1.2. What about vocabulary? Why did you rate it 2nd?**

It's the same for vocabulary because there are many questions on words which have various meanings that we have to learn and they may ask us the meaning of a word which we may not know.

It's the same for reading and usually we have two reading passages in the SPE exam which are heavy and difficult for us and therefore we should practice vocabulary and grammar so that we can learn it. Reading and the way you pronounce are very important in English institutes as well.

Cloze test is important because it includes both vocabulary and grammar but mainly vocabulary. For example we should know whether a preposition is suitable for a blank or a verb, and if verb what tense it should be.

And concerning Sentence Structure and Language Functions, it's not necessary to spend much time on these two sections and I only do practice tests.

**2.1. Do you do these oral activities (using tapes and CDs, watching films or TV programmes, listening in general or listening to the radio, and speaking)?**

Yes, to some extent.

**2.2. Why do you do these activities?**

I want to understand what they say so that I can use them in speaking in the future.

Generally I like to do these in order to learn English well, but they are also helpful for success in the exam.

I think watching English programs and films are very useful for reading comprehension, listening and understanding what people say. Of course the exam also has reading comprehension. Usually I also explain reading texts in English.

I use tapes and CDs which they give us in English institutes.

They are good for listening so that we know how structures and vocabulary are used in our books or elsewhere.

In class I talk to my teacher and classmates in English and at home I talk to my sister.

**2.3. So you believe that these activities are useful for structure, vocabulary, and reading in the exam as well. Is that correct?**

That's correct.

**2.4. Why did you say you do these activities 'to some extent'?**

I don't do them as much as before because I don't have enough time.

**3.1. Do you confirm your response to this questionnaire item** (on which of the following types of materials do you spend more time? Please tick the appropriate box)**?**

Yes:

| Pre-university textbook | Extra materials | I spend equal amount of time on them |
|---|---|---|
| + | ++ | |

**3.1. Why do you use extra materials more than the pre-university book?**

This is because the number of SPE questions is more than GE questions.

**3.2. Where do you get these materials from?**

I've bought some reading books which are collections of reading passages similar to those of the SPE exam, some passages are longer, some are shorter and some have different questions but in general they are similar to those of the SPE exam. We have some pamphlets from the English institute and the reading passages which we do in our preparation class. I should read those pamphlets as well because in the institute they give us exams based on those pamphlets. There are some new words and some grammar points which are useful for both the institute and the SPE exam.

**3.3. To what extent is the pre-university book important?**

The pre-university book is suitable for the GE section of the exam. The revised edition of the book has good reading passages, grammar is explained and if we do the exercises it's good. Generally the revised edition is good.

**4.1. So you are attending both a preparation class and an English institute. Correct?**

Yes, I used to go to an English institute and still go.

**5.1. Do you confirm your responses to this questionnaire item** (on which of the following subjects do you spend more time? Please tick the appropriate box)**?**

Yes:

| English | General subjects | I spend equal amount of time on them |
|---|---|---|
| ++ | + | |

**5.2. Why do you do these activities?**

English has more weight than non-English subjects. Besides, English is what we need in future. However, non-English subjects are only useful for passing the exam and raising our rank but not useful for our future.

**6.1. Do you do these written activities** (reading newspapers and magazines, reading stories, and writing diaries or daily notes)**?**

I use newspapers and magazines but not as much as tapes and CDs. Writing daily notes and diaries are useful for the SPE exam because when we translate things from Farsi into English, we can learn and review them. Concerning reading stories, sometimes we read short stories as assignments in the English institutes.

**7.1. Do you confirm your response to this questionnaire item** (I translate texts into Farsi while reading)**?**

Not quite, because I don't always translate into Farsi but only when there are parts which are difficult to understand. Sometimes, I use synonyms and explain the text in

English.

**8.1. Can you just confirm your responses to these questionnaire items** (I pay attention to the topic of the text when I read, I skim the text first to get a general idea and then go back to read it more carefully)**?**

Yes, I look at the topic and discover the meaning of the whole text by adding the initial meaning gained from the first sentence to the meaning from the final sentence. These sentences are very important. Then I read the passage more carefully and then answer the questions.

**9.1. Do you confirm your responses to this questionnaire item** (I try to guess the meaning of new words from context)**?**

Yes, because there might be a word in a sentence in the exam which we may not know and I can use my general understanding of the sentence to guess the meaning of the unfamiliar word. Therefore this enables us to guess the meaning of the word.

**10.1. Do you confirm your responses to these questionnaire items** (I memorise grammatical rules, I use texts to learn grammar)**?**

Before I study the grammar, I look to see what type of grammar it is. First I memorise the formula then I use it in examples.

We find grammatical structures in texts where many things become clear for example what the tense of the sentence is and why the sentence is used here and for difficult grammar points I make example sentences on my own so that the grammar points become concrete and doing practice tests is very important, for example when you don't know the correct answer and when you check it you can learn better.

**11.1. Do you confirm your responses to this questionnaire item** (I practice with sample test questions)

Yes

**12.1. Do you confirm your responses to this questionnaire item** (Among extra materials, on which of the following do you spend more time? Please tick the appropriate box in each row)**?**

Yes, I use both.

| *Bridging the Gap* | Other materials | I spend equal amount of time on them |
|---|---|---|
| | | √ |

**13.1. What skills are emphasised in your textbooks?**

Usually it is vocabulary, grammar, and reading.

**13.2. What about your teacher?**

The same

**13.3. Does he or she speak English in class?**

No, but there is no reason for us to converse in English, so the teacher should speak Farsi because here it is a matter of learning. In a preparation class Farsi should be used more but in a conversation class, English.

**14.1. You do oral skills for SPE and general language learning and you do grammar and vocabulary for SPE. Do you think grammar and vocabulary are necessary for language learning in general or in other words for a situation in which there is no exam?**

Yes, all skills would be useful, grammar, vocabulary, listening, speaking, etc.

**14.2. So you confirm your responses to the questionnaire items** (the most important part of learning a foreign language is learning vocabulary words, the most important part of learning a foreign language is learning the grammar, it is important to practice with tapes and CDs)**?**

Yes, they are useful for both learning English in general and for the SPE exam because anyone who is going to take the SPE exam has to use them and has to improve their English.

**15.1. Do you confirm your responses to this questionnaire item** (It's o.k. to guess if you don't know a word in English)**?**

Yes

**15.2. Why do you guess?**

Well, they may ask us a word which we don't know.

**15.3. Does this reason apply to a situation where there is no exam for example when you are reading a text for your own interest?**

Yes, there may be new words in any text.

**15.4. How do you guess?**

First I read the passage to identify the new words. It's good if I can guess the meaning of the new words, if not I look them up. Then I read the passage again and answer the questions.

Paying attention to the root of the word could be helpful. There are words which have similar noun and verb forms and there are words whose parts of speech are different, for example in one sentence it may be an adjective and in another it may be a verb. If

we read all the sentences we can discover the meaning of the text. Therefore based on the meaning of the sentence we can guess.

## Appendix 12: The English version of the main study questionnaire

**Note:** The following questionnaire was used for SPE students. However, the same questionnaire was used for GE students except that 'SPE' was replaced by 'GE' in the items, titles, etc.

### The Specialised English Test (SPE) Questionnaire

Dear Students,

I am a full-time student doing a PhD in Applied Linguistics in Lancaster University, UK. I would like you to fill in the following questionnaire for my dissertation. The purpose of my study is to explore how you learn English and how you prepare for the SPE Exam. Please answer as honestly as you can, based on how you really feel.
There are five sections in this questionnaire asking you about your background, your views about language learning, your views about the SPE Test, things you do to prepare for this test, and your knowledge of the test. As you will take the University Entrance Exam this year, your opinions will be invaluable to my study. The information you provide will help us design better tests in future. All your responses and information will be kept strictly confidential and will only be used for this research. I can share the results with you if you are interested. Finally, I would appreciate it if you signed your consent before answering the questionnaire.
Your help will be greatly appreciated.

Yours faithfully,

M. M. Abbasabadi

By completing this questionnaire, I signal that I agree to participate in this research.

.....

(Your signature)

*For researcher's use:*

| Questionnaire Code |
|---|
| |

**Section 1: Background Information**

1. Are you male or female? Please tick the appropriate box.

| Male | Female |
|---|---|
| | |

2. Please place a tick in the box indicating your type of school.

| Non-profit | Public |
|---|---|
| | |

3. Please place a tick in the box indicating your field of study at high school.

| Mathematics & Physics | Natural sciences | Humanities |
|---|---|---|
| | | |

4. Have you attended English language institutes so far? Please tick the appropriate box.

| Yes | No |
|---|---|
| | |

If your answer was "Yes", please tick one of the following boxes to indicate how many terms you have attended English language institutes.

| a) 1-5 | b) 6-10 | c) 11-15 | d) 16-20 | e) 21-25 | f) 26-30 | g) Over 30 |
|---|---|---|---|---|---|---|
| | | | | | | |

5. Are you going to take the Specialised English Test (SPE)?

| Yes | No |
|---|---|
| | |

6. How do you evaluate your English ability in general?

| Weak | Below average | Average | Above average | Good |
|---|---|---|---|---|
| | | | | |

## Section 2: My views about language learning and learning English

*Please tick the most appropriate boxes to express your views about language learning and learning English.*

| | Views | Strongly agree | Agree | No opinion | Disagree | Strongly disagree |
|---|---|---|---|---|---|---|
| 1 | I believe that I will learn to speak English very well. | | | | | |
| 2 | It is important to speak English with an excellent pronunciation. | | | | | |
| 3 | It is necessary to know about English-speaking cultures in order to speak English. | | | | | |
| 4 | I enjoy practicing English with the English speakers I meet. | | | | | |
| 5 | It's o.k. to guess if you don't know a word in English. | | | | | |
| 6 | I have a special ability for learning foreign languages. | | | | | |
| 7 | The most important part of learning a foreign language is learning vocabulary words. | | | | | |
| 8 | People in my country feel that it is important to speak English. | | | | | |
| 9 | I feel timid speaking English with other people. | | | | | |
| 10 | The most important part of learning a foreign language is learning the grammar. | | | | | |
| 11 | I would like to learn English so that I can get to know English speakers better. | | | | | |
| 12 | It is important to practice with tapes and CDs. | | | | | |
| 13 | Learning a foreign language is different than learning other academic subjects. | | | | | |
| 14 | The most important part of learning English is learning how to translate into Farsi. | | | | | |
| 15 | If I learn English very well, I will have better opportunities for a good job. | | | | | |
| 16 | I want to learn to speak English well. | | | | | |
| 17 | I would like to have English-speaking friends. | | | | | |
| 18 | Language learning involves a lot of memorization. | | | | | |
| 19 | Speaking and listening to English are more useful than reading and writing English. | | | | | |
| 20 | Language learning takes a long time. | | | | | |
| 21 | It is important to find as many ways as possible to use English. | | | | | |

22. Tick the statement that you think is most accurate.
The English language is .....

| a) very difficult | b) difficult | c) of medium difficulty | d) easy | e) very easy |
|---|---|---|---|---|
| | | | | |

## Section 3: My views about learning for the SPE Exam

*Please tick the most appropriate boxes to express your views about learning for the SPE Exam.*

| | Views | Strongly agree | Agree | No opinion | Disagree | Strongly disagree |
|---|---|---|---|---|---|---|
| 1 | Success in the SPE Exam is important for me. | | | | | |
| 2 | The SPE Exam makes me try. | | | | | |
| 3 | I enjoy learning for the SPE Exam. | | | | | |
| 4 | I encounter a lot of stress during the preparation. | | | | | |
| 5 | I'm afraid of getting a bad mark on the SPE Exam. | | | | | |
| 6 | I don't like to study for the SPE Exam. | | | | | |

7. What do you think is your motivation level for learning for the SPE? Please tick the appropriate box.

| a) Highly motivated | b) Well-motivated | c) Motivated | d) Slightly motivated | e) Not at all motivated |
|---|---|---|---|---|
| | | | | |

8. Please number from one to five the most important reasons why you are preparing for the SPE Test where 1= the most important and 5= the least important.

I am preparing for the SPE Test because.....

| | |
|---|---|
| a) I will use English resources at university | |
| b) I will have more and better opportunities for my job in the future | |
| c) I want to pass the SPE Test | |
| d) I will be able to communicate in English | |
| e) I want to travel abroad | |
| f) I will be able to appreciate cultural products such as films, art and literature | |
| g) I'm interested in English | |
| h) I want to live in an English-speaking country | |
| i) I will use English in many ways in future | |

9. Please tick the appropriate box.
I think the SPE Test is.....

| a) very difficult | b) difficult | c) of medium difficulty | d) easy | e) very easy |
|---|---|---|---|---|
| | | | | |

## Section 4: Things I do to prepare for the SPE Exam

*How often do you do the following activities? Please tick the most appropriate box.*

| | Things I do | always | often | Sometimes | rarely | never | A good idea but I don't do it |
|---|---|---|---|---|---|---|---|
| 1 | I read English newspapers and magazines. | | | | | | |
| 2 | I use tapes or CDs to practice English. | | | | | | |
| 3 | I memorise word meanings. | | | | | | |
| 4 | I write things like diaries, notes, etc in English. | | | | | | |
| 5 | I practice grammar by making sentences. | | | | | | |
| 6 | I watch English films or programmes. | | | | | | |
| 7 | I translate texts into Farsi while reading. | | | | | | |
| 8 | I listen to English programs on the radio. | | | | | | |
| 9 | I use the Internet in English. | | | | | | |
| 10 | I learn English through computer games. | | | | | | |
| 11 | I memorise grammatical rules. | | | | | | |
| 12 | I look for people I can talk to in English. | | | | | | |
| 13 | I try to guess the meaning of new words from context. | | | | | | |
| 14 | I take mock SPE Exams. | | | | | | |
| 15 | I use texts to learn grammar. | | | | | | |
| 16 | I read English story books. | | | | | | |
| 17 | I pay attention to the topic of the text when I read. | | | | | | |
| 18 | I use new words in sentences that I make. | | | | | | |
| 19 | I read the text first to get a general idea and then go back to read it more carefully. | | | | | | |
| 20 | I review and practice what I learned. | | | | | | |
| 21 | I read various English texts as much as I can. | | | | | | |
| 22 | I practice with sample test questions. | | | | | | |
| 23 | I encourage myself to write or speak English even when I'm afraid of making mistakes. | | | | | | |
| 24 | I send and receive emails in English. | | | | | | |
| 25 | I make summaries of the information that I read in English. | | | | | | |
| 26 | I use English-to-English dictionary. | | | | | | |

27. Are you attending a preparation class?

| Yes | No |
|---|---|
| | |

28. Please number from one to six the sections of the SPE Test on which you spend most time during test preparation, where 1= most time and 6= least time.

| Grammar | Vocabulary | Sentence Structure | Language Functions | Cloze Passage | Reading |
|---|---|---|---|---|---|
| | | | | | |

29. Please number from one to five the following language skills on which you spend most time during test preparation, where 1= most time and 5 = least time.

| Listening | Speaking | Reading | Grammar | Vocabulary |
|---|---|---|---|---|
| | | | | |

30. On which of the following types of materials do you spend more time? Please tick the appropriate box.

| Pre-university textbook | Extra materials | I spend equal amount of time on them |
|---|---|---|
| | | |

31. Among extra materials, on which of the following do you spend more time? Please tick the appropriate box.

| *Bridging the Gap* | Other extra materials | I spend equal amount of time on them |
|---|---|---|
| | | |

32. On which of the following subjects do you spend more time? Please tick the appropriate box.

| English | Non-English subjects | I spend equal amount of time on them |
|---|---|---|
| | | |

**Section 5: What is the SPE Test**

*Are the following statements about the SPE Test correct? Please tick the most appropriate box.*

| | Item | Yes | No | I'm not sure |
|---|---|---|---|---|
| 1 | A section in the SPE Test includes listening and speaking. | | | |
| 2 | Each section of the SPE Test has the same number of questions. | | | |
| 3 | SPE Test has six sections. | | | |
| 4 | Vocabulary and Reading have the same weight. | | | |
| 5 | The total number of questions in the SPE Test is 70. | | | |
| 6 | The majority of the SPE questions come from the Pre-university book. | | | |
| 7 | Marks allocated to English are more than non-English subjects. | | | |
| 8 | Each question has the same mark as the other, regardless of which section it belongs to. | | | |

**Thank you for your cooperation**

## Appendix 13: Extracts and instructions for checking the inter-rater reliability of the coding

### Background to my study

The aim of my study is to explore the "washback effect" of a high stakes test, the Specialised English Test (SPE), which is used to admit students to universities for BA programmes in English in Iran since its introduction in 2002. This test was developed from the General English Test (GE) which was formerly used for this purpose. However, the GE Test is still being used for some groups of students. The two tests are parts of the larger University Entrance Examination and therefore, there are other non-English subjects as well which students have to take at the same time. However, while for the SPE Group the weight of English is higher than non-English subjects, for the GE Group the weight of non-English subjects is higher. While SPE Group takes both SPE and GE tests, the GE Group does not. There are other differences between the two tests as well. The following table shows the number of items in the two tests.

**Table 1: Number of items/weight of each section in SPE and GE tests**

| Skills/Test Sections | GE | SPE |
|---|---|---|
| Vocabulary | 10 | 20 |
| Reading | 5 | 15 |
| Cloze | 5 | 15 |
| Grammar | 5 | 10 |
| Language Functions | 0 | 5 |
| Word Order | 0 | 5 |
| **Total** | **25** | **70** |

The table shows that the SPE Test has 6 sections and the GE Test has 4 sections. In other words, four sections of the GE Test have been retained, but two new sections were also added to the SPE Test. The number of items in the retained sections increased, but the increase was not equal in all the sections. Grammar items were increased less than those of other sections. Two other changes included requirement for the use of extra materials (materials in addition to school materials) for the SPE Test and rise in its difficulty level compared to the GE Test.

To investigate the washback effect of the SPE Test, I examined the learning activities of the students, both SPE test takers and GE test takers, to see what they did to prepare for the tests and why. I used three instruments to collect the data: questionnaire, letters of preparation advice to prospective test takers, and interviews.

### The task

The purpose of the task which you will see on pages 5- 8 is to check the inter-rater reliability of my coding two sets of data (letters and interviews) from the two groups of students (SPE and GE groups). Letters were concerned with what the students recommended for test preparation, and interviews were concerned with 'why the students did the activities they did' i.e. whether what they said they did was because of the test or other factors.

You will see selections from the two sets of data in tables 4- 7. I would like you to use the codes in Table 3 to code the selections. Tables 4- 7 include three columns. The first column includes the transcripts. You will take three steps to code the transcripts i.e. fill in the other two columns. First, you will underline the sections of the transcripts which you are going to code. Then, you will enter the number of the underlined text in the second column. Third, you will examine the codes in Table 3 to decide which one (ones) matches the text which you have underlined and then enter the code(s) in the same row in the third column. There could be four possibilities in underlining and numbering the texts and coding them. The following table illustrates how to underline and enter the figures:

**Table 2: Possibilities: how to underline and enter the figures**

| | Transcripts | Underlined text | Code |
|---|---|---|---|
| 1 | Kkkkkkkkkkkkkkkkkkkkkk (1) kkkkkkkk. Kkkkkkkk | (1) | 22 |
| 2 | Kkkkkkkkkkkkkkkkkkkkkk (2) kkkkkkkk. Kkkkkkkk (3) | (2)<br>(3) | 7<br>7 |
| 3 | Kkkkkkkkkkkkkkkkkkkkkk (4) kkkkkkkkkkk. Kkkkkkkk | (4)<br>(4) | 11<br>18 |
| 4 | Kkkkkkkkkkkkkkkkkkkkkk (5) (6) kkkkkkkkkkkkkkkkk | (5)<br>(6) | 32<br>19 |

In Possibility 1, one code may match one underlined text, which is probably the most common.
In possibility 2, same code matches two underlined texts.
Possibility 3 is the reverse of Possibility 2, where two different codes match one underlined part.
Possibility 4 shows that there might be texts which should be assigned more than one code but may have a structure which is inseparable, for example in 'good teacher and textbook', the adjective 'good' refers to both 'teacher' and 'textbook'. In this case, you can underline the whole text as one part but insert more than one number after the underlined part.
***However, the above possibilities are only suggestions. If you find them confusing, please feel free to do the coding in your own way.***

**Table 3: Codes and their meanings for inter-rater reliability**

| Codes | Meanings |
|---|---|
| 15 | Comparative weight or importance of test sections- vocabulary |
| 31 | Comparative weight or importance of test sections- grammar |
| 4 | Number of items/weight- spending certain amount of time on a skill because of the number of items/weight |
| 7 | Learning materials- spending certain amount of time on a skill because of learning materials |
| 8 | English background- spending certain amount of time on a skill because of English background |
| 32 | Point of time- spending certain amount of time on a skill because of the certain point of time in the preparation period |
| 28 | Predictability of questions- spending certain amount of time on a skill because of predictability of questions |
| 23 | School or teacher's emphasis/effect of school or teacher- spending certain amount of time on a skill because of school or teacher's emphasis |
| 25 | Reducing mistakes- spending certain amount of time on a skill because of desire for reducing mistakes |
| 20 | Belief- spending certain amount of time on a skill because of the belief that test sections/skills affect each other |
| 24 | Belief- spending certain amount of time on a skill because of belief about the nature of a specific skill regardless of its effect on or its relationship with other skills |
| 5 | Test demand: good command of GE |
| 10 | Test demand: comparative difficulty of/level of knowledge for SPE and GE |
| 17 | Pre-university textbook- using the pre-university textbook or recommending it |
| 6 | Extra materials- using or recommending extra materials i.e. in addition to school books or the pre-university book |
| 11 | Time management |
| 12 | Vocabulary-learning vocabulary or recommending it |
| 14 | Vocabulary-learning method- memorising words or recommending this |
| 22 | Vocabulary-learning method- using sentences or recommending this |
| 9 | Vocabulary-learning method- making sentences with words or recommending it |
| 2 | Vocabulary-learning method- using flashcards or recommending them |
| 3 | Vocabulary-learning method- learning accurately or recommending it |
| 29 | Vocabulary-learning method- using Leitner Box or recommending it |
| 39 | Vocabulary- reason for learning vocabulary |
| 41 | Vocabulary- reason for using flashcards |
| 16 | Grammar- learning grammar or recommending it |
| 36 | Grammar- learning quality/depth of learning |
| 37 | Grammar- learning towards test |
| 1 | Reading-learning reading or recommending it |
| 40 | Reading- reason for learning reading |
| 34 | Cloze- learning method- reading a cloze passage in stages |
| 35 | Cloze- learning method- determined by teacher |
| 13 | Word Order- studying 'word order' or recommending it |
| 19 | Preparation class- attending preparation classes (for practice tests) or recommending it |
| 21 | Purpose or reason for exam preparation/success |
| 26 | Mock exam- taking mock exams or recommending them |
| 33 | Mock exam- reason- for taking mock exams |
| 27 | Practice tests- using practice tests or recommending them |
| 30 | Reviewing- reviewing learned materials or recommending it |
| 38 | Reviewing- reason for reviewing |

## Selections of data to be coded

### A. Letters

**Reminder:** Instructions for letter writers were: What test preparation advice do you give to your friend who is going to take the SPE/GE Test next year?

**Table 4: Letter from SPE student # 18**

| Transcripts | Underlined text | Code |
|---|---|---|
| First, you should know more than average students i.e. you should be skilled in Specialised English | | |
| because Specialised English deals with materials outside school textbooks. | | |
| Second, you should learn vocabulary in sentences or through making sentences for example using flashcards | | |
| because flashcards provide for review of the words and for recording them in long-term memory. | | |
| In addition, this makes the students not spend their useful time but dead time on learning words. | | |
| Third, you should also know General English at a high level | | |
| because knowledge of General English is prerequisite to Specialised English. | | |
| Fourth, also use non-school resources and books because Specialised English depends on extensive resources. | | |
| Fifth, you'd better use Leitner Box for accurate learning of vocabulary. | | |

**Table 5: Letter from GE student # 47**

| Transcripts | Underlined text | Code |
|---|---|---|
| In my opinion, the most important is that you shouldn't have any weaknesses in vocabulary and be rich in vocabulary | | |
| at least in order to understand the exam questions and to translate them correctly. | | |
| For a short term, it is necessary to memorise vocabulary | | |
| so that you can understand both the question and the correct answer | | |
| because it is said that most of the questions are based on meaning | | |
| The second stage is learning grammar, | | |
| not in ways in which it is learned in high school or pre-university textbook | | |
| but in ways helpful for the test as well as deeply and meaningfully | | |
| so that you can understand it and know it by heart like our Persian grammar, | | |
| in which case you will definitely be able to answer the exam questions correctly. | | |
| However, without help, practice test books, and practice test classes, this would be a little difficult. | | |
| To become familiar with the exam and exam questions and to reduce your stress, | | |
| you can take part in mock exams. | | |
| In my opinion, the best way to succeed in the exam is if you are really interested in English. | | |

## B. Interviews

**Reminder:** The first interview question which is not included in the following tables asked the students to rate the amount of time they spent on each skill. The interview questions which are included in the tables are based on that first question. However, the questions will not be coded, but only the students' statements.

**Table 6: Interview with SPE student # 3**

| Transcripts | Underlined text | Code |
|---|---|---|
| *Why do you spend most time on Vocabulary?*<br>SPE exam is mainly grammar and vocabulary and because in school also we are asked questions on these | | |
| *Why do you spend less time on Reading?*<br>Unlike grammar which is fixed, vocabulary and reading are different. | | |
| We have to study them regularly so that the number of our mistakes is reduced. | | |
| *Why do you spend less time on Grammar?*<br>Grammar doesn't take much time because it is fixed. | | |
| There is not much difference between pre-university grammar and university level grammar. | | |
| There is a base to which, at most, a little material is added. | | |
| *Why do you spend less time on Cloze?*<br>Cloze depends on vocabulary and grammar, but more on vocabulary | | |
| i.e. words with similar meanings but different uses. Cloze also requires reading skill. | | |
| *Why do you spend less time on Word Order?*<br>We don't work on Word Order very much. | | |
| Only 1 month to the exam the preparation class teacher is going to get us to practice this section. | | |
| It deals with adjectives, etc and is not very important. | | |
| *Why do you spend less time on Language Functions?*<br>Language Functions is conversational and requires idioms. | | |
| As I've attended language schools I have no particular problem with this section. | | |

**Table 7: Interview with GE student # 3**

| Transcripts | Underlined text | Code |
|---|---|---|
| *Why do you spend most time on Vocabulary?*<br>Vocabulary is the most important for the exam because it has 10 questions. | | |
| *Why do you spend less time on Reading?*<br>There are 5 reading comprehension questions in the exam. Our teacher also emphasises reading in class. | | |
| *Why do you spend less time on Grammar?*<br>Grammar is important for the exam. It is explained briefly in the book, which is enough for me. | | |
| *Why do you spend less time on Cloze?*<br>There are 4 or 5 questions on cloze test. It is similar to reading | | |
| and I do practice tests which have both cloze tests and reading tests. | | |
| Teacher suggested studying cloze like reading | | |
| i.e. first reading for a general idea and second reading for answering the questions. | | |
| *Why do you spend no time on Word Order and Language Functions?*<br>I haven't heard of such sections in the GE Test. | | |

## Appendix 14: Inter-rater reliability of the coding

| Items | A | B | C | D | Agreement in % |
|---|---|---|---|---|---|
| **1** | 10 | 10 | 10 | 10 | 100% |
| **2** | 6 | 6 | 6 | 6 | 100% |
| **3** | 12 | 12 | 12 | 12 | 100% |
| **4** | 22 | 22 | 22 | 22 | 100% |
| **5** | 9 | 9 | | 9 | 75% |
| **6** | 2 | 2 | 2 | 2 | 100% |
| **7** | 30 | 30 | 30 | | 75% |
| **8** | 41 | 41 | 41 | 41 | 100% |
| **9** | 41 | 41 | 39 | 41 | 75% |
| **10** | 11 | 11 | 11 | 10 | 75% |
| **11** | 5 | 5 | 5 | 5 | 100% |
| **12** | 10 | 10 | 10 | 10 | 100% |
| **13** | 6 | 6 | 6 | 6 | 100% |
| **14** | 6 | 6 | 6 | 10 | 75% |
| **15** | 29 | 29 | 6 | 29 | 75% |
| **16** | 3 | 3 | 3 | 3 | 100% |
| **17** | 12 | 12 | 12 | 12 | 100% |
| **18** | 39 | 39 | 39 | 39 | 100% |
| **19** | 39 | 39 | 39 | 39 | 100% |
| **20** | 14 | 14 | 14 | 14 | 100% |
| **21** | 39 | 39 | 20 | 39 | 75% |
| **22** | 15 | 39 | 39 | 39 | 75% |
| **23** | 16 | 16 | 16 | 16 | 100% |
| **24** | 23 | 23 | 23 | 23 | 100% |
| **25** | 7 | 7 | 7 | 7 | 100% |
| **26** | 17 | 17 | 17 | 17 | 100% |
| **27** | 37 | 37 | | 37 | 75% |
| **28** | 36 | 36 | 36 | | 75% |
| **29** | 36 | 36 | 36 | 36 | 100% |
| **30** | 36 | 36 | 36 | 36 | 100% |
| **31** | 27 | 27 | 27 | 17 | 75% |
| **32** | 27 | 27 | 27 | 27 | 100% |
| **33** | 19 | 19 | 19 | | 75% |
| **34** | 33 | 33 | 33 | 33 | 100% |
| **35** | 33 | 33 | 33 | 33 | 100% |
| **36** | 26 | 26 | 26 | 26 | 100% |
| **37** | 21 | 21 | 21 | 21 | 100% |
| **38** | 31 | 31 | 31 | 31 | 100% |
| **39** | 16 | 16 | 16 | 16 | 100% |
| **40** | 15 | 15 | 15 | 15 | 100% |
| **41** | 12 | 12 | | 12 | 75% |
| **42** | 23 | 23 | 23 | 23 | 100% |
| **43** | 24 | 24 | 24 | 24 | 100% |
| **44** | 24 | 24 | 24 | 24 | 100% |
| **45** | 24 | 24 | 24 | 24 | 100% |
| **46** | 12 | 12 | 12 | 11 | 75% |
| **47** | 1 | 1 | 1 | 1 | 100% |
| **48** | 25 | 25 | 25 | 25 | 100% |
| **49** | 24 | 24 | 24 | 24 | 100% |
| **50** | 24 | 24 | 24 | 24 | 100% |
| **51** | 24 | 24 | 24 | 24 | 100% |
| **52** | 20 | 20 | 20 | 20 | 100% |

| Items | A | B | C | D | Agreement in % |
|---|---|---|---|---|---|
| **53** | 20 | 20 | 15 | 20 | 75% |
| **54** | 20 | 20 | 20 | 20 | 100% |
| **55** | 20 | 20 | 20 | 20 | 100% |
| **56** | 13 | 13 | 13 | 13 | 100% |
| **57** | 32 | 32 | | 32 | 75% |
| **58** | 23 | 23 | 23 | 23 | 100% |
| **59** | 28 | 28 | 28 | 28 | 100% |
| **60** | 20 | 20 | 20 | 20 | 100% |
| **61** | 8 | 8 | 8 | 8 | 100% |
| **62** | 15 | 15 | 15 | 39 | 75% |
| **63** | 12 | 12 | 12 | 12 | 100% |
| **64** | 4 | 4 | 4 | 4 | 100% |
| **65** | 4 | 4 | 4 | 4 | 100% |
| **66** | 23 | 23 | 23 | 23 | 100% |
| **67** | 16 | 16 | 16 | 16 | 100% |
| **68** | 7 | 7 | 7 | 7 | 100% |
| **69** | 17 | 17 | 17 | 17 | 100% |
| **70** | 4 | 4 | 4 | 4 | 100% |
| **71** | 20 | 20 | 20 | 20 | 100% |
| **72** | 27 | 27 | 37 | 27 | 75% |
| **73** | 35 | 35 | 23 | 35 | 75% |
| **74** | 20 | 20 | 20 | 20 | 100% |
| **75** | 34 | 34 | 34 | 34 | 100% |
| **76** | 4 | 10 | 10 | 10 | 75% |
| **Average** | | | | | **93%** |

Rows/ items = 76
56 (out of 76) × 100% = 56
20 (out of 76) × 75% = 15
Total: 56 + 15 = 71
**Average agreement:** 71 ÷ 76 = 93%

## Appendix 15: Consistency of the students' responses across different instruments

_ please see the notes at the end of this document.

**Table 1: Spending time on vocabulary**

| St | SPE | | | | | GE | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Q | Let | List | Int | Cons | Q | Let | List | Int | Cons |
| 1 | 2nd | + | √ | 2nd | 100% | 1st | + | √ | 1st | 100% |
| 2 | 1st | + | √ | 1st | 100% | 1st | + | √ | 1st | 100% |
| 3 | 1st | - | √ | 1st | .75% | 1st | - | √ | 1st | .75% |
| 4 | 1st | + | √ | 1st | 100% | 1st | + | √ | 1st | 100% |
| 5 | 2nd | + | √ | 1st | 100% | 1st | + | √ | 1st | 100% |
| 6 | 1st | + | √ | 1st | 100% | 1st | + | √ | 1st | 100% |
| 7 | 1st | + | √ | 1st | 100% | 1st | + | √ | 1st | 100% |
| 8 | 1st | - | √ | 1st | .75% | 1st | + | √ | 1st | 100% |
| 9 | 1st | - | √ | 1st | .75% | 1st | + | √ | 1st | 100% |

**Table 2: Spending time on grammar**

| St | SPE | | | | | GE | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Q | Let | List | Int | Cons | Q | Let | List | Int | Cons |
| 1 | 1st | + | √ | 1st | 100% | 4th | + | √ | 3rd | 100% |
| 2 | 2nd | + | - | 3rd | .75% | 4th | + | - | 3rd | .75% |
| 3 | 2nd | - | √ | 3rd | .75% | 3rd | - | √ | 3rd | .75% |
| 4 | 3rd | + | - | 3rd | .75% | 2nd | + | √ | 2nd | 100% |
| 5 | 1st | + | √ | 2nd | 100% | 3rd | + | √ | 3rd | 100% |
| 6 | 3rd | + | √ | 4th | 100% | 2nd | + | √ | 2nd | 100% |
| 7 | 2nd | + | √ | 2nd | 100% | 3rd | + | √ | 2nd | 100% |
| 8 | 2nd | - | √ | 2nd | .75% | 2nd | + | √ | 2nd | 100% |
| 9 | 2nd | - | √ | 2nd | .75% | 2nd | + | √ | 2nd | 100% |

**Table 3: Spending time on reading**

| St | SPE | | | | | GE | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Q | Let | List | Int | Cons | Q | Let | List | Int | Cons |
| 1 | 3rd | + | √ | 3rd | 100% | 2nd | + | - | 2nd | .75% |
| 2 | 3rd | - | √ | 2nd | .75% | 2nd | + | √ | 2nd | 100% |
| 3 | 3rd | - | - | 2nd | .50% | 2nd | + | √ | 2nd | 100% |
| 4 | 2nd | + | √ | 2nd | 100% | 3rd | - | √ | 3rd | .75% |
| 5 | 3rd | + | √ | 3rd | 100% | 2nd | - | - | 2nd | .50% |
| 6 | 2nd | - | √ | 2nd | .75% | 3rd | - | √ | 3rd | .75% |
| 7 | 3rd | - | √ | 3rd | .75% | 2nd | - | √ | 3rd | .75% |
| 8 | 3rd | - | - | 3rd | .50% | 3rd | + | √ | 3rd | 100% |
| 9 | 3rd | - | √ | 3rd | .75% | 3rd | - | √ | 3rd | .75% |

**Table 4: Spending time on cloze**

| St | SPE | | | | | GE | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Q | Let | List | Int | Cons | Q | Let | List | Int | Cons |
| 1 | 5th | - | √ | 4th | .75% | 3rd | - | - | 4th | .50% |
| 2 | 4th | - | √ | 4th | .75% | 3rd | - | - | 4th | .50% |
| 3 | 4th | - | - | 4th | .50% | 4th | + | √ | 4th | 100% |
| 4 | 4th | - | √ | 4th | .75% | 4th | - | √ | 4th | .75% |
| 5 | 4th | - | √ | 4th | .75% | 4th | - | - | 4th | .50% |
| 6 | 4th | - | √ | 3rd | .75% | 4th | - | √ | 4th | .75% |
| 7 | 4th | - | - | 4th | .50% | 4th | - | - | 4th | .50% |
| 8 | 4th | - | √ | 4th | .75% | 4th | - | √ | 4th | .75% |
| 9 | 4th | - | √ | 4th | .75% | 4th | - | √ | 4th | .75% |

**Table 5: Spending time on Language Functions**

| St | SPE | | | | | GE | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Q | Let | List | Int | Cons | Q | Let | List | Int | Cons |
| 1 | 4th | - | √ | 5th/6th | .75% | NR | - | - | - | .75% |
| 2 | 5th | - | √ | 5th/6th | .75% | NR | - | - | - | .75% |
| 3 | 6th | - | √ | 7th | .75% | NR | - | - | - | .75% |
| 4 | 6th | - | √ | 5th/6th | .75% | NR | - | - | - | .75% |
| 5 | 6th | - | √ | 5th/6th | .75% | NR | - | - | - | .75% |
| 6 | 5th | - | √ | 5th/6th | .75% | 6th | - | - | - | .75% |
| 7 | 6th | - | √ | 5th | .75% | NR | - | - | - | .75% |
| 8 | 6th | - | √ | 7th | .75% | NR | - | - | - | .75% |
| 9 | 5th | - | √ | 7th | .75% | NR | - | - | - | .75% |

**Table 6: Spending time on Sentence Structure**

| St | SPE | | | | | GE | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Q | Let | List | Int | Cons | Q | Let | List | Int | Cons |
| 1 | 6th | - | √ | 5th/6th | .75% | NR | - | - | - | .75% |
| 2 | 6th | - | √ | 5th/6th | .75% | NR | - | - | - | .75% |
| 3 | 5th | - | √ | 7th | .75% | NR | - | - | - | .75% |
| 4 | 5th | - | √ | 5th/6th | .75% | NR | - | - | - | .75% |
| 5 | 5th | - | √ | 5th/6th | .75% | NR | - | - | - | .75% |
| 6 | 6th | - | √ | 5th/6th | .75% | NR | - | - | - | .75% |
| 7 | 5th | - | √ | 6th | .75% | NR | - | - | - | .75% |
| 8 | 5th | - | √ | 7th | .75% | NR | - | - | - | .75% |
| 9 | 6th | - | √ | 7th | .75% | NR | - | - | - | .75% |

**Table 7: Spending time on listening**

| St | SPE | | | | | GE | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Q | Let | List | Int | Cons | Q | Let | List | Int | Cons |
| 1 | 5th | - | √ | + | .75% | 4th | - | - | + | .50% |
| 2 | 4th | - | √ | + | .75% | NR | - | - | - | .75% |
| 3 | 5th | - | - | - | .75% | 4th | - | - | - | .75% |
| 4 | NR | - | √ | + | .50% | 5th | - | - | - | .75% |
| 5 | 5th | - | - | - | .75% | NR | - | - | - | .75% |
| 6 | 5th | - | - | - | .75% | NR | - | - | + | .50% |
| 7 | NR | - | - | - | .75% | 5th | - | - | - | .75% |
| 8 | NR | - | - | - | .75% | 4th | - | - | - | .75% |
| 9 | NR | - | - | - | .75% | 3th | - | - | - | .75% |

**Table 8: Spending time on speaking**

| St | SPE | | | | | GE | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Q | Let | List | Int | Cons | Q | Let | List | Int | Cons |
| 1 | 4th | - | √ | + | .75% | 4th | + | √ | + | 100% |
| 2 | 3th | - | √ | + | .75% | 5th | - | √ | - | .50% |
| 3 | 4th | - | - | - | .75% | NR | - | √ | - | .50% |
| 4 | NR | - | √ | + | .50% | 4th | - | √ | - | .50% |
| 5 | 5th | - | - | - | .75% | NR | - | - | - | .75% |
| 6 | 4th | - | √ | - | .50% | NR | - | √ | + | .50% |
| 7 | NR | - | - | - | .75% | 4th | - | - | - | .75% |
| 8 | NR | - | - | - | .75% | 3th | - | √ | - | .50% |
| 9 | NR | - | - | - | .75% | NR | - | - | - | .75% |

**Table 9: Using tapes and CDs**

| St | SPE | | | | | GE | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Q | Let | List | Int | Cons | Q | Let | List | Int | Cons |
| 1 | Oft | - | √ | + | .75% | Som | - | √ | + | .75% |
| 2 | Oft | - | √ | + | .75% | Rar | + | √ | - | .75% |
| 3 | Som | - | - | - | .75% | Som | - | - | - | .75% |
| 4 | Som | - | √ | + | .75% | Rar | - | - | - | .75% |
| 5 | Rar | - | - | - | .75% | Nev | - | - | - | .75% |
| 6 | Som | - | - | - | .75% | Rar | - | - | + | .50% |
| 7 | Nev | - | - | - | 100% | Nev | - | - | - | 100% |
| 8 | Rar | - | - | - | .75% | Gd | - | - | - | 100% |
| 9 | Rar | - | - | - | .75% | Nev | - | - | - | 100% |

**Table 10: Watching films**

| St | SPE | | | | | GE | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Q | Let | List | Int | Cons | Q | Let | List | Int | Cons |
| 1 | Som | - | - | + | .50% | Alw | + | √ | + | 100% |
| 2 | Som | + | √ | + | 100% | Som | - | - | - | .75% |
| 3 | Rar | - | - | - | .75% | Gd | - | - | - | 100% |
| 4 | Som | + | √ | + | 100% | Gd | - | - | - | 100% |
| 5 | Som | - | √ | - | .50% | Nev | - | - | - | 100% |
| 6 | Gd | - | - | - | 100% | Som | - | √ | + | .75% |
| 7 | Rar | - | - | - | .75% | Nev | - | - | - | 100% |
| 8 | Rar | - | - | - | .75% | Rar | - | - | - | .75% |
| 9 | Gd | - | - | - | 100% | Nev | - | - | - | 100% |

**Table 11: Listening to the radio**

| St | SPE | | | | | GE | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Q | Let | List | Int | Cons | Q | Let | List | Int | Cons |
| 1 | Som | - | - | + | .50% | Alw | + | √ | + | 100% |
| 2 | Som | - | - | + | .50% | Rar | - | √ | - | .50% |
| 3 | Nev | - | - | - | 100% | Oft | - | - | - | .75% |
| 4 | Gd | - | - | + | .75% | Gd | - | - | - | 100% |
| 5 | Rar | - | - | - | .75% | Nev | - | - | - | 100% |
| 6 | Gd | - | - | - | 100% | Rar | - | - | + | .50% |
| 7 | Nev | - | - | - | 100% | Gd | - | - | - | 100% |
| 8 | Gd | - | - | - | 100% | Gd | - | - | - | 100% |
| 9 | Nev | - | - | - | 100% | Nev | - | - | - | 100% |

**Table 12: Reading newspapers and magazines**

| St | SPE | | | | | GE | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Q | Let | List | Int | Cons | Q | Let | List | Int | Cons |
| 1 | Rar | - | √ | + | .75% | Som | + | √ | - | .75% |
| 2 | Som | - | - | + | .50% | Rar | - | √ | - | .50% |
| 3 | Gd | - | - | - | 100% | Gd | - | - | - | 100% |
| 4 | Rar | - | √ | + | .75% | Gd | - | - | - | 100% |
| 5 | Rar | - | √ | - | .50% | Gd | - | - | - | 100% |
| 6 | Rar | - | - | - | .75% | Rar | - | - | - | .75% |
| 7 | Nev | - | - | - | 100% | Gd | - | - | - | 100% |
| 8 | Rar | - | - | - | .75% | Gd | - | - | - | 100% |
| 9 | Nev | - | - | - | 100% | Nev | - | - | - | 100% |

**Table 13: Reading stories**

| St | SPE | | | | | GE | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Q | Let | List | Int | Cons | Q | Let | List | Int | Cons |
| 1 | Oft | - | - | + | .50% | Som | - | √ | - | .50% |
| 2 | Som | - | √ | + | .75% | NR | - | √ | - | .50% |
| 3 | Nev | - | - | - | 100% | Gd | - | - | - | 100% |
| 4 | Oft | - | √ | + | .75% | Gd | - | - | - | 100% |
| 5 | Gd | - | - | - | 100% | Nev | - | - | - | 100% |
| 6 | Rar | - | √ | - | .50% | Rar | - | - | - | .75% |
| 7 | Rar | - | - | - | .75% | Nev | - | - | - | 100% |
| 8 | Nev | - | - | - | 100% | Gd | - | - | - | 100% |
| 9 | Nev | - | - | - | 100% | Nev | - | - | - | 100% |

**Table 14: Writing diaries**

| St | SPE | | | | | GE | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Q | Let | List | Int | Cons | Q | Let | List | Int | Cons |
| 1 | Som | - | √ | + | .75% | Rar | - | - | - | .75% |
| 2 | Som | - | √ | + | .75% | Som | - | √ | - | .50% |
| 3 | Gd | - | - | - | 100% | Gd | - | - | - | 100% |
| 4 | Som | - | - | + | .50% | Gd | - | - | - | 100% |
| 5 | Nev | - | √ | - | .75% | Nev | - | √ | - | .75% |
| 6 | Gd | - | - | - | 100% | Rar | - | - | - | .75% |
| 7 | Rar | - | - | - | .75% | Nev | - | - | - | 100% |
| 8 | Nev | - | - | - | 100% | Gd | - | - | - | 100% |
| 9 | Nev | - | - | - | 100% | Nev | - | - | - | 100% |

**Table 15: Translating**

| St | SPE | | | | | GE | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Q | Let | List | Int | Cons | Q | Let | List | Int | Cons |
| 1 | Oft | - | - | + | .50% | Nev | - | - | - | 100% |
| 2 | Rar | - | - | - | .75% | Som | - | - | + | .50% |
| 3 | Oft | - | - | + | .50% | Oft | - | - | + | .50% |
| 4 | Rar | - | - | + | .50% | Som | - | - | + | .50% |
| 5 | Alw | + | √ | + | 100% | Alw | - | - | + | .50% |
| 6 | Alw | - | √ | + | .75% | Oft | - | √ | + | .75% |
| 7 | Alw | - | - | + | .50% | Oft | - | √ | + | .75% |
| 8 | Alw | - | √ | + | .75% | Alw | - | √ | + | .75% |
| 9 | Alw | - | √ | + | .75% | Rar | - | √ | + | .75% |

**Table 16: Paying attention to the topic of passage while reading**

| St | SPE | | | | | GE | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Q | Let | List | Int | Cons | Q | Let | List | Int | Cons |
| 1 | Oft | - | √ | + | .75% | Oft | - | √ | + | .75% |
| 2 | Alw | - | - | + | .50% | Rar | - | - | + | .50% |
| 3 | Alw | - | √ | + | .75% | Alw | - | - | + | .50% |
| 4 | Alw | - | - | + | .50% | Oft | - | √ | + | .75% |
| 5 | Alw | - | √ | + | .75% | Som | - | √ | + | .75% |
| 6 | Alw | - | - | + | .50% | Oft | - | √ | + | .75% |
| 7 | Alw | - | √ | + | .75% | Gd | - | - | + | .75% |
| 8 | Oft | - | √ | + | .75% | Oft | - | √ | + | .75% |
| 9 | Som | - | √ | + | .75% | Som | - | - | + | .50% |

**Table 17: Reading for general and deeper understanding**

| St | SPE | | | | | GE | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Q | Let | List | Int | Cons | Q | Let | List | Int | Cons |
| 1 | Alw | - | √ | + | .75% | Rar | - | √ | + | .75% |
| 2 | Oft | - | √ | + | .75% | Oft | - | √ | + | .75% |
| 3 | Alw | - | - | + | .50% | Oft | - | √ | + | .75% |
| 4 | Oft | - | √ | + | .75% | Gd | - | - | + | .75% |
| 5 | Alw | - | √ | + | .75% | Alw | - | √ | + | .75% |
| 6 | Oft | - | - | + | .50% | Som | - | √ | + | .75% |
| 7 | Alw | - | - | + | .50% | Gd | - | - | + | .75% |
| 8 | Alw | - | - | + | .50% | Som | - | - | + | .75% |
| 9 | Som | - | √ | + | .75% | Nev | - | - | - | 100% |

**Table 18: Guessing the meaning of new words**

| St | SPE | | | | | GE | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Q | Let | List | Int | Cons | Q | Let | List | Int | Cons |
| 1 | Oft | - | √ | + | .75% | Alw | - | √ | + | .75% |
| 2 | Alw | - | √ | + | .75% | Som | + | √ | + | 100% |
| 3 | Alw | - | √ | + | .75% | Som | - | √ | + | .75% |
| 4 | Oft | - | √ | + | .75% | Som | - | √ | + | .75% |
| 5 | Alw | - | √ | + | .75% | Som | - | √ | + | .75% |
| 6 | Oft | - | √ | + | .75% | Alw | - | √ | + | .75% |
| 7 | Som | - | √ | + | .75% | Som | - | √ | + | .75% |
| 8 | Oft | - | √ | + | .75% | Rar | - | √ | - | .50% |
| 9 | NR | - | √ | - | .50% | Rar | - | √ | - | .50% |

**Table 19: Memorising grammar ruels**

| St | SPE | | | | | GE | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Q | Let | List | Int | Cons | Q | Let | List | Int | Cons |
| 1 | Som | - | - | + | .50% | Nev | - | - | - | 100% |
| 2 | Alw | - | - | + | .50% | Oft | - | - | + | .50% |
| 3 | Som | - | - | + | .50% | Som | - | √ | + | .75% |
| 4 | Alw | - | √ | + | .75% | Som | - | √ | + | .75% |
| 5 | Alw | - | - | + | .50% | Som | - | √ | + | .75% |
| 6 | Nev | - | √ | + | .75% | Oft | - | √ | + | .75% |
| 7 | Alw | - | √ | + | .75% | Gd | - | - | + | .75% |
| 8 | Oft | - | - | + | .50% | Oft | - | √ | + | .75% |
| 9 | Alw | - | - | + | .50% | Rar | - | √ | + | .75% |

**Table 20: Using texts for grammar learning**

| St | SPE | | | | | GE | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Q | Let | List | Int | Cons | Q | Let | List | Int | Cons |
| 1 | Alw | + | √ | + | 100% | Nev | - | √ | + | .50% |
| 2 | Alw | - | √ | + | .75% | Som | - | √ | + | .75% |
| 3 | Rar | - | √ | + | .75% | Oft | - | √ | + | .75% |
| 4 | Alw | - | √ | + | .75% | Oft | - | √ | + | .75% |
| 5 | Alw | - | - | + | .50% | Oft | - | √ | + | .75% |
| 6 | Alw | - | √ | + | .75% | Oft | - | √ | + | .75% |
| 7 | Alw | - | √ | + | .75% | Nev | - | - | + | .75% |
| 8 | Alw | - | √ | + | .75% | Oft | + | √ | + | 100% |
| 9 | Oft | - | √ | + | .75% | Nev | - | √ | + | .50% |

**Table 21: Using practice tests**

| St | SPE | | | | | GE | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Q | Let | List | Int | Cons | Q | Let | List | Int | Cons |
| 1 | NR | - | √ | + | .50% | Rar | - | - | + | .50% |
| 2 | Oft | - | - | + | .50% | Alw | + | - | + | .75% |
| 3 | Som | + | √ | + | 100% | Oft | + | √ | + | 100% |
| 4 | Alw | + | √ | + | 100% | Gd | + | √ | + | 100% |
| 5 | Oft | - | √ | + | .75% | Oft | + | √ | + | 100% |
| 6 | Oft | - | √ | + | .75% | Alw | + | √ | + | 100% |
| 7 | Alw | - | √ | + | .75% | Som | - | - | + | .50% |
| 8 | Alw | - | √ | + | .75% | Som | - | √ | + | .75% |
| 9 | Oft | - | - | + | .50% | Rar | - | - | + | .50% |

**Table 22: Using extra materials VS the Pre-university book (SPE Group)**

| St | Q | | Let | | List | | Int | | Cons |
|---|---|---|---|---|---|---|---|---|---|
| | Ext | Pre | Ext | Pre | Ext | Pre | Ext | Pre | |
| 1 | ++ | + | - | - | √ | √ | ++ | + | .75% |
| 2 | ++ | + | - | - | - | - | ++ | + | .50% |
| 3 | ++ | + | - | - | √ | √ | ++ | + | .75% |
| 4 | ++ | + | + | - | √ | - | ++ | + | .75% |
| 5 | ++ | + | - | - | √ | - | ++ | + | .75% |
| 6 | ++ | + | + | - | √ | √ | ++ | + | 87.5% |
| 7 | + | ++ | + | + | - | √ | + | + | 87.5% |
| 8 | + | ++ | - | - | √ | √ | + | + | .75% |
| 9 | + | + | - | - | √ | √ | + | + | .75% |

**Table 23: Using extra materials VS the Pre-university book (GE Group)**

| St | Q | | Let | | List | | Int | | Cons |
|---|---|---|---|---|---|---|---|---|---|
| | Ext | Pre | Ext | Pre | Ext | Pre | Ext | Pre | |
| 1 | ++ | + | - | - | √ | √ | + | + | .75% |
| 2 | + | ++ | + | - | - | √ | - | + | .62.5% |
| 3 | + | ++ | - | + | √ | √ | - | + | .75% |
| 4 | + | ++ | - | - | √ | √ | - | + | .62.5% |
| 5 | + | ++ | - | - | - | √ | - | + | 50% |
| 6 | + | + | - | - | √ | √ | - | + | .62.5% |
| 7 | + | + | - | - | - | - | - | + | 37.5% |
| 8 | + | + | - | + | - | √ | - | + | .62.5% |
| 9 | + | ++ | - | - | - | √ | - | + | .50% |

**Table 24: Using *Bridging the Gap* (VS other extra materials)**

| St | SPE | | | | | | GE | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Q | | Let | List | Int | Cons | Q | | Let | List | Int | Cons |
| | BrG | Ext | | | | | BrG | Ext | | | | |
| 1 | + | + | - | √ | + | .75% | - | - | - | - | - | 100% |
| 2 | + | ++ | - | √ | + | .75% | - | - | - | - | - | 100% |
| 3 | + | ++ | - | √ | - | .50% | - | - | - | - | - | 100% |
| 4 | + | + | - | - | + | .50% | - | - | - | - | - | 100% |
| 5 | + | ++ | - | √ | - | .50% | - | - | - | - | - | 100% |
| 6 | + | ++ | - | √ | + | .75% | - | - | - | - | - | 100% |
| 7 | + | ++ | - | √ | + | .75% | - | - | - | - | - | 100% |
| 8 | + | + | - | √ | + | .75% | - | - | - | - | - | 100% |
| 9 | + | + | - | √ | + | .75% | - | - | - | - | - | 100% |

**Table 25: Studying non-English subjects (VS English)**

| St | SPE | | | | | | GE | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Q | | Let | List | Int | Cons | Q | | Let | List | Int | Cons |
| | N | E | | | | | N | E | | | | |
| 1 | + | ++ | - | - | + | .50% | ++ | + | - | - | ++ | .50% |
| 2 | + | + | - | √ | = | .75% | ++ | + | - | √ | ++ | .75% |
| 3 | + | ++ | - | √ | ++ | .75% | ++ | + | - | - | ++ | .50% |
| 4 | + | + | + | √ | + | 100% | ++ | + | - | - | ++ | .50% |
| 5 | + | + | - | √ | + | .75% | ++ | + | - | - | ++ | .50% |
| 6 | ++ | + | - | √ | + | .75% | ++ | + | - | - | ++ | .50% |
| 7 | + | + | - | √ | = | .75% | ++ | + | - | - | ++ | .50% |
| 8 | + | ++ | - | √ | + | .75% | ++ | + | - | √ | ++ | .75% |
| 9 | + | ++ | - | √ | + | .75% | ++ | + | - | √ | ++ | .75% |

**Table 26: Attending preparation class**

| St | SPE | | | | | GE | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Q | Let | List | Int | Cons | Q | Let | List | Int | Cons |
| 1 | + | - | √ | + | .75% | + | - | - | - | .75% |
| 2 | - | - | - | - | 100% | NR | - | - | - | .75% |
| 3 | + | + | √ | + | 100% | + | + | √ | + | 100% |
| 4 | + | + | - | + | .75% | - | - | √ | - | .75% |
| 5 | + | - | √ | + | .75% | - | - | √ | - | .75% |
| 6 | + | - | √ | + | .75% | - | - | - | - | 100% |
| 7 | + | - | √ | + | .75% | + | - | - | + | .50% |
| 8 | + | + | √ | + | 100% | - | - | √ | - | .75% |
| 9 | + | + | - | + | .75% | - | - | √ | - | .75% |

**Table 27: Attending English institutes**

| St | SPE | | | | | GE | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Q | Let | List | Int | Cons | Q | Let | List | Int | Cons |
| 1 | 21-25 | - | - | + | .50% | 21-25 | - | - | - | .75% |
| 2 | 21-25 | - | - | - | .75% | 11-15 | - | √ | - | .50% |
| 3 | 21-25 | - | - | - | .75% | 6-10 | - | - | - | .75% |
| 4 | 21-25 | - | √ | - | .50% | 6-10 | - | - | - | .75% |
| 5 | 16-20 | - | - | - | .75% | 0 | - | √ | - | .75% |
| 6 | 16-20 | - | - | - | .75% | 0 | - | - | - | 100% |
| 7 | 6-10 | - | - | - | .75% | 0 | - | - | - | 100% |
| 8 | 6-10 | - | - | - | .75% | 0 | - | - | - | 100% |
| 9 | 0 | - | - | - | 100% | 0 | - | - | - | 100% |

**Table 28: Consistency of students' responses on questionnaire and interview in terms of belief about guessing**

| St | SPE | | | GE | | |
|---|---|---|---|---|---|---|
| | Q | Int | Cons | Q | Int | Cons |
| 1 | Str agr | + | 100% | Str agr | + | 100% |
| 2 | Agr | + | 100% | Agr | + | 100% |
| 3 | Str agr | + | 100% | Agr | + | 100% |
| 4 | Agr | + | 100% | Agr | + - | 50% |
| 5 | Str agr | + | 100% | Agr | + | 100% |
| 6 | Agr | + | 100% | Agr | + | 100% |
| 7 | Agr | + | 100% | Agr | - | 0% |
| 8 | Agr | + - | 50% | Agr | + | 100% |
| 9 | No op | + | 0% | Str agr | + | 100% |

**Table 29: Consistency of students' responses on the questionnaire and interviews in terms of belief about vocabulary learning**

| St | SPE | | | GE | | |
|---|---|---|---|---|---|---|
| | Q | Int | Cons | Q | Int | Cons |
| 1 | Agr | + | 100% | Agr | + | 100% |
| 2 | Agr | + | 100% | Agr | + | 100% |
| 3 | No op | + | 0% | Agr | + | 100% |
| 4 | Disagr | + | 0% | Agr | + | 100% |
| 5 | Str agr | + | 100% | Str agr | + | 100% |
| 6 | Agr | + | 100% | Disagr | + | 0% |
| 7 | Str agr | + | 100% | Str agr | + | 100% |
| 8 | Agr | + | 100% | Agr | + | 100% |
| 9 | Agr | + | 100% | Str agr | + | 100% |

**Table 30: Consistency of students' responses on the questionnaire and interviews in terms of belief about grammar learning**

| St | SPE | | | GE | | |
|---|---|---|---|---|---|---|
| | Q | Int | Cons | Q | Int | Cons |
| 1 | Agr | + | 100% | Str dis | - | 100% |
| 2 | No op | - | 0% | No op | + | 0% |
| 3 | No op | + | 0% | Agr | + | 100% |
| 4 | Disagr | - | 100% | Agr | + | 100% |
| 5 | Agr | + | 100% | Disagr | - | 100% |
| 6 | Disagr | + | 0% | No op | + | 0% |
| 7 | Disagr | + | 0% | Str dis | + | 100% |
| 8 | Disagr | + | 0% | No op | + | 0% |
| 9 | No op | + | 0% | Agr | + | 100% |

**Table 31: Consistency of students' responses on the questionnaire and interviews in terms of belief about learning by tapes and CDs**

| St | SPE | | | GE | | |
|---|---|---|---|---|---|---|
| | Q | Int | Cons | Q | Int | Cons |
| 1 | Str agr | + | 100% | Str agr | + | 100% |
| 2 | Agr | + | 100% | Str agr | + | 100% |
| 3 | Str agr | + | 100% | Str agr | + | 100% |
| 4 | Agr | + | 100% | Agr | + | 100% |
| 5 | Str agr | + | 100% | Str agr | + | 100% |
| 6 | Str agr | + | 100% | Agr | + | 100% |
| 7 | Agr | + | 100% | Agr | + | 100% |
| 8 | Str agr | + | 100% | Agr | + | 100% |
| 9 | Agr | + | 100% | Agr | + | 100% |

**Notes:**

1. Average consistency of SPE responses:
   $69 \times 100\% = 69$
   $144 \times 75\% = 108$
   $46 \times 50\% = 23$
   $9 \times 0\% = 0$
   $2 \times 87.5\% = 1.75$
   $69 + 108 + 23 + 1.75 = 201.75 \div 279 = .72$ (consistency)
2. Average consistency of GE responses:
   $106 \times 100\% = 106$
   $115 \times 75\% = 86.25$
   $44 \times 50\% = 22$

5 × 0%= 0
4 × 62.5%= 2.5
1 × 37.5%= .375
106 + 86.25 + 22 + 2.5 + .375 = 217.125 ÷ 279 = .77 (consistency)

3. Abbreviations: St= student, Q= questionnaire, Let= letter, Int= interview, Cons= consistency, Gd= it's a good idea but I don't do it, Nev= never, Rar= rarely, Som= sometimes, Oft= often, Alw= always, Agr= agree, No op= no opinion, Disagr= disagree, Str agr= strongly agree, Str dis= strongly disagree, NR= no response, Pre= the Pre-university textbook, Ext= extra materials, BrG= the book, *Bridging the Gap*, N= non-English subjects, E= English.
4. + in 'Let' columns= activity recommended to be done
5. - in 'Let' columns= not mentioned
6. + in 'Int' columns= activity which was done
7. - in 'Int' columns= activity not done
8. + in Tables 28- 31= believing in guessing, vocabulary and grammar learning, and using tapes and CDs
9. - in Tables 28- 31= not believing in guessing, vocabulary and grammar learning, and using tapes and CDs
10. +/- in Tables 28- 31= 'partly' believing in guessing, vocabulary or grammar learning, or using tapes and CDs
11. ++ in 'Q' and 'Int' columns in Tables 22- 25= More time spent on, + = Less time spent on, and = indicates equal amount of time
12. √ in 'List' columns= activity considered important
13. - in 'List' columns= activity not considered important
14. In Tables 1- 6, the differences between the ranks in 'Q' and 'Int' columns were minimal i.e. ranging from 1- 2 and were considered consistent. As concerns the comparison between the rankings in these two columns with the 'Let' and 'List' columns, as long as an activity was ranked (no matter 1 or 6), it was considered an activity reported to be done and was considered consistent with + and √. This was because in Tables 1-6, there were two columns in which the activities were rank ordered and two columns in which the activities were considered 'impotant' or 'not important', or 'recommended' or 'not recommended'. Therefore, they were not really comparable with the rank ordered activities.
15. Similarly in Tables 9- 21, the adverbs of frequency, 'rarely', 'sometimes', 'often', and 'always' were considered consistent with + and √, and 'never' and 'it's a good idea but I don't do it' were considered consistent with – which meant that the activity was not done or was not considered important.
16. In Tables 22- 25, the differences such as 'more', 'less', or 'equal amount of time' were ignored and what was considered was whether the activity was done or not. The reason was because the letters and the list data were not based on comparison. Therefore, all the three signs of ++ (more), + (less), and = (equal) indicate that the activities were done.
17. As the above explanations indicate, the consistency figures in Notes 1 and 2 are rough estimations.