

MAKING WASHBACK WORK

Lee King Siong and Wong Fook Fei.
Universiti Kebangsaan Malaysia

ABSTRACT

Washback refers to “the extent to which the introduction and use of a test influences language teachers and learners to do things they would otherwise not do that promote or inhibit language learning” (Messick 1996:243). This paper presents a summary of the findings of a small-scale study of the initial washback effects of the recently introduced SPM 1119 (English) examination on teaching. The findings indicate that the new examination seems to have succeeded in bringing about generally positive washback. However, other aspects of the examination, such as the marking procedure and the lack of transparency concerning the assigning of two grades for one examination, are seen to be directly responsible for negative washback, which will defeat the good intentions behind the 1997 introduction of the paper. The writers see potential in the new paper for positive washback and offer suggestions to teachers and the testing authority on making washback work.

Introduction

The phenomenon of “washback” or “backwash” has been with us ever since the first examination, which, according to Bernard Spolsky (1997), was a Chinese invention. The examination was introduced during the Han Dynasty (2 BC) for the purpose of selecting suitable candidates for the Chinese Civil Service. The content of the examination and the way it was conducted soon became the “syllabus” for those who aspired to join the ranks of the “Mandarin”. The same phenomenon has been observed everywhere in the world where a public examination is the gateway to further opportunities: the examination controls teaching and learning. Washback therefore works forwards, rather than backwards, the examination being the starting point (see Pearson, 1988:98).

The various studies that have been carried out to investigate this phenomenon, notably the Sri Lanka study (see Wall and Alderson, 1993; also Shohamy, Donitsa-Schmidt & Ferman, 1996; Watanabe, 1996; Andrews, 1995; Lam, 1995; Alderson & Hamp-Lyons, 1996) have established clear evidence of washback, and it is now known that while washback is neutral, its effects can be potentially positive or negative. However, it is also accepted that “the generalized assertions about washback that abound in the literature and in the ESL/EFL profession are too simplistic” (Alderson & Hamp-Lyons, 1996:295) and the influence of other factors such as the status of the language, the uses of the test and teacher factors must also be acknowledged.

Nevertheless, the powerful and inevitable influence that examinations have on teaching and learning has led to ideas about harnessing washback to bring about change in the curriculum, that is, innovations in the language curriculum can be induced by innovations in language testing (Alderson 1986). Pearson (1988), for example, referred to washback as a “lever for change”, while Swain (1988), Hughes (1989) and Bailey (1996) urge educators and testwriters to “work for washback”. It is in this context of curriculum innovation that this paper is situated.

This paper is based on data obtained from a small-scale study carried out in 1999 to investigate the washback effects of a “new” English examination on teaching. The chosen focus is the impact on teaching because it is the most immediate factor to receive the impact of the new test. (The terms “examination” and “test” in this paper refer to the same assessment instrument, i.e. the SPM 1119).

Background to the SPM 1119

Prior to 1997, all Form 5 students in Malaysian schools had to take the SPM 1322 English examination, and those who were interested in pursuing their tertiary education in British institutions of higher learning took, in addition, the Cambridge O Levels 1119 English paper. Thus, while the SPM 1322 was a compulsory paper, the 1119 was a matter of individual choice and there was no provision for formal instruction in government schools to prepare students for it. In 1995, the Ministry of Education announced that these two English examinations were to be amalgamated into one paper called the SPM 1119. According to newspaper reports (*The New Straits Times*, 23 May 1995, *The Star*, 23, 25 & 28 May 1995), the introduction of this joint paper was prompted by falling pass rates in the SPM English examination and was intended to bring about two important changes: 1) an increase in the general level of English proficiency of Malaysian secondary students, and 2) an increase in the motivation of rural students to improve their competency. This two-in-one deal meant that all candidates would now obtain two grades, one for SPM English and another for O levels, by taking just one examination, the SPM 1119.

Differences between the SPM 1322 and the SPM 1119

Both papers test mainly three skills: Oral, Reading and Writing. The Listening skill is not tested and the Oral component remains unchanged in the SPM 1119, so only changes in the testing of Reading and Writing can be compared.

The most obvious changes are in the format, content and the test types used (see Figure 1 and Figure 2). It can be seen that there is a movement towards more direct testing of reading and writing skills, and a clearer demarcation between skills tested in Paper 1 and Paper 2. There is certainly more emphasis on testing production rather than recognition skills. At the same time, discrete testing of grammar and language functions has been reduced substantially. It was not possible to compare the weightage of the components in the old and new tests because the allocation of marks for each section was not indicated in the SPM 1322 papers.

The main changes introduced into the SPM 1119 may be summed up thus:

1. Use of short answer questions (SAQ) in the testing of reading comprehension
2. Use of shorter authentic, communicative text types (such as sales and job advertisements, notices, signs etc.) in testing reading comprehension
3. An integrated reading and writing task, i.e. summary writing
4. An open-ended writing task (composition)

Figure 1: Comparison of SPM 1322 and SPM 1119: Content

Old Paper, SPM 1322	New Paper, SPM 1119
<p style="text-align: center;"><u>Paper 1</u></p> <p style="text-align: center;">Reading comprehension</p> <p style="text-align: center;">Language functions</p> <p style="text-align: center;">Grammar</p>	<p style="text-align: center;"><u>Paper 1</u></p> <p style="text-align: center;">Reading comprehension</p> <p style="text-align: center;">Language functions</p> <p style="text-align: center;">Grammar</p>
<p style="text-align: center;"><u>Paper 2</u></p> <p style="text-align: center;">Grammar</p> <p style="text-align: center;">Language functions</p> <p style="text-align: center;">Directed writing</p>	<p style="text-align: center;"><u>Paper 2</u></p> <p style="text-align: center;">Directed writing</p> <p style="text-align: center;">Summary writing</p> <p style="text-align: center;">Free composition</p>

Figure 2: Comparison of SPM 1322 and SPM 1119: Test Methods

Old Paper, SPM 1322	New Paper, SPM 1119
MCQ	MCQ
Gap filling	SAQ
Word insertion	Gap filling
Error identification	Directed writing
Error correction	Summary writing
Guided response (prompts provided)	Free composition
Directed writing	

On the whole, the SPM 1119 seems to be more aligned to the syllabus objectives than the SPM 1322, and the test methods used are definitely more communicative and more focused on skills.

The Study

To assess the impact of the SPM 1119 English paper on classroom teaching, an essentially qualitative approach was used. This involved the use of questionnaires, semi-structured interviews, class observations, teaching logs and examination of relevant documents such as textbooks and revision books used by the teachers, and past years' examination papers. Our sample consisted of 18 English teachers from six secondary schools in the Klang Valley.

'Summary of the Findings

In order to understand how teachers teach in response to the introduction of the SPM 1119 examination, we needed to know teachers' perceptions of the examination's status and the extent of their knowledge of the new examination. We found that 50% (n=9) of the teachers attached different degrees of importance to the two papers. To them, Paper 1 had higher stakes than Paper 2 because they perceived that Paper 1 determined the SPM English grade and

Paper 2 the 1119 grade. The SPM grade was deemed more important as it could be used in the calculation of the overall aggregate whilst the 1119 grade was supplementary – “*good to have but not essential.*”

Most of the teachers considered themselves either well or reasonably informed of the new test as they had attended official briefings. During such briefings, teachers were given sample papers and the marking scheme. Some teachers had also attended workshops where they were asked to come up with test items that replicated those in the sample papers. 61%(n=11) of the teachers had marked the paper and were familiar with the marking scheme.

Although teachers professed to be well-informed about the examination, it appeared that they were more concerned about the types of test items that might appear rather than the sub-skills being tested in each test-type. This could be because the focus of the briefings was on the format and very little was said about the specific sub-skills that would be tested and how teachers could prepare students to acquire those skills.

Our findings indicate that the impact of the examination is mainly on *what* teachers teach and not so much on *how* they teach. The discussion on the washback effects is divided into two sections. The first summarizes the positive washback effects and the second, the negative effects.

Positive Washback

The positive washback is summarized below:

- **More Focus On The Teaching Of Writing Skills**

In line with the increased emphasis on writing in Paper 2, more than half (55.56%) of the teachers reported giving more attention to writing. Some teachers said they had started to teach specific writing skills. They felt that for the free writing task, students needed to be taught composition skills, for example, techniques for generating ideas and skills for organizing ideas. Previously, they did not have to teach students these skills since the question was a directed writing task and the stimulus given was sufficient, and correctly sequenced. A few teachers also claimed that they used the process writing approach and conducted one-to-one conferencing.

- **More Emphasis On The Teaching Of Grammar And Vocabulary**

Some teachers claimed to have intensified their teaching of grammar (22.22%) and vocabulary (33.33%) as these were seen as basic requirements to good writing skills.

- **More Emphasis On Teaching Specific Reading Skills**
Some teachers began to devote more time to teaching specific reading skills such as skimming and scanning, reading for gist, identifying main ideas, using contextual clues, and inferencing. This was in response to the introduction of the SAQ. Teachers felt that students needed to know how to process the text, locate the answer, select what is relevant to write out as an answer and paraphrase the answers. Previously, MCQ was the only test method used and teachers felt that there was less need to teach all these skills because the distractors helped students to identify the correct answers.
- **More Task-Based Lessons**
Teachers said that they tried to use the text types that appeared in the new paper (graphic materials and short authentic communicative texts) in their teaching. Thus, a number of teachers brought newspaper reading into the classroom, and lessons became more task-based.

Negative Washback

The negative washback effects are summarized below:

- **Weak Classes Were Not Taught Writing Skills**
Some teachers said that they were devoting more time to teaching writing only for the better classes while the weak classes were not taught writing at all because many students could not even write a complete sentence correctly. For these classes, the components of Paper 2, the writing paper, were not taught. This is tied to the teachers' belief that the SPM English grade was derived solely from Paper 1 whilst performance in Paper 2 determined the 1119 grade.¹
- **Teaching Test-Taking Strategies that are Detrimental to Good Teaching and Learning**
A teacher said that he instructed his students to lift directly from the text when writing the summary. He said that there were cases where students who were quite good were penalized because they paraphrased inaccurately. On the other hand, less able students could perform better by just lifting relevant parts from the original text. Another teacher said that she asked her students to memorize some essays so that they could write something even if it was not directly relevant to the question. She said some marks would be awarded for the essay even if it might be out-of-point. Some teachers said that they instructed students not to use their own words when answering the

¹ A check with an education officer from the Malaysia Examination Board proved this belief to be erroneous. The SPM English Grade and the 1119 grade are derived from a composite of scores from Paper 1, 2 and 3.

reading comprehension questions (SAQ) although the question explicitly stated that they had to. According to them, students were not penalized even if they lifted directly from the text, so there was no point in attempting to answer using one's own words.

- **Rejection of Textbooks**

Many of the teachers felt that textbooks were inadequate for preparing students for the examination. They preferred to source their materials mainly from commercial revision books. This indicates that teachers were more guided by the examination format, test methods and test types in deciding what to teach. Indeed teachers seemed to be more concerned with providing as much practice as possible on exam-like exercises than with teaching the students to develop the skills required to tackle the examination.

Conclusion

Although there has been no explicit official statement from the Ministry of Education regarding the theoretical framework which supports the changes introduced in the SPM 1119, our interpretation is that this new examination is a "tool" to bring about change in teaching and learning. Our view is that the SPM 1119 seems to have potential for positive washback, hence the message to teachers reflected in the title "Making Washback Work".

We offer the following suggestions to teachers on how the SPM 1119 can bring about positive change in their teaching.

Suggestions to Teachers:

1. **Go back to the syllabus objectives.**

Teachers, in their zeal and anxiety to prepare students for the examination, may have lost sight of the syllabus objectives. Notice that the communicative syllabus and the textbook have not been changed, only the examination has changed. Therefore, do not ignore the syllabus nor abandon the textbook. Giving students a lot of practice in exam-mimicking exercises is not teaching. It is also imprudent to rely solely on the judgment and "expertise" of commercial revision-book writers whose main concerns are sales and profit rather than real language learning.

2. **Focus on skills (but do teach test-taking strategies too)**

Be alert to the fact that the SPM 1119 is concerned with *direct testing* of skills, therefore there is no better way to train students than to teach them the

skills that are spelt out in the syllabus. For example, teaching summary writing involves teaching students to distinguish between main ideas and supporting details, to select those points that are relevant to the question, and then to synthesize them into a coherent piece of writing. Such cognitive activities are far more challenging and valuable to students than merely looking for matching words between question and text and lifting whole sections of the text.

We believe there is certainly a need for teaching test-taking strategies, but these should come second to the teaching of skills.

3. Develop good classroom management and rapport

The fact that the examination has now become more aligned to the communicative syllabus only points to the emphasis on using the communicative approach to language teaching. This calls for changing the conventional way that classes have been managed. For example, teachers could try out a different seating arrangement that would allow more eye contact and facilitate interaction among students. In an exam-oriented, teacher-directed class, students are passive and concerned with only getting his or her own answers right. In the communicative classroom, students see learning as a shared process, and are motivated to use the language by the very activities or tasks they perform.

Suggestions to the Testing Authority

We would like the testing authority to consider these suggestions on how positive washback may be further enhanced.

1. Be transparent in all aspects of the examination.

In our study, we discovered that teachers were relatively ignorant of the rationale for the change. They were also not aware of any test specifications, and if they had never been involved in marking the paper, they would not know about the marking scheme either.

- We have shown how lack of transparency can have negative repercussions. The teachers' perception that there are two papers leading respectively to two separate grades has influenced a few teachers to adopt a pragmatic approach to concentrate only on Paper 1 when teaching classes of very low proficiency. Teachers are entitled to know exactly how grades are derived so that they can make informed decisions on what strategies to take.

We suggest that information about all aspects of the examination (including test objectives, specifications, format, weightage) should be easily available to anyone who has a stake or interest in the exam.

2. Involve teachers (change agents) from the start

The authorities should recognize the important role teachers play as change agents. They should be involved from the start in determining the objectives of the test, the rationale for the test design, the test methods and even the marking scheme. Teachers who are informed and who are involved in some way are likely to be more effective in making washback work.

3. Provide training for teachers in teaching methodology, classroom management

In our study, we were told that there were briefings where teachers were informed about the format of the new exam, and that workshops were conducted where teachers wrote practice exercises that mimicked the sample exam items given to them. It would be more useful to run refresher courses for teachers in teaching approaches, classroom management, even in testing, all geared towards more effective teaching of the syllabus. Training in teaching summary writing and writing skills would be especially pertinent since these are new components in the examination.

4. Keep the channels of communication open

The teachers we interviewed were unhappy that their feedback regarding the examination or the marking scheme seemed to be ignored. At the same time, they felt that they should have got feedback from the examination authorities regarding the areas that students generally performed poorly in, so that they could review their teaching to remedy the problems. It is important for the authorities to keep the communication channels open between them because teachers have much that is valuable to contribute to the improvement of the examination.

5. Build washback into the test design.

Finally, if our perception about the examination being used as a tool to engineer curriculum innovation is correct, then it is imperative that washback be thoughtfully worked into the design of the test. For example, although the present SPM 1119 seems to be more communicative than its predecessor is, we are puzzled by the inclusion of the items testing language functions, and the rational cloze. These items, in our opinion, are unlikely to bring about favourable washback because it is not clear just what skills are being tested, and it is doubtful that the test methods used are communicative. There is also a need to improve on the authenticity of test tasks such as those in the summary and open-ended writing questions. The past years' questions

on Writing reveal no indication of audience or purpose while the Summary questions lack communicative intent because they are not task-based.

In conclusion, we see that the SPM 1119 has potential for positive washback. At the same time, we are mindful that it takes more than the introduction of a new examination to bring about change in teaching and learning.

References

- Alderson, J.C. 1986. Innovations in Language Testing, in M. Portal (ed). *Innovations in Language Testing*. London: NFER/Nelson.
- Alderson, J.C. & Wall, D. 1993. Does Washback Exist? *Applied Linguistics*, 14, 115-129.
- Alderson, J.C. and Hamp-Lyons, L. 1996. TOEFL Preparation Courses: A study of Washback. *Language Testing*, Vol. 13, 280-297.
- Andrews, S. 1995. Washback or Washout? : The Relationship between Examination Reform and Curriculum Innovation. In D. Nunan, R. Berry and V. Berry, (eds). *Bringing about Change in Language Education* (Proceedings of the ILEC 1994). Hong Kong.
- Bailey, K. 1996. Working for Washback: A review of the Washback Concept in Language Testing. *Language Testing*, Vol. 13, 257-279.
- De Paul, V. 1995. Tougher English Exam. *The New Straits Times*, 23 May, 1.
- Goh, L. 1995 .Two Grades for Joint Paper. *The Star*, 23 May, 1 & 5.
- Hughes, A. 1989. *Testing for Language Teachers*. Cambridge: Cambridge University Press.
- Lam, H.P. 1995. Methodology Washback – An Insider’s View, in D. Nunan, R. Berry, and V. Berry (eds). *Bringing about Change in Language Education* (Proceedings of the ILEC 1994). Hong Kong.
- Messick, S. 1996. Validity and Washback in Language Testing. *Language Testing*, Vol. 13, 241-245.
- Ministry of Education. 1989a. *Curriculum Specifications for English Language Form Four*. Kuala Lumpur: Ministry of Education.

- Ministry of Education. 1989b. *Curriculum Specifications for English Language Form Five*. Kuala Lumpur: Ministry of Education.
- Pearson, I. 1988. Tests as Levers for Change. In D. Chamberlain, and R. Baumgartner (eds). *ESP in the Classroom: Practice and Evaluation*. ELT Document 128. London: Modern English Publications.
- Shohamy, E. 1993. The Power of Tests: The Impact of Language Tests on Teaching and Learning. NFLC Occasional Paper. Washington DC: National Foreign Language Center.
- Shohamy, E., Donitsa-Schmidt, S. and Ferman, G. 1996. Test Impact Revisited: Washback Effect Over Time. *Language Testing*, Vol. 13, 297-317.
- Spolsky, B. 1997. The ethics of gatekeeping tests: what have we learned in a hundred years? *Language Testing*, Vol.14, 242-247.
- Swain, M. 1985. Large-Scale Communicative Language Testing: A Case Study. In Lee Y. P. et al. (eds). *New Directions in Language Testing*. Hong Kong: Pergamon Press.
- Wall, D. and Alderson, J.C. 1993. Examining Washback: The Sri Lankan Impact Study. *Language Testing*, Vol.10, 41-69.
- Watanabe, Y. 1996. Does Grammar Translation Come From the Entrance Examination? Preliminary Findings from Classroom-Based Research. *Language Testing*, Vol.13, 318-333.
- Wong, H. 1995. New Paper Will Test Students' True Ability. *The Sunday Star*, 28 May, 28.