



The Nature and Impact of Textbook-based EFL Vocabulary Tests on Teaching and Learning

Η Φύση και Επιρροή των Διαγωνισμάτων Λεξιλογίου των Διδακτικών Εγχειριδίων στη Διδασκαλία και Εκμάθηση της Αγγλικής ως Ξένης Γλώσσας

Dina TSAGARI & Athena PAVLOU

The current study examines the nature and impact of the vocabulary tests included in the EFL textbook used in the 5th grade of primary state schools in Greece. To achieve its aims, the study employed both qualitative and quantitative research methods. Analysis of the vocabulary tests showed that these place emphasis on students' lexical competence while vocabulary teaching encouraged in the textbook focuses mostly on overall vocabulary performance. Responses to questionnaires revealed that a significant number of teachers has a negative attitude towards the vocabulary tests under study due to their incompatibility with their students' language level. Consequently, teachers adapt these tests or design their own according to their teaching practices and students' linguistic level. However, they, too, seem to favour assessment of specific lexical elements in the tests they construct. This raises questions with regard to the effectiveness of teachers' instructional and assessment practices and the extent to which these are in line with the writers' intentions enacted in the textbook materials, a central issue in the field of materials design and use. The paper concludes with suggestions for teachers, trainers, materials designers and researchers.

∞

Η παρούσα μελέτη εξετάζει τη φύση και τις επιπτώσεις των δοκιμασιών (τεστ) λεξιλογίου που περιλαμβάνονται στο διδακτικό εγχειρίδιο της Αγγλικής ως ξένης γλώσσας το οποίο χρησιμοποιείται στην 5η τάξη δημοτικού σε δημόσια σχολεία στην Ελλάδα. Για την επίτευξη των στόχων της, η μελέτη χρησιμοποίησε τόσο ποιοτικές όσο και ποσοτικές μεθόδους έρευνας. Ανάλυση των τεστ λεξιλογίου κατέδειξε ότι αυτά δίνουν έμφαση στην αναγνώριση και παραγωγή λεξιλογίου των μαθητών, ενώ στην διδασκαλία λεξιλογίου του εγχειριδίου ενθαρρύνονται κυρίως οι συνολικές επιδόσεις του λεξιλογίου. Οι απαντήσεις σε ερωτηματολόγια κατέδειξαν ότι ένας σημαντικός αριθμός των εκπαιδευτικών έχει αρνητική στάση απέναντι στα τεστ λεξιλογίου λόγω ασυμβατότητάς τους με το επίπεδο των μαθητών τους. Ως εκ τούτου, οι εκπαιδευτικοί προσαρμόζουν τα εν λόγω τεστ ή σχεδιάζουν τα δικά τους σύμφωνα με τις πρακτικές διδασκαλίας τους και το γλωσσικό επίπεδο των μαθητών.

Οι εκπαιδευτικοί επίσης φαίνεται να ευνοούν την εξέταση μεμονωμένων λεξιλογικών στοιχείων στα τεστ που σχεδιάζουν. Αυτό εγείρει ερωτήματα σχετικά με την αποτελεσματικότητα των πρακτικών διδασκαλίας και αξιολόγησης των εκπαιδευτικών και με το βαθμό ως προς τον οποίο είναι σύμφωνες με τις προθέσεις των διδακτικών εγχειριδίων, ένα κεντρικό ζήτημα στον τομέα του σχεδιασμού και χρήσης διδακτικών υλικών. Η μελέτη καταλήγει με προτάσεις για τους εκπαιδευτικούς, επιμορφωτές, συγγραφείς διδακτικών υλικών και ερευνητές.

Key words: Textbook material evaluation, vocabulary assessment, test impact, teachers' beliefs and practices.

1. Introduction

Evaluation of textbook material has been in the forefront of research in language teaching and learning in recent years (Tomlinson, 2012) providing invaluable insights to foreign language pedagogy. Such evaluation is particularly important for teachers and materials designers as it provides them with constructive feedback with regard to the effectiveness of teaching materials (Dickins & Germaine, 1992). However, review of the literature in materials evaluation has shown that although extensive research has been conducted so far (Dickins & Germaine, 1992; Ellis, 1998; Harwood, 2010; Littlejohn, 1998; Mc Donough & Shaw, 1993; Mc Grath, 2002; Sheldon, 1987; Tomlinson, 2012), few researchers have dealt with the evaluation of vocabulary in coursebooks (Cunningsworth, 1984; 1995; Hedge, 2000). What is more, hardly any research has focused on the evaluation of vocabulary tests that constitute part of the 'coursebook package' or their effects on teachers and students (Tomlinson & Masuhara, 2010). It seems that the field of language teaching and testing has not yet recognized the salient place that textbook-based tests have recently gained especially with regard to their impact on classroom teaching and learning.

The present study aims to address this gap in the literature by investigating the nature and impact of vocabulary tests of the EFL textbook used in the 5th grade of primary state schools in Greece through multiple sources of empirical data. The paper presents the results of the study and discusses implications which offer useful insights to EFL teachers, teacher trainers and textbook designers.

2. Theoretical perspectives in vocabulary assessment

Researchers have shown a growing interest in the study of lexical knowledge in recent years. With the advent of the communicative approach, grammar lost its salient place while emphasis was placed on vocabulary learning and assessment (Alderson, 2000; Nation & Newton, 1997; Read & Chapelle, 2001; Read, 2000; 2004; Schoonen & Verhallen, 2008; Thrasher, 2000). Vocabulary assessment in particular witnessed several developments. Conventional vocabulary tests usually contained discrete-point items assessing students' knowledge of the meaning and usage of a set of words. Such tests focused on assessing vocabulary elements in isolation or at sentence level that required learners to recognize the meaning of a given word or recall it from memory (Read, 2000). More recently, however, comprehensive or embedded vocabulary tests were designed that involved vocabulary assessment through tasks which simulate real-life situations (referred to as 'meaningful tasks' in this study). These vocabulary tests assess whether students can recognize certain words when they encounter them in context, e.g. within a listening or a reading task, or

whether they can recall and use them either orally or in written form; as such, students' vocabulary knowledge and use are assessed receptively and productively within a larger construct (Read, 2000). In view of this, more precise knowledge of a given lexical item is required for language use (Nation, 2001) which makes meaning-focused tasks more demanding for students. Therefore, to achieve successful balance between vocabulary acquisition and use, vocabulary assessment needs to involve evaluation of both specific lexical items, namely, language-focused test tasks as well as students' overall language performance, that is, meaning-focused test tasks. The same balance is also encouraged in language learning programmes. For example, Nation (2001, p. 232) stresses that a well-designed language learning programme should balance opportunities to learn from message-focused activities as well as from the direct study of language items, with language-focused activities covering a smaller proportion of the total learning programme (also supported by Hulstijn, 2001). The aim of the present study is to explore the extent to which real-life and discrete-point test tasks are included in vocabulary tests produced as part of a textbook series, their relation with the corresponding textbook activities and impact on vocabulary learning.

The next section will focus on defining the qualities of vocabulary tests. This will be used as a point of reference in the presentation and discussion of the results at the end of this paper.

2.1. Qualities of good vocabulary tests

There are certain qualities which affect the degree of usefulness and fairness of language tests. Generally speaking, a test is considered useful and fair by the extent to which it possesses a balance of qualities such as reliability, validity, practicality, utility, authenticity and impact (Bachman & Palmer, 1996; Hughes, 1989; Mishan, 2010; Purpura, 2004). More specifically, a test is reliable when it yields consistent results (Harmer, 2001; Hughes, 1989). With regard to vocabulary tests, language-focused tests tasks, which assess individual lexical items, are considered to be objective and highly reliable. They are also practical to score and provide diagnostic information and other type of feedback to teachers and learners concerning students' vocabulary knowledge (Bachman & Palmer, 1996; Herman, Aschbacher & Winters, 1992; West, 2004). Meaning-focused test tasks are subjective as vocabulary is measured within a larger construct, and therefore, these tests are of high validity. Feedback collected sheds light on learners' future language performance beyond the test to language use in other meaningful contexts outside the classroom. These vocabulary tasks are authentic (Abdullah, 1998 & Tomlinson 1998b, cited in Mishan, 2010; Bachman & Palmer, 1996; West, 2004) as language is used for genuine purposes (Guariento & Morley, 2001). Hence, positive influence is exerted on students' learning (Peacock, 1997).

With regard to validity, Read (2000, p. 151) states that a test is valid when we are able to draw meaningful conclusions from its results. More specifically, a vocabulary test exhibits face validity when it measures the lexical items or skills it is supposed to measure. In addition, a vocabulary test has predictive validity when it assesses future language performance, namely when it predicts the degree of success of learners in using the language in the future through meaning-focused test tasks (Hughes, 1989; Underhill, 1987). What is more, a test is said to have content validity if its content reflects the language skills, the syllabus or the students' language needs. The greater the content validity of a test, the more likely it is to measure accurately the language abilities of students or the construct it is designed to measure and, hence, exhibits construct validity (Davies 1968, cited in Alderson & Hughes, 1979/1981; Hughes, 1989; West, 2004). The degree of fairness a vocabulary test

may exhibit depends on the extent to which the test is reliable and valid (Hamp-Lyons, 2000; Shohamy, 2000).

Another quality of a 'good' test is the impact it exerts on students, teachers and even society. The direct impact of testing on teaching and learning is referred to as 'washback' (Alderson & Wall, 1993). This can be either positive or negative. When a test is based on the learning objectives of the curriculum (Heaton, 1990) then it has a positive washback effect; otherwise negative washback is exerted on teaching and learning (Bachman & Palmer, 1996; Harmer, 2001; Harris & Mc Cann, 1994). Negative washback is also reflected in test scores, as they are influenced by the characteristics and contents of the test tasks (Griffin 1992, cited in Nation 2001; Cheng & Curtis, 2004). In addition, when teachers limit their teaching practices to test requirements, then negative influence is likely to be exerted on teaching as teachers run the risk of 'teaching to the test' and thus ignore other equally important learning aspects (Cheng & Curtis, 2004; Leung & Andrews, 2012; Tsagari, 2009, 2012).

A further notion of impact pertains to predictive validity. When students are assessed through real-world tasks that measure and predict their future language performance (Wesche, 1983) then they exert beneficial washforward effect on students (Peacock, 1997; West, 2004) which also needs to be emphasized in teachers' practices (Watanabe, 2004). It is crucial, therefore, that a vocabulary test exerts positive washback on teaching and learning and has beneficial washforward effect, as the influence of a test exercised on learning will feed back into students' overall language performance.

In the present study we explore and discuss the extent to which textbook vocabulary tests follow good practice in language testing. The next two sections provide a brief overview of the research context and describe the research design employed in the study undertaken.

2.2. The current teaching and assessment context

The language level of the 5th grade primary students attending EFL classes in the Greek state educational system, claimed to be Basic Users of the English language (Pedagogical Institute, 2011), actually ranges from learners who belong to the Breakthrough stage (A1) to those who can be traced at the weak end of the Waystage level (A2.1) (Council of Europe, 2001). With regard to their lexical competence at this level, students are expected to have sufficient vocabulary for the expression of basic communicative needs and control a narrow repertoire relevant to everyday needs (Council of Europe 2001, p. 112).

The English textbook series employed recently in the 5th grade of primary schools of the present context is a series called 'English 5th' (Κολοβού & Κραγιώτου, 2009), published by the Ministry of Education. It consists of a Pupil's book, a Workbook, a Teacher's book and audio CDs. The Pupil's book contains ten Self-assessment tests presented at the end of every chapter. There are also three Progress tests in the Teacher's book (at the end of every three chapters) and one final Progress test in the Pupil's book (upon completion of all ten chapters).

The Cross-Curricular Thematic Framework (Pedagogical Institute, 2003), used in the Greek primary state education, the Presidential Decree of 8/1995 and the CEFR (Council of Europe, 2001) formed the bases of the design of the present textbook. Vocabulary assessment is not explicitly mentioned within the Cross-Curricular Thematic Framework. However, vocabulary assessment within a larger construct is encouraged through the communicative skills of

speaking, listening, reading and writing. Recently, within the new Syllabus of Foreign Languages, namely the Common Curriculum of Foreign Languages (Pedagogical Institute, 2011), certain illustrative descriptors are employed which define what the learners can do (also in line with the CEFR). In the new syllabus, assessment of overall language performance is encouraged, too (Pedagogical Institute, 2011, pp. 28-31).

3. Research design

3.1. Rationale

When 'English 5th' was distributed in schools in 2009, teachers welcomed it as an aid towards accomplishing the objectives of their lessons (Mc Grath, 2002). However, teachers expressed doubt and skepticism towards the vocabulary materials included in the textbook complaining that the instructional materials and the accompanying vocabulary tests did not correspond to their students' level (see also Tsagari and Sifakis, 2012). Consequently, many teachers complained that they had to adapt the vocabulary test tasks of the textbook or design their own.

Given this state of affairs, the researchers decided to systematically analyse and evaluate the vocabulary tests included in the textbook. The study also aspired to elicit teachers' beliefs and instructional practices towards vocabulary learning and assessment as, other than initial teachers' complaints, there was scant empirical evidence about what teachers actually believe or do when they work on vocabulary in the classroom (also in Tomlinson, 2012). The identification of these areas led to the formulation of certain research questions that constitute the basis of the present study, e.g.:

1. What is the nature of the vocabulary tests included in the 5th grade EFL textbook used in the primary schools in Greece?
2. Are these vocabulary tests in accordance with vocabulary teaching advocated in the textbook?
3. What is the impact of these vocabulary tests on student achievement?
4. What are teachers' attitudes and practices towards vocabulary teaching?
5. What are teachers' attitudes and practices towards vocabulary test tasks included in the textbook?

It is necessary to highlight that the purpose of the study was not to highlight the negative aspects of the textbook material under examination but rather to explore the effectiveness of the vocabulary tests of the textbook under study and discuss implications for language teaching and assessment (Cunningsworth, 1979, cited in Mc Grath, 2002; Tsagari, 2009).

3.2. Research Methodology

The present study employed an in-depth analysis method for the collection and analysis of textbook and test data (Mc Grath, 2002). The in-depth method was used in three phases as these have been proposed by Breen (1989), illustrated below:

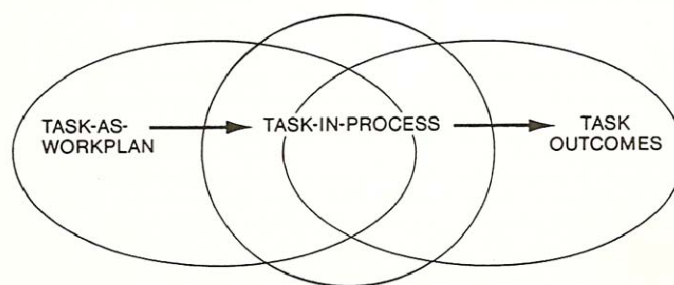


Figure 1. Three phases of a task

Task-as-workplan

At this phase, content analysis was conducted focusing on the nature of the vocabulary activities included in

- a. all the Progress tests of the Teacher's book
- b. all the Self-assessment tests of the Pupil's book and
- c. the first three chapters of the Pupil's book

Owing to the length of the textbook material (10 chapters in total – 152 pages), it was impractical to analyse the entire content of the vocabulary activities of the Pupil's book (Cunningsworth, 1995; Ellis, 1998; Littlejohn, 1998). Therefore, for the purposes of the study, only the first three chapters were analysed.

The data collected from this analysis was expected to answer the first two research questions (Section 3.1). The analysis framework was informed by relevant bibliography (Bülent & Stoller, 2005; Byrd, 2001; Hill, 2004; Luoma, 2004; Purpura, 2004; Ur, 1996; West, 2004) that will be discussed in detail in Section 3.3.

Task-in-process

During the second phase, information was collected about the ways in which teachers respond to and use vocabulary tasks and tests so that conclusions could be drawn about their effectiveness (Dickins & Germaine, 1992). This data was collected through questionnaires administered to EFL primary state school teachers. This part of the research intended to answer the last two research questions (Section 3.1).

Task outcomes

Finally, the third phase of the evaluation cycle aimed to analyse what had been learned (Breen, 1989). Therefore, the researchers collected and analysed the results obtained from the textbook vocabulary tests administered to students and then compared these with previous test task analysis and teachers' responses. After all, test scores are influenced by the characteristics and contents of the test tasks (Cheng & Curtis, 2004; Griffin 1992, cited in Nation, 2001). For practicality reasons again, results were collected from the administration of the vocabulary tests included in:

- a. the first three Self-assessment tests of the Pupil's book and
- b. the first Progress test included in the Teacher's book

The results were analysed quantitatively (Burns, 1990; Dickins & Germaine, 1992). This last phase also aimed at answering the third research question (Section 3.1) with regard to the impact of the vocabulary tests on students' learning.

3.3. Research instruments

The evaluation checklist

The first step in textbook evaluation is to identify the general categories and/or criteria to be used in the analysis of any given EFL textbook (Tomlinson 2003b, cited in Tomlinson, 2012). In the present study, the generation of specific criteria was based on the study of the relevant literature on published checklists (Cunningsworth, 1984, 1995; Hedge, 2000; Mc Donough & Shaw, 1993; Miekley, 2005; Williams, 1983) and was informed by the literature on vocabulary assessment (Section 2.1.). Overall, the checklist designed was divided into two parts, each consisting of 11 questions (Appendix I). The first part focused on vocabulary teaching and learning, whereas the second one on vocabulary assessment. Both parts employed quantitative and qualitative systems of elicitation (Tsagari, 2009). These consisted of closed questions ('Yes' and 'No' questions) which also provided space for comments so that the checklist users (in this case the present researchers) could further expand on their responses (Mc Grath, 2002).

The teachers' questionnaire

The teachers' questionnaire (Appendix II) aimed at collecting teachers' beliefs and practices towards vocabulary teaching and assessment in general as well as those based on the textbook. Following good practice in research design, the questionnaire underwent piloting. A group of ten EFL teachers working in primary state education, as similar as possible to that in the main study (Burns, 1990), answered the questionnaire. The results of this pilot work led to a number of revisions such as the inclusion of scaled items as there were too many open-ended questions which were answered vaguely, therefore exhibiting poor reliability.

In its final version, the questionnaire contained two sections: the first one, divided in five parts, included questions based on a five-point Likert scale (e.g. 1: 'strongly agree', 5: 'strongly disagree') so as to achieve greater uniformity of measurement, and therefore, greater reliability (Bartels, 2005; Burns, 1990). The second section included only free-response questions to encourage the respondents to express their beliefs and attitudes concerning the topic under study (Oppenheim, 1992).

Sixty-seven teachers took the questionnaire. The majority of the respondents were female English language teachers working in primary state education. Overall, teachers had between 6 to 10 years teaching experience (1-5 years spent in primary schools). All respondents held a university degree, which is a prerequisite for the employment of teachers in the state school sector. A small number held a Master's and PhD degrees. Collection of data commenced on December 24, 2009 and ended on January 30, 2010. Teachers took about 15-20 minutes to complete the questionnaire.

Learners' outcomes

Finally, students' test scores were obtained. These scores were based on the administration of the first three Self-assessment tests and the first Progress test (based on Units 1-3). Test

administration commenced on 15th October, 2009 and ended on 18th January, 2010. The 34 students involved in this study were eleven-year-old 5th graders of mixed ability and learning styles, of Greek and Albanian origin residing on the east coast of South Greece.

4. Presentation of results

4.1. Textbook analysis

Given the space confines of the present paper, we will present findings with regard to the types of vocabulary tasks in textbook materials and tests analysed. The interested reader can consult Pavlou (2010) for the presentation of the full set of data.

Types of tasks	Vocabulary activities		
	Self-assessment tests - Pupils' book (Units 1-10)	Progress tests - Teacher's book (Units 1-10)	Pupils' book (Units 1-3)
Meaning-focused activities (overall language performance)	3 (21%)	3 (23%)	39 (81%)
Language-focused test tasks (specific vocabulary items)	11 (79%)	10 (77%)	9 (19%)
Total	14	13	48

Table 1. Frequency of vocabulary tasks in textbook materials and tests

Overall, the analysis showed that both meaning- and language-focused vocabulary activities are included in the textbook tests and the Pupil's book (Table 1). However, the vocabulary test tasks in both the Self-assessment and Progress tests place more emphasis on the assessment of individual lexical items (79% and 77% of tasks respectively) rather than on vocabulary assessment within a larger construct. By contrast, meaning-focused tasks (where emphasis is placed on communication and message transmission) cover a much larger proportion in the textbook (81% of the tasks included). This means that the textbook materials emphasize overall language use in the teaching of vocabulary.

4.2. Test score analysis

This part of the analysis addresses the second research question concerning the impact of the vocabulary tests on student learning/performance.

	Self- assessment Test 1	Self- assessment Test 2	Self- assessment Test 3	Progress Test 1
Mean score	75.09	90.36	88.10	75.44
Mode	100	100	100	89
Variance	470.75	184.96	317.02	385.54
Standard Deviation (SD)	21.70	13.60	17.81	19.64

Table 2. Distribution of test scores

Results from the analysis of test scores (Table 2) indicated that most students produced high scores in Self-assessment tests 2 and 3, where vocabulary was assessed in limited context through language-focused test tasks that required recall or recognition of word forms. Actually most of the students scored high in these test tasks. This is confirmed by the normal distribution of the data in Self-assessment tests 2 and 3 as the test scores are clustered around the mean scores (SD: 13.60 and 17.81 respectively) (Carr, 2008; Niles, n.d).

Contrary to the above, students were presented with two meaning-focused test tasks based on a reading text in Self-assessment test 1 and Progress test 1 (e.g. a dual choice and an information transfer task). As a result students exhibited slightly lower performance. The test scores are more spread apart (SD: 21.7 for Self-assessment 1 and SD: 19.64 for Progress test 1) exhibiting a relatively larger standard deviation (Table 2). What this means is that a significant number of students had difficulty and, therefore, received relatively lower marks (Carr, 2008).

The following section reports results from the Teachers' questionnaire which attempted to answer the fourth and fifth research question (Section 3.1)

4.3. Questionnaire results

The questionnaire results (Table 3) show that, in general, the majority of the teachers are in favour of meaning-focused vocabulary activities (96%) and tests tasks (90%) (Qu. 2 & 4.3). Teachers also prefer language-focused activities and test tasks at a slightly smaller extent (85% & 72%) (Qu. 3 & 4.2). However, teaching and assessment of isolated words (Qu. 1 & 4.1) is supported by a much smaller percentage of respondents (27% & 19% respectively).

Questions	Stongly Agree (1)	Agree (2)	Neutral (3)	Disagree (4)	Strongly disagree (5)
<i>Part 2: Vocabulary Instruction and Acquisition</i>					
Qu. 1: Vocabulary should be learned through systematic vocabulary teaching	6 (8,96%)	12 (17,91%)	19 (28,36 %)	22 (32,84%)	8 (11,94%)
Qu. 2: Vocabulary should be learned in context focusing on the message to be transmitted	39 (58,21%)	25 (37,31%)	1 (1,49%)	2 (2,99%)	0
Qu. 3: Vocabulary should be practiced through vocabulary exercises such as matching words with definitions, etc.	24 (35,82%)	33 (49,25%)	8 (11,94 %)	2 (2,99%)	0
<i>Part 3: Vocabulary Assessment</i>					
Qu. 4.1: Vocabulary should be assessed as isolated words	1 (1,49%)	12 (17,91%)	18 (26,87 %)	21 (31,34%)	0
Qu. 4.2: Vocabulary should be assessed at sentence level.	8 (11.94%)	40 (59.70%)	16 (23.88 %)	2 (2.99%)	1 (1.49%)
Qu. 4.3: Vocabulary should be assessed through meaningful tasks.	40 (59.70%)	20 (29.85%)	5 (7.46%)	1 (1.49%)	1 (1.49%)
<i>Part 4: Textbook Material</i>					
Qu. 5: Vocabulary covered in the	1	14	13	28	11

textbook is appropriate for my learners' current level.	(1,49%)	(20,90%	(19,40 %)	(41,79%)	(16,42%)
Qu. 6: The work done on vocabulary through reading, listening, speaking and writing is meaningful and engages learners in real-life communication.	3 (4,48%)	24 (35,82%)	22 (32,84 %)	15 (22,39%)	3 (4,48%)
Qu. 7: The vocabulary tests are relevant to my students' level.	2 (2,99%)	17 (25,37%)	18 (26,87 %)	25 (37,31%)	5 (7,46%)
Qu. 8: The vocabulary tests assess the words and/or phrases students have practised in class.	9 (13,43%)	37 (55,22%)	17 (25,37 %)	3 (4,48%)	1 (1,49%)
Qu. 9: The vocabulary test tasks are similar to the textbook activities.	3 (4,48%)	35 (52,24%)	15 (22,39 %)	13 (19,40%)	1 (1,49%)
Qu. 10: The vocabulary test tasks are meaningful engaging learners in real-world language use.	1 (1,49%)	10 (14,93%)	27 (40,30 %)	24 (35,82%)	5 (7,46%)
<i>Part 5: Use of Textbook Tests</i>					
Qu. 11a: I use the Self-assessment tests to assess my students' vocabulary knowledge and use.	16 (23,88%)	25 (37,31%)	12 (17,91 %)	8 (11,94%)	6 (8,96%)
Qu. 11b: I use the Progress tests to assess my students' vocabulary knowledge and use.	19 (28,36%)	34 (50,75%)	8 (11,94 %)	3 (4,48%)	3 (4,48%)
Qu. 11c: I use the Final Progress test to assess my students' vocabulary knowledge and use.	9 (13,43%)	23 (34,33%)	16 (23,88 %)	12 (17,91%)	7 (10,45%)
Qu. 12a: I employ these tests to collect feedback on my students' progress and organize remedial work.	28 (41,79%)	34 (50,75)	3 (4,48%)	2 (2,99%)	0
Qu. 12b: I employ these tests because they are quick and practical to score.	19 (28.36%)	27 (40.30%)	14 (20.90 %)	5 (7.46%)	2 (2.99%)
Qu. 12c: I employ these tests because they are consistent in their scoring.	11 (16,42%)	28 (41,79%)	20 (29,85 %)	7 (10,45%)	1 (1,49%)
Qu. 12d: I employ these tests because they measure my students' vocabulary knowledge.	9 (13,43%)	33 (49,25%)	21 (31,34 %)	4 (5,97)	0
Qu.12e: I employ these tests because they measure learners' future language performance.	2 (2,99%)	8 (11,94%)	30 (44,78 %)	20 (29,85%)	7 (10,45%)
Qu. 12f: I employ these tests to assess my students' vocabulary knowledge and use because they reflect the vocabulary learners have been taught in class.	8 (11,94%)	43 (64,18%)	11 (16,42 %)	5 (7,46%)	0

Table 3. Answers to scaled items of the teachers' questionnaire

Most teachers (69%) (Qu. 8) believe that the textbook tests assess vocabulary previously practised in class. Furthermore, the majority of teachers believe that the vocabulary test

tasks are similar to the textbook activities (57%) (Qu. 9). However, irrespective of teachers' positive attitudes towards the textbook-based tests, a significant percentage of teachers also believes that the textbook vocabulary activities and the test tasks do not correspond to their students' current language level (58% & 45%) (Qu. 5 & 7). This is also confirmed by Qu. 6, according to which only 40% of the respondents who believe that the textbook activities are meaningful, and by Qu. 10, where only 15% believe that the vocabulary test tasks meaningfully engage learners in real-world language use. Interestingly an equal percentage (40%) holds a neutral view on this. (This is also in line with analysis of test tasks that showed that these are far from meaningful, see Table 1). Interestingly, an equal percentage (40%) holds a neutral view on this.

<i>Qu. 13</i>	<i>I adapt the vocabulary tests because</i>	<i>Frequency</i>
a	They do not reflect my students' language level and needs.	31 (81,58%)
b	I want to measure specific aspects of vocabulary I have practiced in class.	3 (7,89%)
c	I want to give my students more real-life tasks.	4 (10,53%)
<i>Qu. 14</i>	<i>When I adapt the vocabulary test tasks because</i>	
a	I try to assess new vocabulary in meaningful contexts through reading, listening, speaking and writing tasks.	7 (35%)
b	I try to simplify the tasks to suit my students' level.	6 (30%)
c	I use tasks my students have practised in class and they are familiar with.	5 (25%)
d	I add both easy and difficult tasks to cover every language level of my students.	2 (10%)
<i>Qu. 15</i>	<i>I design my own vocabulary tests</i>	
a	According to my students' language level.	15 (60%)
b	Because I wish to measure specific aspects of vocabulary I have practiced in class.	10 (40%)
<i>Qu. 16</i>	<i>The vocabulary test tasks I prefer to design are</i>	
a	Language-focused test tasks	32 (78,05%)
b	Meaning-focused test tasks because students should have a realistic aim when they learn a foreign language and they should learn vocabulary in context.	7 (17,07%)
c	Both meaning- and language-focused test tasks because I want to assess my students' vocabulary knowledge and use.	2 (4,88%)
<i>Qu. 16.1</i>	<i>Reasons for employing language-focused test tasks</i>	
a	They are practical to score and I can identify which words my students have learned.	10 (31,25%)
b	Students are familiar with them through previous classroom practice.	4 (12,50%)
c	Students find them easier to deal with and thus more interesting.	18 (56,25%)

Table 4. Answers from free response items of the teachers' questionnaire

With regard to the ways teachers assess their students' vocabulary knowledge and use, the findings showed that most respondents do employ the tests offered in the textbook though at different frequency, e.g. 79% of the teachers uses the Progress tests (Qu. 11b), 61% uses the Self-assessment tests (Qu. 11a) while 48% makes use of the Final Progress test (Qu. 11c). Teachers said they use the tests because:

- they want to check their learners' progress and organize remedial work (93%) (Qu. 12a),
- the tests reflect vocabulary taught in class (76%) (Qu. 12f),
- for reasons of practicality (69%) (Qu. 12b),
- teachers want to measure learners' vocabulary knowledge (63%) (Qu. 12d) and
- the tests are highly reliable (58%) (Qu. 12c).

As to the tests' predictive validity (Qu. 12e), 40% supports that the tests fail to measure students' future language performance while a significant percentage (45%) holds a neutral stance.

Teachers further confirmed and explained the reasons for the perceptions they have towards vocabulary test tasks through the open-ended items included in the second part of the questionnaire (Table 4).

In line with the answers to the scaled questions presented in Table 3, a significant number of teachers further confirmed that the vocabulary tests do not correspond to students' linguistic level (Table 4, Qu. 13a) and, therefore, need to adapt test tasks or design their own according to their learners' level (Qu. 15a). However, as teachers explained, they prefer language-focused test tasks mostly (Qu. 16a & b; 16.1). This accords with their teaching practices in class (Qu. 15b). As a result, fewer teachers use a meaning-focused orientation when they adapt the vocabulary tests or design their own (Qu. 13c, 14a, 16b & c). However, this practice contradicts the results collected from the analysis of Part 2 and 3 of the questionnaire where teachers supported quite the opposite (Qu. 2 and 4.3).

The next section summarises of the main findings and elaborates on the implications of the present research study.

5. Summary and discussion of results

Motivated by the scant empirical research evidence in textbook-based vocabulary teaching and testing as well as the complaints about the effectiveness of the vocabulary tests of a specific EFL textbook expressed by teachers, a research programme was undertaken to investigate the nature and impact of such materials on teaching and learning.

Data collected from the analysis of the test tasks included in the vocabulary tests of the 5th grade EFL textbook used in primary schools Greece show that these mostly focus on students' lexical competence, rather than overall language performance typical of conventional vocabulary tests. On the contrary, most textbook activities place emphasis on meaning-focused activities (communication and message transmission). These findings can be interpreted in both a positive and negative light. For example, the tests can be said to be objective and, thus, reliable and practical to score. To this extent, they can be described as fair, too (Hamp-Lyons, 2000; Shohamy, 2000). However, due to their limited use of meaning-focused tasks, the tests fail to maintain the necessary balance of vocabulary task types that

can ensure appropriate assessment and teaching of vocabulary learning. As such, the tests exhibit low content and construct validity, e.g. the tests partially reflect teaching practices and fail to fully measure students' language performance as described in the textbook. In addition, the fact that the vocabulary test tasks do not simulate real-life situations renders them of poor predictive validity. In other words, since the tests partially cater for students' real language needs (Nation, 2001; West, 2004), they are likely to exert harmful washforward effect on students' future language performance (Hughes, 1989; West, 2004). Finally, given the test's under-representativeness of test tasks, negative washback effect is likely to be exerted on teaching too. In other words, it is quite likely that teachers might supplement the textbook activities with task types included in the tests, running the risk of teaching to the test; hence, teachers' instructional choices are likely to be affected negatively by the requirements of the tests (Cheng & Curtis, 2004; Leung & Andrews, 2012).

Analysis of the test results showed that students scored higher in the vocabulary tests which included only language-focused test tasks. The lower scores that students achieved in the meaning-focused tests tasks indicate that students faced difficulties when encountering such tasks in their tests. Several possible reasons can explain students' low performance. For example, it might be the case that meaning-focused language tasks are more demanding for students. This might be one reason vocabulary use is assessed less frequently than vocabulary knowledge in the vocabulary tests of the new 5th grade EFL textbooks. Lower test scores could also be attributed to the frequency of exposure to such tasks in their instructional programme. It is also likely that teachers might not practice meaning-focus vocabulary tasks given the requirements of the test tasks included in the vocabulary test tasks. As Griffin (1992, cited in Nation, 2001) clearly states, 'learners score higher when the testing format matches the learning format' (ibid: 33).

The results from the teachers' questionnaire yielded very interesting results with regard to teaching and assessment practices of vocabulary learning which confirmed initial teachers' reactions. For example, despite their generally positive views towards real-life authentic meaning in learning and assessing vocabulary, most teachers are either negative or neutral towards the vocabulary tests of the textbook. Teachers are not particularly satisfied with the extent to which the test and teaching tasks of the textbook engage students in real-world use of vocabulary. Teachers also stressed that the tests do not correspond to their learners' level. As a result, they need to adapt these tests or design their own to meet their students' needs.

However, the questionnaire yielded an amount of contradiction in teachers' responses. For example, even though teachers believe that vocabulary materials and tests are beyond their students' language level, most teachers believe that the tests assess vocabulary previously practised in class. Furthermore, even though previous task analysis revealed a weak relationship between vocabulary test tasks and textbook activity types, the majority of the teachers believe that the vocabulary test tasks are similar to the textbook activities. In the same vein, despite their preference towards balanced assessment practices between meaning- and language-focused activities, teachers tend to prefer to design test tasks which focus on the assessment of specific vocabulary items. To the teachers, language-oriented tasks are more practical to design and score and can offer useful feedback on students' vocabulary knowledge. Therefore, in theory teachers seem to encourage vocabulary assessment and teaching both at sentence level and in larger context. However, even if they hold these beliefs, when it comes to their actual assessment practices, they prefer language-focused test tasks. It is obvious that teachers' beliefs are opposed to their actual assessment procedures. Studies have shown that teachers' general beliefs on instruction as compared to

actual teaching are often incompatible (Foss & Kleinsasser, 2001; Kennedy, 1996; Zeichner & Tabachnick, 1981, cited in Bartels, 2005). This means that teachers tend to opt for the ideal in their beliefs, namely targeting towards overall language development through both real-life and language-focused tasks, whereas they seek more practical solutions in their assessment and instructional practices, in this case by designing discrete-point language tasks. This raises several questions with regard to the extent to which teachers have the skills or knowledge to evaluate textbook materials systematically or use them in the ways that are expected by textbook writers. This points to a lack of sufficient training in assessment and language pedagogy overall that could provide teachers with the necessary knowledge of crucial assessment concepts.

To conclude, given the evidence gathered in the present study, it is important to stress, first of all, that the present textbook-based vocabulary tests can be used to collect useful feedback about students' performance in specific areas of vocabulary learning. The degree of beneficial or harmful impact of these tests on teachers and students depends on teachers' instructional decisions and assessment practices. Therefore, teacher education in materials evaluation, design and use is essential. In the final section, recommendations are made to all stakeholders involved either directly or indirectly in the present study.

6. Limitations and Implications

This study investigated the nature and influence of textbook-based vocabulary tests on teaching and learning through a variety of methods. Even though the samples chosen are not extensive (mainly due to time limitations and other practical restrictions), nevertheless they were representative of the textbook material and teacher and student population. In addition, there was an effort to enhance research credibility by triangulating the sources of data that have led to interesting tendencies. For example, we feel that the detailed analysis of the materials yielded interesting results and can lead to several pedagogical and research implications. These will be discussed below in the form of several suggestions which are addressed to several key participants identified in the present study. However, they are not meant to be exhaustive or definitive. Also the practicality and effectiveness of each suggestion merits further exploration and research.

To begin with, suggestions can be made to those who produce textbook materials. Textbook writers/publishers need to follow a sound theoretical and practical framework in their approach to textbook material design (including textbook tests). For example, their textbook materials need to exhibit a comprehensive and accurate coverage and balance of language skills/areas and task/test types through appropriate teaching/learning techniques linked to the actual rather than perceived level of the students. This can be achieved through rigorous needs analysis which will also help record additional practicalities and constraints in the system. Also material writers need to sequence their materials in terms of difficulty. Given the importance of textbook use in language teaching and learning, textbook designers should also provide EFL teachers with supplementary guidelines concerning assessment procedures as well as with additional graded testing material that can meet the needs of students at different levels of ability.

Materials writers and publishers also need to pilot textbook materials (test tasks included) before introducing them to teachers and students to make sure that the materials are at the level of the students and can methodologically cater for the needs of teachers, too. In addition, publishing houses should also have mechanisms that monitor and support teachers

in the way they operationalise learning objectives through their textbook materials. Finally, even though there is no imperative for the textbook publishing industry to do so, we suggest that, in the broader discussion of ethics and 'good practice' in the field, writers/publishers should not only have a concern for the design of their textbook materials but also provide empirical evidence of the link between their textbooks and the CEFR levels.

Suggestions can also be made for teacher trainers (also addressed to curriculum planners/textbook designers). For example, there is a need for teacher educator programmes to include components that familiarise (pre- and in-service) EFL teachers with the rationale and principles behind various instructional and assessment techniques in textbook materials and train teachers in ways of using them effectively. Training teachers into the use of textbooks is a very important issue as the teacher variable plays an important role in the learning process (see also Tsagari, 2009). Also training programmes need to develop teachers' skills and strengthen their self-confidence in adapting the existing textbook materials or design their own to their students' advantage. Such training initiatives should also clarify misconceptions with regard to the intentions of material writers as these are reflected in the instructional and assessment materials of the textbook.

7. Recommendations for future research

Several practical recommendations could be made to future language testing and teaching researchers interested in conducting research in the field of materials design and evaluation as the influence of teaching materials in EFL courses merits further research. Research into the area could follow a number of directions.

For example, a first step towards understanding the nature and impact of teacher-made tests on teaching and learning and teacher effectiveness would be to collect teachers' vocabulary test samples and student scores based on such tests. In addition, further studies into materials design could look into the ways in which (experienced and non-experienced) teachers use textbook-based test tasks, e.g., the use or adaptation of such tests, their impact on student learning, the way results are reported, etc or the way they structure their lessons when using textbook materials, e.g. in terms of instructional time, parts of the materials they omit, etc. In learning more about the relationship between teacher implicit theories and practices as these are reflected in textbook use, feedback can be provided to teacher educators, materials writers, and most importantly, assistance to pre- and in-service teachers.

It would also be valuable to contact textbook writers and publishers (also discussed in Tsagari, 2009), after the analysis of textbook materials, to explore their rationale in designing textbook materials the way they do and the constraints they face. It may be valuable to look more closely at what exactly material writers think and do when they design instructional and test material using, for instance, through verbal protocols (see also Johnson, 2000; Johnson, 2003) or interviews (Tsagari and Sifakis, 2012). Given the increasing size of the textbook industry, it would also be important to determine whether textbooks with their accompanying materials, e.g. teachers' books, workbooks, test booklets, etc would have any real effect on improving learning and raising students' test scores in language tests (internal or external to the textbook).

To conclude, the views of publishers/textbook writers, those contributing to the professional literature, teacher educators and actual teachers' practice of teaching materials can be

synthesized to establish a 'theory' of how teachers can best fulfil their roles vis-a-vis materials and learners. However, it is important to take into consideration students' perspective, too. In-depth exploration of students' opinions and reactions towards their textbook materials would be a worthwhile venture as it can create mutual understanding between materials producers and users.

References

- Alderson, J. C. (2000). *Assessing reading*. Cambridge: Cambridge University Press.
- Alderson, J. C. & Hughes, A. (1979/81). 'Issues in language testing', *ELT documents: The British Council*, pp. 8-25.
- Alderson, J. C. & Wall, D. (1993). 'Does washback exist?', *Applied Linguistics*, 14/2: 115-129.
- Bachman, L. F. & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- Bartels, N. (2005). 'Researching applied linguistics in language teacher education.' In N. Bartels (Ed.), *Applied linguistics and language teacher education*. New York: Springer Science & Business Media, Inc, pp. 1-26.
- Breen, M. (1989). 'The evaluation cycle for language learning tasks'. In R. K. Johnson (Eds.), *The second language curriculum*. Cambridge: Cambridge University Press, pp.187-206.
- Bülent, A. & Stoller, F. L. (2005). 'Maximizing the Benefits of Project Work in Foreign Language Classrooms', *English Teaching Forum*, 43/4: 10-21.
- Burns, R. B. (1990). *Introduction to research methods*. London: Sage Publications.
- Byrd, P. (2001). 'Textbooks: evaluation for selection and analysis for implementation'. In M. Celce-Murcia (Ed.), *Teaching English as a second or foreign language*. Boston: Heinle & Heinle, pp. 415-427.
- Carr, N. T. (2008). 'Using Microsoft Excel to calculate descriptive statistics and create graphs'. *Language Assessment Quarterly*, 5/1: 43-62.
- Cheng, L. & Curtis, A. (2004). 'Washback or backwash: a review of the impact of testing on teaching and learning'. In L. Cheng, Y. Watanabe & A. Curtis (Eds.), *Washback in language testing: research contexts and methods*. New Jersey: Lawrence Erlbaum Associates, pp. 3-18.
- Council of Europe (2001). *Common European framework of references for languages: learning, teaching, assessment*. Cambridge: Cambridge University Press.
- Cunningsworth, A. (1995). *Choosing your coursebook*. Oxford: Heinemann.
- Cunningsworth, A. (1984). *Evaluating and selecting EFL teaching materials*. Oxford: Heinemann.
- Dickins, P. R. & Germaine, K. (1992). *Evaluation*. Oxford: Oxford University Press.
- Ellis, R. (1998). 'The evaluation of communicative tasks'. In B. Tomlinson (Ed.), *Materials development in language teaching*. Cambridge: Cambridge University Press, pp. 217-238.
- Guariento, W. & Morley, J. (2001). 'Text and task authenticity in the EFL classroom'. *ELT Journal*, 55/4: 347-353.
- Hamp-Lyons, L. (2000). 'Fairness in language testing'. In A. J. Kunan (Ed.), *Studies in language testing 9: fairness and validation in language assessment*. Cambridge: Cambridge University Press, pp. 30-34.
- Harmer, J. (2001). *The practice of English language teaching* (3rd ed.). Harlow: Longman.

- Harris, M. & Mc Cann, P. (1994). *Handbooks for the English classroom: assessment*. Oxford: Heinemann.
- Harwood, N. (2010). *English language teaching materials: theory and practice*. Cambridge: Cambridge University Press.
- Heaton, J. B. (1990). *Classroom testing*. Harlow: Longman.
- Hedge, T. (2000). *Teaching and Learning in the language classroom*. Oxford: Oxford University Press.
- Herman, J. L., Aschbacher P. R. & Winters L. (1992). *A practical guide to alternative assessment*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Hill, M. (2004). 'Teaching speaking: applying theory to practice'. In N. Sifakis, A. Georgountzou & M. Hill, *Language learning skills and materials (oracy and literacy): teaching listening and speaking, vol. 1*. Patras: Hellenic Open University, pp. 309-402.
- Hughes, A. (1989). *Testing for language teachers (2nd ed.)*. Cambridge: Cambridge University Press.
- Hulstijn, J. H. (2001). 'Intentional and incidental second language vocabulary learning: a reappraisal of elaboration, rehearsal and automaticity'. In P. Robinson (Ed.), *Cognition and second language instruction*. Cambridge: Cambridge University Press, pp. 258-286.
- Johnson, K. (2000). 'What task designers do'. *Language Teaching Research*, 4/3: 301-321.
- Johnson, K. (2003). *Designing language teaching tasks*. Basingstoke: Palgrave.
- Κολοβού, Ε. Κ. & Κρασιώτου, Α. (2009). *Αγγλικά Ε' δημοτικού: βιβλίο μαθητή*. Αθήνα: ΟΕΔΒ.
- Leung, C. Y., & Andrews, S. (2012). 'The mediating role of textbooks in high-stakes assessment reform'. *ELT Journal*, 66/3: 356-365.
- Littlejohn, A. (1998). 'The analysis of language teaching materials: inside the Trojan Horse'. In B. Tomlinson (Ed.), *Materials development in language teaching*. Cambridge: Cambridge University Press, pp. 190-216.
- Luoma, S. (2004). *Assessing speaking*. Cambridge: Cambridge University Press.
- Mc Donough, J. & Shaw, C. (1993) *Materials and methods in ELT*. Oxford: Blackwell.
- Mc Grath (2002). *Materials evaluation and design for language teaching*. Edinburgh: Edinburgh University Press.
- Miekley, J. (2005). 'ESL textbook evaluation checklist'. *The Reading Matrix*, 5/2: 1-9. Retrieved on 16.12.2012, from http://www.readingmatrix.com/reading_projects/miekley/project.pdf.
- Mishan, F. (2010). 'Withstanding washback: thinking outside the box in materials development'. In B. Tomlinson & H. Masuhara (Eds.), *Research for Materials development in language learning: evidence for best practice*. India: Replika press, pp. 353-365.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nation, P. & Newton, J. (1997). 'Teaching vocabulary'. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition*. Cambridge: Cambridge University Press, pp. 238-254.
- Niles R. (n.d) 'Standard deviation'. Retrieved on 16.12.2012 from <http://www.robertniles.com/stats/stdev.shtml>.
- Oppenheim, A. N. (1992). *Questionnaire design, interviewing and attitude measurement*. London: Bittles Ltd.
- Pavlou, A. (2010). *Assessing vocabulary knowledge in primary education: an adaptation of the 5th class self-assessment tests*. Unpublished MA dissertation. School of Humanities, Hellenic Open University, Patras, Greece.
- Peacock, M. (1997). 'The effect of authentic materials on the motivation of EFL learners'. *ELT Journal*, 51/2: 144-156.

- Pedagogical Institute (2011). *Common syllabus of foreign languages*. Retrieved on 16.12.2012, from <http://digitalschool.minedu.gov.gr> and /or <http://www.pi-schools.gr>
- Pedagogical Institute (2003) *Cross-curricular thematic framework*. Retrieved on 16.12.2012, from <http://www.pi-schools.gr>
- Purpura, J. E. (2004). *Assessing grammar*. Cambridge: Cambridge University Press.
- Π.Δ. 8/1995 'Η Αξιολόγηση Μαθητών του Δημοτικού Σχολείου'. *Εφημερίδα της Κυβερνήσεως της Ελληνικής Δημοκρατίας*, ΦΕΚ 3, Τεύχος Πρώτο, pp. 22-24.
- Read, J & Chapelle, C.A. (2001). 'A framework for second language vocabulary Assessment'. *Language Testing*, 18/1: 1-32.
- Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.
- Read, J. (2004). 'Plumbing the depths: how should the construct of vocabulary knowledge be defined?' In P. Bogaarts & B. Laufer (Eds.), *Vocabulary in a second language*. Amsterdam: John Benjamins Publishing Company, pp. 209-224.
- Schoonen, R. & Verhallen, M. (2008). 'The assessment of deep word knowledge in young first and second language learners'. *Language Testing*, 25/2: 211-236.
- Sheldon, L. E. (1987). *ELT textbooks and materials: problems in evaluation and development*. Oxford: Modern English Publications.
- Shohamy, E. (2000). 'Fairness in language testing'. In A. J. Kunan (Ed.), *Studies in language testing 9: fairness and validation in language assessment*. Cambridge: Cambridge University Press, pp. 15-19.
- Thrasher, R. (2000). 'Test theory and test design'. *Language Testing*, International Christian University. Retrieved on 16.12.2012, from <http://subsite.icu.ac.jp/people/randy/Test%20text%20grammar%20mcw.pdf>
- Tomlinson, B. & Masuhara, H. (2010). 'Published research on materials development for language learning'. In B. Tomlinson & H. Masuhara (Eds.), *Research for materials development in language learning: evidence for best practice*. India: Replika press, pp. 1-7.
- Tomlinson, B. (2012). 'State of the art article: materials development for language learning and teaching'. *Language Teaching*, 45/2: 143-179.
- Tsagari, D. & Sifakis, N. (2012). 'Course-book evaluation in Greek primary state EFL Schooling: lessons learned.' Paper presented at the 11th Conference for the European Society for the Study of English (ESSE), Bogazici University, Istanbul, Turkey.
- Tsagari, D. (2012). 'FCE-exam preparation discourses: insights from an ethnographic study'. *University of Cambridge ESOL Examinations Research Notes*, 49: 36-47. Retrieved on 16.12.2012, from http://www.cambridgeesol.org/rs_notes/rs_nts47.pdf
- Tsagari, D. (2009). *The Complexity of test washback*. Frankfurt am Main; Berlin: Peter Lang.
- Underhill, N. (1987). *Testing spoken language: a handbook of oral testing techniques*. Cambridge: Cambridge University Press.
- Ur, P. (1996). *A course in language teaching: practice and theory*. Cambridge: Cambridge: University Press.
- Watanabe, Y. (2004). 'Methodology in washback studies'. In L. Cheng, Y. Watanabe & A. Curtis (Eds.), *Washback in language testing: research contexts and methods*. New Jersey: Lawrence Erlbaum, pp. 19-36.
- Wesche, B. M. (1983). 'Communicative testing in a second language', *Modern Language Journal*, 67/1: 41-55.
- West, R. (2004). *Testing and assessment in language learning: principles and approaches to language testing, vol. 1*. Patras: Hellenic Open University.
- Williams, D. (1983). 'Developing criteria for textbook evaluation'. *ELT Journal*, 37/3: 251-255.

Dina Tsagari (tsagari.konstantia@ucy.ac.cy) holds a B.A. in English Language and Literature from the University of Athens, an M.A. in Linguistics for ELT and a PhD in Linguistics from the University of Lancaster, UK. She works as lecturer in Applied Linguistics/TEFL at the Department of English Studies, University of Cyprus. Her research interests include language testing and assessment, language teaching and learning, teacher education, distance learning, syllabus and materials design.

Athena Pavlou (pavlouath@yahoo.gr) holds a B.A in English Language and Literature from the National and Kapodistrian University of Athens. She holds an M.Ed. in Teaching English to Speakers of Other Languages (TESOL) from the Hellenic Open University. She works as an English language teacher in Primary Education. Her research interests include textbook material evaluation, teaching and assessment of young learners and teacher education.

APPENDIX I

Evaluation Checklist for the Analysis of Vocabulary

<i>A. Vocabulary Teaching and Learning (Textbook tasks)</i>	Yes	No	Comment
1. Is vocabulary taught explicitly through language-focused activities?			
2. Is vocabulary taught implicitly within the four skills?			
a. listening tasks?			
b. reading tasks?			
3. Is vocabulary practised through meaning-focused activities, namely:			
a. speaking tasks?			
b. writing tasks?			
4. Do these tasks meaningfully engage learners in real-life language use?			
<i>B. Vocabulary Assessment (Textbook tests)</i>			
5. Is vocabulary assessed in limited context through language-focused test tasks?			
6. Is vocabulary assessed within a larger construct through:			
a. listening tasks?			
b. reading tasks?			
c. speaking tasks?			
d. writing tasks?			
6. Are the vocabulary tests relevant to the students' level?			
7. Do the vocabulary test tasks meaningfully engage learners in real-life communication?			
8. Do the vocabulary test tasks reflect the teaching practices in class?			
9. Do the vocabulary tests shed light on students' future language performance?			

APPENDIX II

Teachers' Questionnaire

Section A

Part 1: Background Information

Please tick (√) the appropriate option.

Gender:

Male [] Female []

Teaching experience:

1-5 years [] 6-10 years [] 11-15 years [] 16-20 years [] over 21 years []

Teaching experience in State Primary Education:

1-5 years [] 6-10 years [] 11-15 years [] 16-20 years [] over 21 years []

Further studies other than your University degree

Master's in..... PhD in.....
Other (please specify).....

Part 2: Vocabulary Instruction and Acquisition

Please read the statements and choose a number from 1 to 5 as specified below.

1. I strongly agree, 2. I agree, 3. neutral, 4. I disagree, 5. I strongly disagree

1. Vocabulary should be learned through systematic vocabulary teaching (e.g. memorization of wordlists with their L1 translation or definition). []
2. Vocabulary should be learned in context focusing on the message to be transmitted. []
3. Vocabulary should be practised through vocabulary exercises such as matching words with definitions, gap filling, categorizing etc. []

Part 3: Vocabulary Assessment

Please read the statements and choose a number from 1 to 5 as specified below.

1. I strongly agree, 2. I agree, 3. neutral, 4. I disagree, 5. I strongly disagree

4. Vocabulary should be assessed ...
 - 4.1. as isolated words. []
 - 4.2. at sentence level. []

4.3 through meaningful tasks (e.g. information transfer, role playing, letter writing etc.) []

4.4 If other, please specify.....

Part 4: Textbook Material

Please read the statements and choose a number from 1 to 5 as specified below.

1. I strongly agree, 2. I agree, 3. neutral, 4. I disagree, 5. I strongly disagree

Pupil's book

5. Vocabulary covered in the textbook is appropriate for my learners' current level. []

6. The work done on vocabulary through reading, listening, speaking and writing is meaningful and engages learners in real-life communication. []

Vocabulary tests included in the Pupil's book and Teacher's book

7. The vocabulary tests are relevant to my students' level. []

8. The vocabulary tests assess the words and/or phrases students have practised in class. []

9. The vocabulary test tasks are similar to the textbook activities. []

10. The vocabulary test tasks are meaningful engaging learners in real-world language use. []

Part 5: Use of Textbook Tests

Please read the statements and choose a number from 1 to 5 as specified below.

1. I strongly agree, 2. I agree, 3. neutral, 4. I disagree, 5. I strongly disagree

11. I use the following methods to assess my students' vocabulary knowledge and use:

a. the Self-assessment tests []

b. the Progress tests []

c. the Final Progress test []

d. If other, please specify.....

.....

12. I employ these tests to assess my students' vocabulary knowledge and use because...

a. I can collect feedback concerning my students' progress and

- organize remedial work (revision). []
- b. They are quick and practical to score. []
- c. They are consistent in their scoring (anyone can correct them at any time and produce the same results) []
- d. They measure what they are supposed to measure, namely my students' vocabulary knowledge (the words and/or phrases they know). []
- e. They measure learners' future language performance (real language use other than the classroom). []
- f. They reflect the vocabulary learners have been taught in class. []
- g. Other (please specify).....

Section B

Free Response Questions

Please complete the statements

- 13. I adapt the vocabulary tests because.....
.....
.....
 - 14. When I adapt the vocabulary test tasks (please specify how you adapt the tasks).....
.....
.....
 - 15. I design my own vocabulary tests because.....
.....
.....
 - 16. The vocabulary test tasks I prefer to design are.....
.....
because.....
.....
 - 17. Please feel free to express any other concerns about the vocabulary component of the textbook tests.
.....
.....
- Thank you for your co operation.