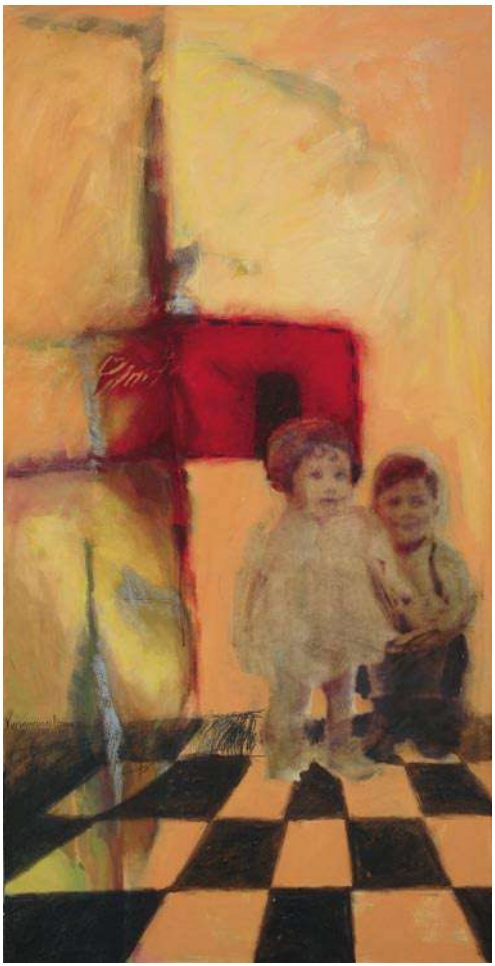

Líneas de fuego, 2004
Collage sobre lienzo, 200 x 100 cm

Washback of the ICFES Exam: A Case Study of Two Schools in the Departamento del Atlántico*[1]

Mg. Norma Barletta**
Mg. Ofelia May***

This study explores the impact of the compulsory state examinations at the end of secondary schooling (the ICFES exams) on English teaching practices in Colombia, since these examinations were modified in 2000. More specifically, we describe the washback effect of the English Test in two public schools. The washback effect is discussed, and results presented that were obtained through the triangulation of information and sources.

Key words: ICFES exam, washback effect, dimensions of washback, English language teaching

Este estudio explora el impacto que tienen las pruebas estatales de conocimiento de nivel secundario (el Examen ICFES) en la enseñanza del inglés en Colombia desde su modificación en el año 2000. Concretamente, se describe y se discute el efecto de rebote (o efecto *washback*) de la Prueba de Inglés en dos colegios públicos y se presentan resultados obtenidos a través de la triangulación de los datos y las fuentes de información.

Palabras clave: Examen ICFES, efecto de rebote, enseñanza del inglés

Cette étude décrit l'impact qu'ont eu les examens d'état de connaissance du niveau secondaire (examens ICFES) dans l'enseignement de l'anglais en Colombie depuis leurs modifications en 2000. Nous avons de façon concrète décrit et discuté l'effet de reflux (effet *washback*) de l'épreuve d'Anglais dans deux collèges publics et nous avons présenté les résultats obtenus grâce à la triangulation des données et des sources de l'information.

Mots clé: examen ICFES, effet de reflux, enseignement de l'anglais

---

* Recibido: 28-03-06 / Aceptado: 31-07-06

1 This paper reports some of the findings of the study "Relación entre las prácticas pedagógicas en el área de inglés y el desarrollo de la competencia comunicativa en esta lengua en colegios del sector público de Barranquilla y el Atlántico" funded by the Dirección de Investigaciones y Proyectos de la Universidad del Norte. The authors belong to the Grupo Lenguaje y Educación at the Departamento de Lenguas, of Universidad del norte, Barranquilla. Some of these findings were presented at the IV Round Table in Second Language Acquisition in Tucson, Arizona in February, 2005.

Introduction

In Colombia, the study of a foreign language is perceived as necessary due to the internationalization of the economy, multiculturalism as well as the scientific and technological development of our age. Different initiatives have been started to promote the teaching and learning of foreign languages among the school-aged population. One such effort is associated with the introduction of a Foreign Language Test as an obligatory component of the State Examination for the Admission into Higher Education. In the year 2000, the ICFES, the governmental agency responsible for the evaluation of the whole educational system in the country, established a new competence-based framework for this State Examination. Since the vast majority of children in Colombia are taught English in schools, English language teaching has been gaining considerable attention.

Due to the importance of the ICFES Examination as a criterion for admission into higher education, as a mechanism to inform students about their competencies in specific areas, as a support in the processes of self-evaluation and improvement of educational institutions, and as an instrument and basis both for research work and for granting certain educational benefits (ICFES, n.d., chap. 5), it is essential to identify its effects.

It is natural to suppose that the introduction of the Foreign Language Test has had some repercussion, since this subject is, for the first time in the history of education in the country, considered as important as any other. So, the following questions arise: Has the National Examination had any effect on the teaching and learning of English? If so, what kind of effect? How is the examination reflected in the classroom? These are the questions that this paper will attempt to address.

In the literature about language testing, the property of the test that concerns its effects is commonly known as washback. An exploration of the large body of research in this field is needed to establish our theoretical framework.

Theoretical framework

Washback, sometimes referred to as backwash (Biggs, 1995, 1996 in Cheng, 2000), can be generally understood as the effect of an examination in the teaching

and learning (Cheng, 2003, Chen, 2002, Hughes, 2003), but not all scholars have agreed as to its definition. Alderson and Wall (1993) restricted the use of the term 'washback' to "classroom behaviors of teachers and learners rather than the nature of printed and other pedagogic material" (p. 118). They would also consider washback as what teachers and learners do that "they would not necessarily otherwise do" (p. 117). Messick (1996) states that in order to be considered washback, good or bad teaching has to be "evidentially linked to the introduction and use of the test" (p. 16). Moreover, Wall (1997 in Cheng and Curtis, 2004:4) makes a clear distinction between washback and test impact. The latter would refer to the effect of a test on "individuals, policies or practices, within the classroom, the school, the educational system or society as a whole".

Other researchers (Andrews et al., 2002) do not make that distinction and they consider that narrow and wider effects can be included under the term washback. For the purposes of this study, washback will be understood in the wider sense that is, including what some scholars call 'impact'.

Messick (1996) claims that if a test is deficient because it has construct underrepresentation, then good teaching cannot be considered an effect of the test, and conversely, if a test is construct-validated, poor teaching, cannot be associated with the test. Only valid tests (which minimize construct underrepresentation and construct irrelevancies), can increase the likelihood of positive washback

Types of Washback

Washback has been shown to act in multiple directions, or, as Watanabe (2004) puts it "a highly complex rather than a monolithic phenomenon" (p.19). Alderson and Wall (1993), for example, distance themselves from a simplistic assumption about the way a test can influence behaviors. Therefore, they developed 15 washback hypotheses according to what is influenced, from teaching to attitudes and the number of teachers and learners affected by a test. Which hypotheses will be put forward depends on the nature of the test, the educational context, and the nature of the decisions resulting from test outcomes. In fact, there seem to be a number of variables in society, education, and schools that determine how washback will appear.

Hughes (1993 in Bailey, 1996) also attempts to operationalize the different workings of the washback effect. In his trichotomy model he distinguishes between participants, processes, and products. Participants in his model include students, teachers, administrators, materials designers and publishers. Processes are the actions undertaken by the participants and that may affect learning. Products refer to "what is learned (facts, skills, etc.) and the quality of the learning (fluency, etc)" (Hughes, 1993, p. 2 cited in Bailey, 1996, p. 262).

One more recent attempt at disentangling the complexity of washback was done by Watanabe (2004), who conceptualizes it in terms of dimension, aspects of learning and teaching that may be influenced by the examination, and the factors mediating the process of washback being generated. Watanabe's dimensions are specificity (to what extent is the specific type of exam or a specific component/aspect of the test that brings about changes in the teaching and learning), intensity (strong or weak washback depending on how much of what happens in the class or how many students or teachers are affected by the exam), length (the period of time an exam is said to influence the teaching and learning), intentionality (whether the social consequences of test interpretation and use were intended or unintended), and value (the value judgment -positive or negative washback- associated with a test by a certain audience).

The aspects of learning and teaching that may be influenced by the exam, Watanabe (2004) argues, can be described in terms of the three aspects considered in Hughes's trichotomy model (1993, see above) and the 15 washback hypotheses Alderson and Walls (1993) put forward (see above).

Finally, Watanabe (2004) provides a number of factors mediating washback: test factors, such as purpose, content, method of the test), importance of the test (status, consequences of the test), personal factors (beliefs, training of participants), and contextual factors (school, town, society where the test is used).

Usually researchers focus on one aspect or type of washback. In Alderson and Wall's study in Sri Lanka (Wall, 1996), the introduction of a test of English as a foreign language proved to produce faster changes in the content of teaching than changes in teaching methodology. Cheng (1997), in the preliminary results of a study of the washback effect of the exam taken by students in Hong Kong

secondary schools, reports that washback effect 'works quickly and efficiently in bringing about changes in teaching materials […] and slowly and reluctantly and with difficulties in methodology (p.1). Cheng introduces the term 'washback intensity' to refer to the "degree of washback effect in an area or a number of areas that an examination affects most" (p.7).

Quesada (2001), in a study carried out in high schools in Costa Rica, found that pressure from the test seemed to explain why teachers and students panic and abandon the syllabus for exclusive training for the examination.

Andrews et al (2002) found that the impact of a test can be immediate or delayed. According to these researchers, washback seems to be associated primarily with 'high-stakes' tests, that is, tests used for taking important decisions that affect different sectors. One such decision is determining who receives admission into further education or employment opportunities (Chapman and Snyder, 2000). Madaus (1990 in Shohamy, Donitza-Shmitdt & Ferman, 1996) identifies as 'high' such situations when admission, promotion, placement or graduation are dependent on the test.

Another aspect that has been studied is whether the test has been used as lever for change (Pearson 1988 in Cheng, 1997), so everything, from textbooks to staff, works to achieve better scores. Cheng (2000) reports on the introduction of tests to improve teaching and learning. In some countries these tests can be considered "the engine for implementing educational policy" (Petrie, 1987:175 in Cheng, 2000: 6).

Factors Affecting Washback

Education is a complex phenomenon. Therefore it is simplistic to believe that a test can bring about all desired changes in teaching and learning. Saif (2000) argues that an analysis of the needs and objectives of learners and educational system should be carried out as a starting point for the research in washback. Shohamy et al. (1996) consider factors like the status of the subject-matter tested, the nature of the test, and the use to which the test scores are put. (p. 300). Wall (1996) provides a list of factors which might have prevented the examination

in Sri Lanka from providing an effective 'lever for change'. These ranged from teachers' factors including lack of understanding of exam, resistance to change and content, to more macro factors like gap between test designers and teachers, lack of well trained teachers, overload of teachers, etc. In addition, according to Andrews et al (2002), the innovating effect of a testing innovation is affected by the teachers and their interpretation of the innovation, which may be different from that of the test conceivers. Another variable can be the published materials in use (Andrews et al., 2002).

The Study

The general objective of the present study is to describe the washback effect of the ICFES English Language Test in the teaching of English in two public schools: one in a 10th grade classroom in Barranquilla (context 1) and the other in an 8th grade class (context 2) in Puerto Colombia, a town situated to the north of Barranquilla. Both are recognized as important traditional schools in the Departament, neither predominantly privileged nor particularly poor or disadvantaged. We used these specific classes because the teachers and their students agreed to participate, and allowed our presence in their classrooms. As with all case studies (Merril, 1998), the idea was to do "an intensive, holistic description and analysis" (p. 28) of the nature of the phenomenon in the particular contexts rather than attempting at external validity.

Because of our interest in participants, processes and outcomes, we needed to obtain information from different sources and informants: documents, practices within the classroom, teachers' and learners' opinions and judgments. This information was then triangulated both in method (researchers' views contrasted) and in source (interviews, tests, observations).

With regards to documents, the framework for the Foreign Language Tests used by the ICFES as well as the actual 2003 and 2004 tests were scrutinized. This information was then compared with the classroom practices recorded from non-participant observations. Five English lessons were observed in the development of a complete unit of the syllabus within each school. Two observers sitting one at the front and one at the back were present in each

classroom. They audio recorded events while taking notes and made copies of the tests applied and graded by the teacher. Recordings (observations and interviews) were completely transcribed.

Specific items were analyzed within each observation: activities carried out, learners' responses, actions undertaken by each teacher to evaluate students' learning, feedback from the teacher and competencies assumed to be developed in each type of activity, following the Foreign Language Framework (See appendix 1)

Short informal interviews after most classes and a formal interview at the end of the observation period were arranged. Additionally, six students from each context were interviewed. Questions to teachers included their perceptions and satisfaction with respect to English instruction in their contexts (goals, topics, materials, rationale of their courses); perceived importance of English outside the classroom; their understanding of communicative competences; importance, objectives and implications of the new ICFES exam. Students were asked about their needs in English and its role in their lives, the importance of the ICFES exam and their perceptions of their learning of the foreign language in class. A content analysis was applied to the interviews with the aim of obtaining information about the participant's mediating effect on the washback of the exam.

For the purpose of the analysis, the trichotomy proposed by Hughes (1993 cited in Bailey, 1996) was found useful in order to distinguish between the different instances affected by the test: participants, processes and outcomes. In the discussion section we follow the framework suggested by Watanabe (2004), to give account of a number of dimensions of the washback: intensity, intentionality and value, as well as of the analysis of the factors that mediate the washback.

Results

In this section, we will initially give a brief description of the characteristics of the ICFES test and then focus our attention on a portrait of the contexts involved.

The Foreign Language Test

An insight into the test was obtained through an analysis of the official document of the Foreign Language Test. In its framework, special attention is given to Canale and Swain's (1980) definition of communicative competence with the distinction between Grammatical, Sociolinguistic, Discursive and Strategic competences. The concept of communicative competence is operationalized for evaluative purposes through Bachman's (1990) and Bachman and Palmer's (1996) notions of linguistic and pragmatic competencies. Linguistic competence includes grammatical competence, textual competence and textual coherence. Grammatical competence, according to Canale (1983, cited in ICFES, n.d.), implies mastering the linguistic code, that is, semantic and morphosyntactic aspects. Textual competence implies the identification of cohesive elements and the rhetorical organization of a text. Textual coherence is understood as the ability to structure discourse, logical sequences of ideas, and construct meaning from a text. It is these three competencies that all students finishing high school are evaluated on for the State Exam. Pragmatic competence, which accounts for the knowledge of the use of the language is tested only for those students that choose Foreign Language as an area in which they desire to be more thoroughly evaluated.

A summary of the types of item ("contextos de evaluación") with their distribution in the 2003 and 2004 versions of the test, as well as the types of competence evaluated is included in Table 1.

Table 1. English Test item types

Contexts of evaluation	# items	Competence
Identification of grammatically correct construction in a given context.	5	Grammatical
Interpretation of graphs.	5	Grammatical
Cloze test where missing words are verbs, prepositions, connectors, adjectives, nouns, etc.	6	Textual
Organization of paragraphs.	2	Textual
Dialogue completion (students identify the missing turn).	3	Textual coherence
Situation comprehension	2	Textual coherence
Text comprehension	12	Textual/Textual coherence.

Out of the 35 items of the exam, ten exclusively test grammatical competence: knowledge of discrete vocabulary items or grammar structure. The cloze test demands knowledge of correct verb forms and prepositions in given sentences. In some cases it also requires understanding of relationships between adjacent sentences or clauses, or interpreting a reference. The dialogue completion task requires students to read the line after the blank of the missing turn in order to answer. The 'organizing paragraphs' items ask students to put sentences or chunks of sentences in a logical order, usually requiring knowledge of how discourse pieces are connected. The 'text comprehension' items include simplified texts and questions that aim at testing global understanding (the general idea of the text, the intention of the text, for example) but also discrete vocabulary items, paraphrasing of phrases and local reference interpretation.

Participants: Teachers

Information about how the participants' mediate the washback effect was obtained through the teachers' description of their teaching situation, their goals for their classes, their individual concept of communicative competence and the ICFES exam, and their decisions regarding syllabus and use of materials. The findings show different degrees of institutional support and teachers' autonomy in the classrooms (see Appendix 1).

The approach to teaching in context 1 is more traditional and text-bound. The teacher has a very vague notion of communicative competence but she considers that grammar is extremely important to speak and write. She thinks there is pressure from the government for teachers to undertake a new kind of teaching although she was not able to define the nature of the change. Even though she expects the textbook to achieve most of its objectives, she believes the book is not completely adequate for her students.

In spite of the teacher's enthusiasm in context 2, she regrets the limited institutional support for the teaching of English. She relies on what she has learned at workshops and seminars and her knowledge of the context to decide what to teach. She has a pragmatic notion of communicative competence, but she also sets forth a broader educational aim (helping students define a 'project of life'). She believes by doing the right thing she will eventually help students

cope with the ICFES exam. Her job, as she puts it, is "to fulfill the students' expectations rather than those of the ICFES […] because how do they measure real performance in English? I think it is sort of incoherent".

Participants: Students

The learner's perception of the foreign language in their particular context and in society in general, as well as their opinions about an exam affect the general impact of a test. In the interviews students of both contexts reflected general awareness of the instrumental benefits of knowing English. They also acknowledged the importance of the ICFES English Test. However, in context 1 they have taken two preparatory tests though they did not report having received any specific feedback on it, or carrying out specific preparation for it. In context 2, these particular eighth graders perceived the exam as a distant requirement their teacher reminds them of every now and then, but with little or no practical consequence in their classrooms. Students in both contexts think English outside school is of limited use in their daily lives: perhaps for watching and reading the captions of movies, listening to songs, meeting an occasional foreigner, or, potentially, chatting on the internet.

Processes

Processes, according to Hughes (1993 in Bailey, 1996) refers to material development, syllabus design, changes in teaching methodology, and testing strategies, among others.

Context 1

As mentioned before, the syllabus of the course was textbook driven, though sometimes the teacher would either skip some of the activities or follow a different order. During the observations, the book was followed while it dealt with discrete points (vocabulary, guided writing). More extensive writing and project development was assigned as homework.

The analysis of the classroom observations (see Appendix 2) showed that from the seventeen activities developed during the time observed, nine of them were

directed towards the development of grammatical competence. Beyond that was a cloze exercise on quantifiers, two guided writing tasks where students were asked to use sequence connectors. Other activities, like oral presentations and dialogue dramatizations, which are potentially suitable for developing other competencies turned out to be rote memorization of dialogues. In other cases the presentations consisted of reading aloud texts that contained a number of mistakes. There was no evidence of the other students' following or understanding the presentations or dialogues. Most of the time the students were practicing rules in artificial exercises that focused on form, usually at sentence level, with predictable responses. Feedback on presentations and dialogues was very little, usually very general. For grammar exercises, the correct answer and sometimes the repetition of the rule were supplied.

During the observation time a test was administered. Table 2 shows that it resembled the ICFES examination in that it addressed the same competencies.

Table 2. Summary of In-class examination (Context # 1)

Type of item	Competence	Example	Percentage of correct answers
Matching	Textual	How many tomatoes do we need? a. Only a few b. Only a little	68
Blank filling	Grammatical	- Are there ____ eggs? - I eat a _____ fish.	39
Dialogue completion Using *how much/ how many/ a few*, etc.	Textual	1. How much water does Betty drink each day? _____ 2. _____ She can buy a lot of apples	31
Writing Describe recipe	Textual coherence	Students had to use sequence words: after that, first, next, finally. Apart from this, only blank lines were given.	42
Fill in blanks with the correct future form	Grammar	-Hi, Mike. What's new? -I'm making plans for the weekend (1.I, go) _____ to the Latin Music Fiesta on Friday.	10
Average	38		

Context 2

As mentioned above, the teacher does not follow one textbook, but occasionally uses photocopies from unreferenced textbooks. Her classes are not governed by a rigid, discrete grammatical itinerary but are focused on topics and their essential vocabulary. She is more concerned with keeping students engaged in using the language through mostly oral activities and games.

The analysis of the five observations showed that during these lessons the teacher set up sixteen activities. Nine of them were clearly focused on the development of lexical grammatical competence. Seven of these focused exclusively on the vocabulary related to the topic of occupations (matching exercises and games). The two other activities required students to construct simple sentences about a person's occupation. The teacher's feedback focused exclusively on form. Three other activities had the potential to address textual competence and textual coherence (learners dealing with texts or dialogues). However, she focused on ensuring students understood discrete lexical items. In an oral activity that required students to answer questions about their mothers' or fathers' jobs, the teachers' attention was centered on correcting pronunciation and providing needed words. All students were asked the same questions or they provided the information voluntarily in a sentence. There was never a follow up question or additional information offered. Spanish was often used by students.

In a truly communicative exercise, the teacher tried to elicit from the students their opinions about some jobs. Despite the interest this exercise arose, she never obtained anything beyond "Yes", "Good" or "Bad" in English plus all sorts of comments in Spanish.

Three written activities - "expressing future plans", "writing mini-dialogues about people's professions"- "writing about their future and that of their community"- undoubtedly required learners to go beyond lexical and grammatical competence, to focus on content and to provide answers that were not necessarily predictable. These proved difficult for the students. They kept asking questions to the teacher but during the class she never got to check on what

they wrote. Though the learners expressed no frustration, they depended on the teacher to provide vocabulary and structure and their language evidenced problems with the use of auxiliaries, pronouns, third person singular, and textual cohesion.

Regarding evaluation, no written test was given to the students. Students prepared mini-dialogues, but the teacher did not make corrections during, or give feedback after their performance. She did collect the compositions of the final activity "writing about their future and that of their community". She said she would correct them, give feedback and have students rewrite it. Students did not report that as a usual procedure.

Products

Context 1

The products in a washback study refer mainly to what is learned. Our interest is the development of the communicative competence. As can be seen from the results of the in-class test (Table 2), and from the analysis of the performance of the students during the class observations (see Appendix 1), students seem to be working exclusively towards the development of linguistic competence with greater emphasis on grammatical subcompetence. However, in the class exam they demonstrate more textual competence (48.5% average for the textual competence items) than grammatical competence (24.5% average for the two items that tested it). That is, they seem more competent dealing with connected texts, of two or more sentences, rather than with discrete grammatical items. Additionally, there is no trace of the development of any other subcompetence within the larger notion of communicative competence.

Context 2

The analysis of students' performance in class showed that they seem to be advancing mainly towards the development of lexical (e.g. vocabulary related to occupations and professions - observations 1-3) and pragmatic competences (e.g. asking and giving information about parents' jobs and speaking and writ-

ing about their future lives - observations 4-5). It was made apparent during the observations that they are more interested in being able to make sense from the language they are exposed to, and to be able to ask for and provide information about themselves than in using correct structures. When the performance of nine pairs was analyzed, it was noted that in general they were able to ask and answer the questions following the model given by the teacher. One pair showed a high level of competence when exchanging information about themselves with no grammatical or lexical errors. All the others made mistakes either in the use of auxiliaries to make questions *(are* or *do* instead of *does,* sometimes omission of auxiliary), in the choice of pronoun *(I am* instead of *He is or She is),* third-person singular –s, in pronunciation, use or articles, question word (*Where does your father do*?) or even in the coherence of the questions and answers (A: *What does our sister do?* – B: *She is a student*. A: *Where does he work?* – B: *He work in the restaurant*).

Discussion

The central question of this study is whether the English Test has had any washback effect on the teaching of English in the specific contexts of this study. We did not deliberately choose a context where we knew everybody was deeply concerned with the exam. As with all case studies (Merril, 1998) we went into the schools which welcomed us with no hope of aiming at generalizability but to obtain in depth understanding of what is happening with respect to the ICFES exam. Both schools can be considered average state-run schools, not particularly privileged, but certainly not especially poor or disadvantaged. The results of the study do not allow us to give a unique unequivocal answer as to what is the effect of the English Test. Three dimensions -specificity, intensity, and value (Watanabe, 2004)- will be addressed. We will not refer to length of the washback effect because the study was just a snapshot of each context and claims beyond that cannot be made. Intentionality of the washback will not be considered a relevant dimension for discussion since the focus is on what is actually happening in two specific contexts.

Specificity

This dimension refers to whether the effect of an exam can be linked to the existence of the test, or to one particular feature thereof. The test was meant as a step in a whole process of improving the quality of teaching in the country and as a measure to quantify the advances made by the schools since the introduction of English as an obligatory subject. Logistically and administratively, some schools are making efforts in providing more teachers and classrooms for the teaching of English in order to have fewer students in each class, allowing only qualified teachers to teach the subject, and providing support for in-service training. Teachers also vaguely discuss the notion of competencies and communicative competence as a new direction in the teaching of languages. All this seems to point to a general type of washback.

Though no pressure, panic or urgency was felt inside any of the two classrooms to work towards better performance in the examination, teachers and students reported awareness of the importance and the demands of the examination. Still, as expected, there are features which distinguish each setting. In context 1, for example, there is more specificity. Students take preparatory tests in order to score higher in the exam. A textbook providing ample practice in linguistic competence was introduced and most classroom time is devoted to this aspect as a result of a certain curriculum alignment with the test. Not surprisingly, the competences that are being developed are to a large extent limited to those tested in the ICFES exam. In context 2, the teacher demonstrates concern for the use of language for communicative purposes, she emphasizes the development of lexical competence and at times, reminds students of the exam ahead. Here, however, no written test is administered to learners and no text is followed. Textual competence and textual coherence are overlooked. There is the notion that grammar should not be the main concern (rather vocabulary) and that pleasure in using English will bring competence. The lack of specificity in context 2 may be due to the fact that in this context students still have three more years to prepare for the test, whereas in context 1 the exam is forthcoming in one year's time, so students need to get familiar with the type of questions and the specific competences tested in the exam. It seems, however, that it is the teacher's

personal beliefs and knowledge of her students that determine why these specific competences are not taught to eighth graders, why basic elements of textual competence are not addressed, and why lexical competence is so prevalent over grammatical.

It can be said that the test has contributed to strengthening the perception of the importance of English inside context 1 and among teachers. It could be argued that if there were no exam, the changes in context 1 would hardly have taken place, and in context 2, the teacher would not be so much focused on what she understands as communicative competence. Consequently, one can claim that the test has had a *general* positive washback, because it has prompted a number of changes towards improving the quality of teaching. This effect cannot be directly associated to any specific feature of the test, more with the fact that *there is* such a test and that its general goal is to test communicative competence, however it is that this notion is understood.

Intensity

In terms of intensity, the washback of a test can be weak or strong. While it can be said that the test has had some impact at administrative, logistic and planning level, the same is not felt inside the classrooms. The data do not show that the test determines everything that happens there or that it is substantially different from what was done before. Unfortunately, there is no reliable data as to how English was taught in these contexts before 2000 or 2003. Practices observed in context 2 are strikingly different from those observed in context 1.

According to Watanabe (2004), the intensity of the washback could be a function of the importance of the test: the higher the stakes, the stronger the intensity. In this study, the exam can nominally be considered a 'high-stakes' test, but it works differently for each of the participants. Schools see the ICFES test as the instrument whereby they are ranked among all the other schools in the country and the department. The teachers are also evaluated through the test because it would be particularly shameful for them if the scores in their area are lower than those obtained in other areas. However, the degree of staff effort is mediated by what the school considers is the expected performance

of the school, and what the teachers consider the general performance of the subject should be as compared with other subjects. So, institutions and teachers will react accordingly. Based on the results from this study, it can be said that the ICFES exam has had a stronger washback for the school and the teacher in context 1 than in context 2, where the teacher seemed to be less influenced and the curriculum does not necessarily revolve around the test's exercise typology. This could be related to the grade in which each of the teacher participants was teaching. The closer the class is to graduation, the stronger the washback effect. Another variable is the size and nature of the town. The larger the town and the number of higher education institutions, the stronger the washback effect is likely to be. Concrete decisions regarding goals, methodology, emphasis, activities, feedback and evaluation practices will depend on the teacher's theory of teaching and learning and her/his interpretation of the students' needs.

Likewise, a number of factors in the students can affect the intensity of the washback of the exam. Even if a school as an institution makes a large effort to help students obtain higher scores in the examinations, individual students may not consider the importance of the test, given that they are not expecting to get into higher education, so the test would not have any practical consequence for them. Mainly, private goals mediate the effect of the test.

Value

Value refers to whether the effect of the exam is positive or negative. In this dimension, the washback is the least clear. The best method is to follow Alderson's (1992 in Watanabe, 2004) suggestion of identifying the audience. Official entities (ICFES) use results of the test to measure the improvement of education. The exam serves that end, and apparently positive changes are being implemented in schools. On the other hand, students need to know what and how much they learn. A special reflection is due here.

Official documents state that the exam is meant to have students show their ability to go beyond applying the rules governing the language. It is supposed to give evidence of test takers' use of the language for communicative purposes.

However, the ICFES tests analyzed are limited to the lower or more restricted competences. Pragmatic subcompetencies are not included in the general test and that certainly has had its effect in the face validity of the exam. Both teachers in this study have the general impression that the test is about grammar and reading comprehension. This gets strongly reflected in the activities carried out in context 1. The teacher in context 2 disregards the lower competencies, except for lexical competence, and seems to aim at free functional use of the language, without her students being really prepared for it.

Thus, the teachers reflect the same kind of ambivalence of the theoretical framework of the test. In context 1 the teacher wants her students to be able to talk, to understand when talked to, to write and to understand written texts, but her focus is on form. In context 2 where apparently students have more opportunities for free expression, they are not given solid bases or enough feedback on their performance. The focus is discrete vocabulary items and communicating simple ideas. The students' oral production consists of single words, lots of Spanish, or very simple sentences with grammatical mistakes. In neither context, feedback provided by teachers is appropriate. In context 1, for instance, feedback on the few things that might have been communicative (presentations, writing of own recipes) was minimal, and most attention was paid to grammatical and pronunciation accuracy. In context 2, the teacher hardly ever gave feedback on students' oral performance, except for pronunciation. These findings in context 1 confirm Messick's (1996) claim that washback is related to the construct validity of the test. Since there is obvious under-representation of the concept of communicative competence, the washback can hardly be positive. In context 2, the teacher feels the exam does not reflect the true nature of communicative competence, so she opts for the other end, the pragmatic side, but with very little success.

Other factors could be argued to influence the value of the washback. In the particular case of context 1, the teacher was evidently grammar-oriented. Her personal beliefs about what is important obviously count when making decisions on textbook, activities, feedback and formal evaluations etc. She strongly believes that proper grammar knowledge is needed to cope with communicative activities. And this is especially negative because the data show that her

students were developing textual competence at a higher level than they were developing grammatical competence (see Table 2), and she might decide it is not worth continuing with the development of other competences until they do better at the grammatical level. In the case of the teacher in context 2, she is convinced that her major role is to make students aware of the importance of their participation in the construction of a more prosperous community. She does not seem to feel the pressure of an external party (The ICFES or the Ministry of Education as a surveillance entity). This lack of pressure gives her the liberty to implement a course, the emphasis of which is not only linguistic, but, and sometimes, above all, attitudinal and cultural. Perhaps, she has realized that only a few of these students will really take advantage of their knowledge of the language in the future. Perhaps she is trying to compensate for this neglected cultural appreciation which is absent not only in the ICFES exam but in the core curriculum of many schools. It is not implied, however, that events in this classroom show a lack of appreciation of the academic subculture. They simply show an awareness of the real value that cultural aspects may have within the school curriculum.

Results obtained from these contexts do not say much about what students can really do with the language, or whether they would be able to use it for authentic purposes. In neither context were they exposed to authentic language or non-academic interactions. Neither does the exam require it.

Table 3 summarizes our model of how washback works in the contexts and the type of washback that the different factors seem to be generating. The test produced general awareness of the importance of English, reduced class size and seems to have contributed to the generation of ideal goals in line with the communicative competence construct. These are in themselves part of the general positive washback effect, which was perceived here as 'strong' and 'positive'. However, since it is 'general' washback, as a factor for the outcome of the test in terms of learning, its effect has shown to be rather weak. What seems to be crucial, is the teachers and the decisions they make. These decisions, however, cannot evidentially be linked to the examination because nothing in the class or in the interviews can incontrovertibly show this direct relation, but there is, nevertheless, a strong correlation.

Table 3. How Washback Works

Value/ specificity	Factors mediating washback	Participants	Processes	Products
General Positive Washback	Macro – context: importance of English	School decisions: number and size of classes		
	Perception of English in the society: educational legislation	Teacher's awareness and personal efforts Student's awareness		
		Ideal goals students set for themselves		
Specific Negative Washback	Importance of exam in the context	Syllabus Specific objectives Classroom activities and interaction In-class evaluation practices and type of feedback		Students' level of communicative competence
	Teacher's beliefs about how language is learnt and what is communicative competence			

Conclusions and recommendations

This study evidences the complexity of the washback phenomenon and exemplifies a tentative proposal of qualitative research design to investigate it. The types of competence tested in the ICFES English Test strongly correlate with the ones actually developed in one of the classrooms observed. The orientation of both the test and the class goals is quite limiting, so students' level of competence is not surprisingly low. It seems then necessary to question the construct validity of the test and the decision made to test only what have been here described as 'low-level' competencies. Though issues of practicality could be one explanation, ideologies about what it is that Colombians need to know about English or what can be expected from our teachers and the teaching of English in our schools

may also be playing a role. It is then essential that the current exam be regarded only as a first step into a process of raising the level of expectations and demands with regard to the learning and teaching of English in the country.

The study shows that both teachers claim they are working towards the development of their students' communicative competence. Neither of them questions this as the goal to pursue but each has her own way of understanding this notion and different views regarding what is necessary to do in classroom to that end. As a result, the teacher in context 1, thinking she is compliant with the policies and instrumental in helping students develop communicative competence, actually focuses on the same restricted notion the exam promotes. Her class contrasts with the one in context 2, where there is an open disregard for the exam. Here the exam is regarded as non-congruent with the students' actual needs and interests in middle school, so the teacher exercises her autonomy to decide and implement what she considers is best for her pupils regardless of the demands of the exam. Though the notion of communicative competence is understood wider than in context 1, some misconceptions about communicative language teaching surface: the prevalence of vocabulary over grammar, the undesirable effect of negative feedback, (written) testing as unnecessary. It is the teacher's accumulated knowledge, training, experience and philosophy which seem to matter most in defining what happens in the classroom, at least in a context where the expected socio-economic activities of the students after finishing high school diminish the importance attached to the State Exam.

It is hoped that in the very near future the authorities decide to include pragmatic competence in the test, so that it gains face and construct validity. It is also essential that, given that what happens in the classrooms depends on teachers' decisions, they get extensive opportunities for professional development not only in relation to what communicative competence is, how it can be developed and tested, but also in ways to assess their contexts and their particular needs so as to make decisions in the classroom that are validated by theoretical and empirical findings. It is also hoped that more direct and authentic language and tasks are included in the test. If students are not asked to speak or write beyond recitation or mere copying, those practices are unlikely be felt as relevant and important in the classrooms. If feedback to performance is not given to students, errors are

likely to stay with them. It also seems essential that an analysis of the needs of English in our context be carried out, so that teaching and testing matches these needs and students feel real necessity to learn the language.

The English Test in its current form can also be seen as an initial step towards a long-term goal of raising the standards of communicative competence in the country, in which case the contexts under study seem to be very slowly going in the right direction. However, there is no evidence in this study as to whether what is being done currently is qualitatively different from what had been done before the exam was introduced. It is important then to keep track of the changes in the test and of the ways in which the schools and teachers react to them, for which longitudinal studies will be called.

References

Alderson, J. C., 1992, "Guidelines for the evaluation of language education". In J. C. Alderson & A. Beretta (Eds.), *Evaluating second language education* (pp.274-304). Cambridge, England: Cambridge University Press.

Alderson, J. C., & D. Wall, 1993, "Does washback exist?", *Applied Linguistics, 14* (115-129).

Andrews, S. et al., 2002, "Targeting washback – a case study", *System, 30 (*207-223).

Bachman, L. F., 1990, *Fundamental considerations in language testing*. Oxford: Oxford University Press.

Bachman, L. F., & A.S. Palmer, 1996, *Language testing in practice*. Oxford: Oxford University Press.

Bailey, D. M., 1996, "Working for washback: A review of the washback concept in language testing", *Language Testing, 13*(257-279).

Biggs, J., 1995, "Assumptions underlying new approaches to educational assessment". *Curriculum Forum, 4*(2), 1-22

Biggs, J., 1996, *Testing: To educate or to select? Education in Hong Kong at the cross-roads.* Hong Kong: Hong Kong Educational Publishing.

Canale, M. & M, Swain, 1980, "Theoretical bases to communicative approaches to communicative teaching and testing"*, Applied* Linguistics, 1 (1-47).

Chapman, D., & C, Snynder, 2000, "Can high stakes national testing improve instruction: Reexamining conventional wisdom", *International Journal of Educational Development, 20*, 457-474.

Chen, L., 2002, *Washback of a public exam on English teaching.* (ERIC Document Reproduction Service No. ED472167)

Chen, L., 2000, *Washback or backwash: A review of the impact of testing on teaching and learning.* (ERIC Opinion papers ED 442280)

Chen, L., 2003, "Looking at the impact of a public examination change on secondary classroom teaching: A Hong Kong case study", *Journal of Classroom Interaction, 38* (1), 1-10).

Cheng, L., & A. Curtis, 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.* Mahwah, N.J.: Lawrence Erlbaum.

Cheng, L., et al (eds.), 2004, *Washback in language testing: Research contexts and methods.* Mahwah, N.J.: London: Lawrence Erlbaum.

Hughes, A., 1993*, Backwash and TOEFL 2000*, Unpublished manuscript, University of Reading in England.

___, 2003, *Testing for language teachers* (2nd Ed.) Cambridge: Cambridge University Press.

Instituto Colombiano para el Fomento de la Educación Superior. (n.d). *Nuevo examen de estado para el ingreso a la educación superior: Cambios para el siglo XXI- Idiomas*. <http://www.icfes.gov.co/> [Retrieved November 1, 2004].

Madaus, G., 1990, *Testing as a social technology*. Paper presented at the inaugural annual Boisi Lecture in Education and Public Policy, Boston College.

Messick, S.,1996, "Validity and washback in language testing", *Language Testing, 13* (3), 241-256.

Merril, S., 1998, *Qualitative research and case study applications in education* (2nd Ed.) San Francisco: Jossey_Bass.

Pearson, G., 1988, "Tests as levers for change". In: D. Chamberlain & R. J. Baumgardner (eds.) *ESP in the classroom practice and evaluation*. Modern English Publications.

Petrie, H. G., 1987. "Introduction to evaluation and testing", *Educational Policy*, *1*, 175-180.

Quesada, R., 2001, Washback overrides the curriculum: An exploratory study on the washback effect on a high-stakes standardized test in the Costa Rican EFL high school context [Abstract]. DAI, 62, 05a, 1814.

Saif, S., 2000, Theoretical and empirical considerations in investigating washback: A study of ESL/EFL learners [Abstract]. DAI, 61, 03, 966.

Shohamy, E., S. Donitsa-Schmidt, and I. Ferman, (1996). "Test impact revisited: Washback effect over time", Language *Testing*, *13*, 3, 298-317.

Wall, D., 1996, "Introducing new tests into traditional systems: insights from general education and from innovation theory", *Language Testing, 13* (3) 334-354.

Wall, D., 1997, "Impact and washback in language testing", in: C. Clapham & D. Corson (Eds.), *Encyclopedia of language and education: Vol. 7. Language testing and assessment* (pp. 291-302), Dordrecht: Kluwer Academic.

Watanabe, Y., 2004, "Methodology in washback studies", in: Cheng, L., Y. Watanabe & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods.* (pp.19-36). Mahwah, N.J.: Lawrence Erlbaum.

The authors

** Norma Barletta is a teacher and researcher at the Instituto de Estudios Superiores en Educación, in Universidad del Norte.
E-mail: nbarlett@email.arizona.edu

*** Ofelia May works as a teacher and researcher at the Instituto de Estudios Superiores en Educación, in Universidad del Norte.
E-mail: omay@uninorte.edu.co

APPENDIX 1

Summary of Teachers' stated Perceptions, Beliefs and Decisions

Topics	Context 1	Context 2
Place of English teaching (school)	There is a recently constituted formal English Department Now only trained FL teachers teach English. In the past other teachers took over to complete required teaching hours. They have put together a project for English in the school. FL classes were made smaller in size. New textbook was chosen.	No cohesive work within the English Department. Teachers don't perceive special institutional effort except support for training. Laments students' little exposure to English as compared to other schools.
Aims of English class	To develop four communicative skills	Understanding texts; giving and supporting opinions. General educational goals related to students' identities and their 'projects of life'.
Syllabus and materials	Mainly follows the chosen textbook. The textbook is communicative because" it includes activities to develop oral skills". She needs to add more grammar exercises from other texts.	Doesn't follow any published textbook. Selected topics do not necessarily follow initial planning. Choice depends on students' interests. Her description of topics centers around grammar aspects (tenses, adverbs). She wants students to use English to describe their 'life projects'.
Notion of Communicative Competence	Good listening, reading, writing and speaking	Being able to use linguistic resources appropriately and effectively for own purposes.
Beliefs about teaching and learning	Grammar is essential. Oral activities cannot be done without it so level has to be lowered. Doing oral activities takes time from the more important.	Students should have fun, should not be afraid of using the language and should eventually like learning it.
ICFES	It is very important but it is basically a reading and grammar test. If schools worked towards communicative competence, the Exam should not be a problem. She tries to include evaluations in the way the ICFES does. She knows she has to change former teaching practice but does not offer concrete examples.	It tests grammar rather than students' real performance and it has little to do with their needs. What she does will help students cope with it in the long run because they will take pleasure in reading texts.

APPENDIX 2

Sample of Analysis of Classroom Observations (Context 1)

Obs.	Activity	Interaction	Feedback by the teacher	Competence	Evidence of students' level of competence
1	Oral presentations	Student's monologue addressing class. Mostly reading out loud from prepared poster.	Refers to appropriateness of visual aid.	Not clear	Students read texts. No evidence of whether it is their work or if other students understand.
1	Vocabulary translation	Teacher says names of food items in English. Students translate into Spanish	Gives correct translation.	Grammatical	Most words are correctly translated.
1.	Pronunciation	Teacher says the word, students repeat it.	None.		
2.	Gap filling exercise.	Students supply missing quantifiers in sentences (board). Students read the sentence with the missing word.	Uses Spanish to ask for alternative answers, writes correct answers (board). Provides brief explanation (Spanish).	Grammatical	Students' answers do not reflect full understanding of quantifiers. They make mistakes, and when asked again, they seem unsure. Teacher says last word.
2.	Graph interpretation	Students fill in missing words according to image, then describe the pictures using appropriate quantifiers (workbook)	Goes through exercise and supplies correct answers.	Grammatical	Students make mistakes (50%) in their workbooks. Images sometimes a bit confusing.

2.	Mini -dialogue practice following given pattern.	Students ask questions with *how many/much* .Answers with *a few/a little* based on examples and pictures given.	Listens to groups and corrects verb form, quantifiers, and also the word 'there' which is often mispronounced as 'these'. No new or reiterated explanation.	Grammatical	Students make a number of mistakes when working in pairs. There is no consistency in the use of quantifiers, verb forms or question forms..
2.	Gap filling	Students copy sentences with blanks from the board and individually fill in *much/many or how much/many*.	Corrects words students write and asks students to think if the concept is countable or uncountable. (Spanish).	Grammatical	Two thirds of the attempts to answer were correct.
3	Gap filling	Students individually (at home) fill in the blanks in a dialogue with *some, any, a little, a few, dozen, pound, enough, many and much.*	Checks answers and supplies the correct word.	Grammatical	Most answers were supplied by the teacher.
3.	Guided writing	Students write a sequence *using first, then, after that* and *finally*.(groups)	Supplies needed vocabulary, helps with sentence construction.	Textual coherence	While focus is on connectors, students have difficulty with word order, vocabulary. They use connectors correctly.
3.	Ordering sentences	Students number given sentences (5 altogether) in the correct order.(workbook)	Checks the order	Textual	No evidence
3	Guided writing	Students write the given sentence in the correct order, adding appropriate sequence marker (workboks)		Textual coherence	No evidence

4	Dramatiz-ation of dialogue	Students (pairs) recite memorized lines of a dialogue from the textbook	Corrects pronunciation	Grammatical Memory?	This was a former task some students had not done, so it is an opportunity to make up and improve grade.
4	Oral present-ations	Students make short oral presentations about favorite food or animals.	Encourages students to ask for revision of what they are going to say, the spelling and grammar in the posters and to practice pronunciation	Not clear	No clear point is made. Lots of mistakes, difficult to understand. Ex: " Iguana reptile from America central similar the lagarto"
4	Cloze exercise	Students in groups work on filling in the words *a few/ a little/ a lot* of in a two-paragraph text	Elicits correct word when the answer is mistaken. She corrects pronunciation a lot.	Textual	Students were very unsure of their answers. They seem not to have a criterion for deciding which word to put. A number of incorrect guesses were heard.
4	Oral translation into Spanish	Teacher asks individual students to translate phrases from the text.	Lets students hypothesize translations and gives the correct one. Focuses on discrete forms and words.	Grammatical	Only one student volunteered to translate. Most translation was supplied by the teacher.
5	Writing sentences	Individually students write sentences describing similarities or differences between foods.	Discusses the similarities or differences (Spanish). Corrects grammar in students' notebooks on individual bases.	Grammatical (syntax)	No clear evidence
5	Reading aloud	Students read short recipes from textbook out loud in turns.	Corrects pronunciation	Grammatical	Students made mistakes in almost every other word.