

# Washback on speaking and spoken interaction: teachers' perceptions in the French as an additional language classroom

MARIA DOLORS CAÑADA PUJOLS

ALBA MILÀ-GARCIA

*Universitat Pompeu Fabra*

Received: 29/07/2021 / Accepted: 19/05/2022

DOI: <https://doi.org/10.30827/portalin.vi40.21814>

ISSN paper edition: 1697-7467, ISSN digital edition: 2695-8244

**ABSTRACT:** The increasing importance of tests in society raises questions about their impact on teaching and learning situations, and washback is precisely concerned with the effects they can have on teaching materials and classroom management. Although research in washback is extensive, there are languages that have not received much attention, such as French. Using a corpus of interviews, we study the influence that a standardised exam at the end of the B1 level in Escoles Oficials d'Idiomes (EOI) in Catalonia (Spain) has on teachers' practices, particularly concerning oral activities. Although most of the teachers believe that their focus is still on teaching rather than testing, the interviews reveal that they may attach more importance to the exam than they realise. Moreover, they seem particularly concerned with the format of the spoken exam, the communicative nature of which they consider questionable. The results of this qualitative analysis further the knowledge of the role of washback in the teaching of additional languages.

**Keywords:** assessment, washback, French as an additional language, speaking, teachers' perceptions

***Washback en la expresión e interacción orales: la percepción del profesorado de francés como lengua adicional en el aula***

**RESUMEN:** La importancia que ha adquirido la certificación de conocimientos en la sociedad actual lleva a plantearse su repercusión en las situaciones de enseñanza/aprendizaje, y el concepto de *washback* se refiere precisamente a los efectos que los exámenes pueden tener en los contextos formativos. Aunque la investigación sobre el *washback* es considerable, su efecto en el caso del francés todavía ha recibido poca atención. A partir de un corpus de entrevistas, estudiamos la incidencia de un examen estandarizado al final del nivel B1 en las Escuelas Oficiales de Idiomas (EOI) de Cataluña en las prácticas del profesorado, especialmente en lo que se refiere a las actividades de expresión oral. Aunque la mayoría de las y los docentes creen que su objetivo sigue siendo la enseñanza de la lengua y no los exámenes, las entrevistas revelan que le otorgan más importancia de lo que creen. Además, están especialmente preocupados por el formato del examen oral, cuya naturaleza comunicativa consideran cuestionable. Los resultados de este análisis cualitativo abundan en el conocimiento actual en torno al efecto del *washback* en la enseñanza y aprendizaje de las lenguas adicionales.

**Palabras clave:** evaluación, washback, francés como lengua adicional, expresión oral, percepción del profesorado

## 1. INTRODUCTION

Washback, understood as the influence of testing on teaching and learning, has received continuous attention since its existence was questioned and proven in Alderson and Wall (1993) and developed, among others, in Cheng et al. (2004). High stakes tests and standardised testing in general fulfil an important social function, as the rewards for passing may be substantial, including professional promotion or access to post-graduate education programs, among other things. One example of this kind of testing is the standardised exam certifying a B1 level according to the *Common European Framework of Reference for Languages* (henceforth CEFR) which is administered by the *Escoles Oficials d'Idiomes* ('Official Language Schools', henceforth, EOI) in Catalonia, currently one of the Autonomous Communities making up Spain. Since 1996, this exam has been standardised under the auspices of the Catalan Department of Education, which designs the exam, and it is administered by all the EOIs in Catalonia. Because a pass mark entitles the exam-taker to an officially recognised certificate widely required in professional settings, there is considerable public demand for this exam. Given the importance attached to it, it is hardly surprising that it produces washback—whether positive or negative—in the classroom. The aim of this paper is to analyse this washback through the perceptions and beliefs of 11 experienced teachers of French as an additional language (henceforth, FAL) who teach B1 level courses, following Winke's (2011, p. 633) view that teachers' perceptions are "valuable pieces of information concerning whether tests affect the curriculum as intended" and their beliefs can be more decisive than other factors and can affect instructional decisions (Winke and Lim 2017).

Through the use of interviews, we will address the following research questions:

- 1) What are FAL teachers' thoughts on the speaking and spoken interaction section of the official B1 exam?
- 2) What are the washback effects of the FAL B1 exam on the preparation of the course and on the choice of activities teachers carry out to enhance learners' oral competence?
- 3) From an affective point of view, how do teachers approach the official exam at the end of the year?

## 2. LITERATURE REVIEW

### 2.1. Washback: the role of standardised tests on language teaching and learning

Washback has been traditionally linked to high-stakes tests, which are commonly used as requisites for university admission, employment or legal residency status (Wall, 2013). However, washback studies have also looked at standardised exams that take place in controlled settings but that do not necessarily have such pivotal importance (see, for example, Geranpayeh & Taylor, 2013, regarding the Cambridge English exams). A distinction is made between positive and negative washback: according to Taylor (2005) the former occurs when the presence of the exam leads to good teaching practices, while the latter takes place when the exam places constraints on the teaching and learning process. Wall (2013) relates this distinction to the connection between what the exam represents and the aims of the curric-

ulum: if they match, there will be positive washback, but if they do not and teachers must make changes to their values and goals in order to adapt, there will be negative washback. As a consequence, for optimal positive washback there should be little if any difference between activities involved in learning the language and activities involved in learning the test. The dominant paradigm in EOIs is the one established in the CEFR, which builds on the Communicative Approach, the dominant paradigm in teaching languages all around the world (Alamri, 2018), to focus on an action-oriented approach. In this approach, curricula and courses are based on learners' real world communicative needs, and assessment revolves around communicative ability in real life (CEFR). Therefore, it follows that the tasks contained in the exam should be communicative and common in the classroom.

On the other hand, some authors question the validity of this distinction by shifting the attention to the various contextual factors that also affect teaching practices. Wall (2013) stresses the need to take context into account in any attempt to study washback; Green (2013) suggests a change in the relationship between tests and teaching and learning, which moves from an influence that the former plays on the latter to an interaction that goes both ways; and Barnes (2017) questions the very notion of good teaching practices and states that, even if an exam is designed so that the practice that leads to it can follow a communicative approach, this does not guarantee that teachers will adopt it in their teaching practices. Along the same line, Bachman and Palmer (2010) venture that an assessment might lead to unintended consequences, which could be detrimental.

In fact, the relevance of contextual and personal factors was first pointed out by Alderson and Wall (1993), with teachers at the very centre (Alderson & Hamp Lyons, 1996). However, Ali and Hamid (2020) warn researchers about the danger of emphasising a causal relationship between testing and teaching that disregards other relevant factors. In his study on teacher professionalism, Runté (1998, p. 166) warned about the different ways in which standardised tests can affect teachers and their practices. For example, while on the one hand a standardised test obviates the need for the teacher to design any final evaluation, on the other hand it requires that the teacher should follow a centralised curriculum that reflects the needs of the exam. Since there are so many ways in which teachers can be influenced by washback, they have become the main focus of attention in washback studies. In washback research they fall under the category of *participants* (also called *stakeholders* in the literature) (Hughes, 1993, as reported in Green, 2013). The other two areas of washback research are the *process*, which revolves around the actions that are part of the learning process, and the *product*, which focuses on the learning outcome. Following Gebril and Eid (2017), this paper is concerned with the first two, participants and process, and it explores teachers' perceptions about the influence a standardised test exerts in a specific context and in a specific area, speaking and spoken interaction. Therefore, within the factors that Spratt (2005) established as affecting the form and intensity of washback, teacher-related factors (which include beliefs, attitudes, education, experience and personality) will be taken into account, together with school factors (the context in which teaching, learning and testing take place) and factors related to the test instrument itself (its stakes, purpose and format).

Regarding research about standardised FAL tests, Vandergrift (2015) points out that there is very little empirical research on the tests themselves or their use with different language learner populations, especially in the case of the DELF (*Diplôme d'études en langue française*). According to this author, washback is usually approached tangentially and

without using empirical data, and Elder (2018) affirms that this is an area in which further research and/or documentation is needed.

Finally, we must acknowledge the desired place that speaking should take within communicative approaches: it is placed at the centre of the learning process, and it should be conceived from a holistic approach (Goh & Burns, 2012). However, introducing it in class meaningfully is a challenging endeavour (Burns, 2012). This can become even more challenging in a course that ends in a certification exam such as the one under study. Often, the focus is set on normal conversation, although it is undoubtedly difficult to establish exactly what constitutes it. In this regard, Fulcher (2015) argues that a choice must be made in test design between normal conversation and domain-specific interaction (i.e., classroom discourse genres), because they are quite different.

## 2.2. Emotional Factors and Speaking Tests

There is abundant literature showing that affective characteristics, such as motivation, can influence language learning processes (e.g., Dörnyei & Ryan, 2015). Furthermore, the need to consider the role of emotions, whether positive or negative, is widely acknowledged. Although anxiety—one of the most prominent of emotions in the additional language learning context—is undoubtedly “multi-faceted” (Horwitz, 2010, p. 154) and must be explored and understood from an ecological perspective (Gkonou, 2017), it has been proven to be debilitating on L2 learning and achievement” (Dewaele et al., 2016; Horwitz, 2017). Among the different skills, speaking activities are the ones that lead to language anxiety more often (Horwitz, 2001; Young, 2013). On the other hand, the negative correlation between test anxiety and grades has also been explored in previous studies (e.g. Hewitt & Stephenson, 2012; Salehi & Marefat, 2014). In connection with testing, Rubio Alcalá (2004) states that, regardless of the assessment system adopted, anxiety can reach the point where it prevents students from showing their true communicative and linguistic competence. Therefore, anxiety and the way teachers address it in the classroom are crucial factors that must be included in any study of washback.

## 3. METHODS

### 3.1. Setting: EOIs and the B1 certificate

All the teachers interviewed in this research work at EOIs, which are publicly-funded non-university educational centres devoted to the specialized teaching and learning of modern languages to teenagers and adults. Catalonia has 46 EOIs teaching from two to fifteen different additional languages, and French is taught in 38 of them (82.6%). However, demand is especially high only for the first three French courses, upon successful completion of which students receive an official B1-level certificate if they pass the standardized test, as will be explained below.

The B1 level established in the EOI curricula follows the Council of Europe recommendations as specified in the CEFR, and the exam requires a minimum score of 65% to

achieve a passing grade. At the time of the interviews (October–December 2017), the spoken production and interaction part of the B1-level official exam was described by the Catalan Department of Education as follows:

The oral skills part of the exam will be simultaneously administered to two candidates, who will interact. It will consist of two tasks.

- Task 1. Description of a picture and narration of an experience that happened in the past. The two candidates must ask each other questions and answer them. They must also answer questions from one or both of the two examiners.
- Task 2. Prompted exchange of opinions between the two candidates.

Duration: 20–25 minutes

Following the CEFR, this exam distinguishes between oral production (speaking), in which students must show their capacity to produce a sustained monologue (task 1), and spoken interaction, where they engage with their fellow candidate and two examiners (task 2). Access to this exam is not limited to those who have completed coursework at the EOI but is open to any members of the public who seek certification of their additional language mastery. In such cases, the candidates are unlikely to be familiar with either the other candidate with whom they interact during the exam or the two examining teachers. With regard to internal candidates (those who have completed the appropriate prior course), they will be familiar with the other candidate, and they will be familiar with one of the examining teachers. Finally, it must be noted that during the first term of the 2017–2018 academic year teachers were notified that changes would be made to specific parts of the oral exam, as will be explained in section 4.1 below.

### 3.2. Procedures

The data for this research is extracted from a corpus of qualitative interviews, that is, “professional conversations [...] where knowledge is constructed in the inter-action between the interviewer and the interviewee [...] about a theme of mutual interest” (Holter et al., 2019, p. 2). Qualitative interviews provide thick descriptions or rich data, which is a hallmark of qualitative inquiry (Brekhus et al., 2005) as well as a requirement for a valid qualitative analysis. Labels used to classify qualitative interviews vary across scholars; following Brinkmann’s classification (Brinkmann, 2007), we have done eleven doxastic interviews (which totalled 12 hours and 45 minutes). Doxastic interviews focus on the interviewee’s experience, attitudes, and understanding of the context, thus the interview is considered a research instrument (Talmy, 2010). The topics addressed were assessment and speaking and spoken interaction in the FAL classroom.

Eleven FAL B1 teachers answered the call to participate in the interviews, which took place during the first term of the 2017–2018 academic year. The interviews, which were conducted in Catalan or Spanish, took place at the teachers’ respective EOIs, on the grounds that they would feel most at ease in familiar surroundings. The fact that the interviewer had in fact previously worked as a French teacher made it easier for her to establish rapport with the interviewees.

### 3.3. Participants

The relevant information about the participants is provided in table 1. As can be seen, they had been teaching French for an average of 16.4 years and had an average of 7.2 years of experience teaching the EOI B1 level.

**Table 1.** *Research participant profiles*

PSEUDONYM	SEX	EOI	EXPERIENCE AS A FAL TEACHER	EXPERIENCE IN B1 FAL COURSES
Alicia	female	A	26 years	3 years
Albert	male	B	6 years	10 years
Arlet	female	C	13 years	9 years
Carmen	female	G	6 years	1 year
Cesc	male	A	26 years	7 years
Charo	female	A	13 years	3 years
Elvira	female	I	9 years	4 years
Fina	female	E	10 years	10 years
Helena	female	D	20 years	11 years
Mireia	female	H	19 years	13 years
Pablo	male	F	32 years	8 years

### 3.4. Analyses

The interviews were transcribed and annotated with Atlas-ti software, using the technique known as content analysis, “a research technique for making replicable and valid inferences from texts [...] to the contexts of their use” (Krippendorff, 2018, p. 24). With the aim of “making sense out of the collected data” (Merriam, 2009, p.178), the two authors independently conducted a first round of coding, looking for themes and subthemes covered in the interviews, such as speaking (including assessment, activities, concept, relevance, interaction, activities) and assessment (criteria, self-assessment, external assessment, feedback, test preparation). A thorough discussion followed in order to compare the annotations, establish clear boundaries for the subthemes and to agree on debated segments, in order to ensure consistency in coding. Finally, the quotes annotated with these (sub)themes were recovered, reviewed and discussed to obtain an elaborate picture of teachers’ perceptions concerning the topics under discussion.

## 4. RESULTS

The interviews raised interesting points regarding teachers' views on exam format, washback and their emotional relationship with the exam. We will address in three different subsections the three research questions formulated at the end of section 1 above.

### 4.1. Assessment of the speaking and spoken interaction format of the B1 official exam

As noted, at the time of the interviews the official exam was undergoing some changes: the Catalan Department of Education had changed the format of some tasks of the oral part. Teachers knew that the new exam would be a bit different, but the specific knowledge they had about those changes varied among them. This change in the format in fact addressed the general discontent expressed by teachers with regard to the previous version of this exam task, which they found too difficult and complex, as Arlet observed: "it is too tricky, confusing". Elvira's comments both show her displeasure and justify her opinion (note that all quotes have been translated from Catalan and Spanish into English):

*No, [in the previous version of the exam] there was a picture, but they couldn't describe it. I mean, I've got a picture, you've got a picture and I have to guess what's on your picture without having you describe it to me. Honestly, that's not quite natural, and that's what we complained about, you see? [...] In real life when you've got a picture the first thing you do is show it, and if you're on the phone, you describe it, right? I mean, we didn't like that at all.* Elvira

Other teachers agreed with Elvira about the artificiality of the task, in the sense that it hardly seemed like an authentic interaction. Cesc's comments are especially sharp, because he criticises first the format and then the institution behind its design:

*Because [...] you must be missing something if you design a spoken task where you force students to use rather artificial structures that end up being repeated and that bear no resemblance to a normal interaction.* Cesc

When informed that from then on the picture-based task would solely consist of a description by each candidate of their picture, which would then trigger an interaction between the two candidates, all teachers rated the new format very positively.

### 4.2. Washback of the B1 official exam in teaching practices

As expected, all teachers were constantly aware of the reality of the official exam at the end of the academic year. However, they had different ways of integrating the exam into the course they taught. Their approaches to the exam were situated on a continuum between two general inclinations: those who refused to be dominated by the exam and those who believed that the exam requires conscious and intentional preparation.

Teachers pertaining to the first group felt that the third course at the EOI is not intended to be a preparation course for the official exam, which focuses on achieving specific skills:

*I always tell [students] “I will teach you as much French as I can, because you must learn French. If you can speak French by the end of the course you will get the certificate because you will pass on your own, we do not have to prepare for it. This is more or less the concept that I try to instil in them. Helena*

*This is what [a teacher he had worked with years before] said. She always said “those teachers who spend the third and fifth year teaching their lessons as if they were at a driving school, focusing only [...] focusing on training their students for the exam, and that is not right”. And I have always believed she was right. Albert*

This group of teachers believed that the only difference between the third course and the other courses was that they were “freed” from the need to qualify their students through a course-specific final exam (although there is still an ongoing assessment of students during the course) because in effect the final exam for the course is the standardised exam; if they pass it, they pass the course and obtain the official B1 certificate. Interestingly, this “freedom” was previously observed by Alderson (foreword to Cheng et al., 2004) in his interviews to TOEFL teachers, who liked the fact of not having to prepare lessons and give feedback in those courses.

As for those teachers in favour of conscious exam preparation, Elvira and Charo were convinced that giving students the necessary language skills is not enough, that the teacher must fully be aware of how the exam is structured and practice for it. Charo felt strongly that mastering the techniques required for the exam was directly linked to passing it: candidates that had received classroom instruction at the EOI were more likely to pass the exam than external candidates:

*That is why I think that internal students have already obtained a big part... a big part of the mark, because they are taught well and well-prepared, because we practice a lot, we practice a lot in order to get them there, they know the exam [...] they already know... They know how the exam works and proceeds. Charo*

Some teachers positioned themselves between the two positions presented above, devoting a certain amount of time to explaining exam format without letting it become a central element of the course. Many teachers reported that they described the exam format to the students, gave them some advice about how to approach it and at times had students practise the different skills involved separately. For teachers that felt the need for a bit of preparation, one activity commonly used was a mock exam. Arlet, Helena and Carmen reported programming only one during the whole course, while Albert and Fina programmed two per course. The other teachers did not explicitly mention using them.

Regardless of the use of mock exams in the classroom, all interviewees reported introducing activities that were similar in format to specific parts of the exam to one extent or another. Pablo, for example, said he practised the spoken part of the exam “quite frequently”. On the other hand, Alicia said she did not want the classroom to become a crash course on the exam, so she introduced it “little by little”, and whenever the program of the course allowed for it, for example by adding the sort of “exchange of opinions” that makes up the spoken interaction part of the exam. Arlet and Helena reported following a similar approach for the same reasons:

*We do it a bit [...], but no, we don't do it systematically, in the sense of "once a week we'll do X", no, no, no. [...] During the course we do some activities and exercises that serve as "Look, we've done this as if it were that of the official exam" [...] so that students know.* Helena

Cesc, Arlet and Fina put forward two additional arguments to justify their working on a specific activity more frequently. Cesc explained that the nature of the second part of the exam had not been introduced in lower-level courses (unlike the description of pictures) and he therefore felt obliged to “work with it a lot” in the classroom. Similarly, the recent changes in exam format led Arlet to shift the focus to dialogues “because that is what students will face at the end of the year”. This suggests that recurrent practices that were tied to the exam were beginning to change: “now I focus on that instead” (Arlet). As for Fina, she was also planning to make some changes to her program: “I need to revise it, because it has changed.”

Besides the novelty of the new format, another reason given by Fina to justify working on a specific exam-related activity more frequently was the difficulty this activity entailed. That is why she had chosen to place a lot of emphasis on speaking activities in the classroom which were very similar to those found in the exam:

*The speaking is also very hard for them —the pictures they give them in the third year, the third year official exam, where they're supposed to describe a picture. Since there are two similar pictures, they prepare the pictures, also in pairs [...].* Fina

This difficulty became an added motivation for students, and Mireia took advantage of this motivation in the form of “mock oral exams”, which she scheduled during her office hours:

*I follow the format of the exam, because that helps them to prepare, and it also motivates them to come and all that; you always need to take them for a ride, so to speak [laughter]. But I offer them feedback, that is, they do the exam in pairs and then I make comments about it [...]. “You should do this, you should do that, you're struggling a bit in this part”.* Mireia

Finally, as shown by Arlet, some teachers had already abandoned some of their previous practices now that those had been excluded from the new exam format. Alicia sums it up succinctly: “This is not in the third year [exam] anymore, so [there is no need to do it in class]”.

### 4.3. Emotional washback in teachers and students

The teachers seemed to be aware that emotions play a crucial role in the process of language teaching and learning. Most of them focused on the personal experience that students undergo regarding the exam, and only Elvira and Helena mentioned how the exam affected them as teachers. Helena seemed to find it “perfect”, while Elvira explained that the official exam made her restless, a negative emotion that she tried not to transmit to her students:

[The exam] *affects everyone [...]. I'm always thinking about the exam in almost everything I do. I try not to make them obsess over it, since I'm already doing it.*  
[laughter] *And yes, it affects me.* Elvira

Although, one way or another, all the teachers were aware of the importance of emotional issues, they seemed to adopt different strategies regarding their possible effects. Albert showed quite an extreme approach, completely denying that the prospect of having to take the exam could possibly stress out his students; he felt that if the “students become super anxious”, then their teacher is to blame. Albert tried to play down the exam to his students:

*I think it more or less depends on what we pass on to students. Of course, if they see you as a teacher who is super focused on the exam, who practically obsesses over it, and I don't, in my case I don't even talk about it that much. We carry on with our B1-level lessons and that's it, and we keep working, and they keep learning, and they know that there's an exam at the end.* Albert

The attitudes of the remaining teachers fell somewhere between Elvira's constant worry and Albert's denial. In general, they were aware of how students perceived the standardized exam. Charo talked, for example, about the “emotional load of an official exam”, to which she attributed two causes. On the one hand, students were going to have to interact with an unknown examiner (the other one being their own teacher), and, on the other—and more importantly—, oral examinations are quite rare in Spain, and therefore students were not used to taking part in them:

*Having two people looking at you, people who don't know you, there's tension. [...] I think that oral exams are one of the biggest challenges for language students in Spain: oral exams are like twice as hard and, moreover, in a foreign language.* Charo

These teachers deployed two different non-exclusive strategies to deal with their students' anxieties about the exam: either they tried to minimize, or they accepted their existence and tried to do something to reduce their students' level of stress. Thus, for example, Helena tried to minimise the pressure created by the upcoming exam by comparing it to a special reward that students would receive at the end of the year (“a piece of candy”). However, this approach did not prevent her from doing classroom activities that modelled the various sections of the official exam in order to reduce her students' feelings of nervousness:

*I try, what I'm saying is that I try to teach them French, and if they can speak French they're going to be able to pass the exam. But even then you have to give them a sense of security. The students have to know where they're going, right?*  
Helena

On the other hand, Carmen's approach was somewhat different. She was also aware of the students' stress, which she said happened because they were facing “something new for them, something that is completely unknown to them”. She likewise acknowledged that

passing the exam and thus being able to claim possession of the B1 certificate on their CV could have an impact on their job prospects (“for them it is like a pressure”). She therefore made sure to do exam-like activities in class on a regular basis:

*What I try to do, in order to calm them down, is to one day bring [activities that follow the] exam template, so that, every three weeks or so, they receive activities like those of the exam. And this seems to calm them down a bit. Carmen*

Our data show that teachers were aware of a specific need that the Department of Education developers of the official exam did not fully address, which is that students want more than a sample exam provided on the official website and information about the exam’s structure. As Alicia states, students want feedback: “the students have a need, which is to be corrected,” because they are focused on the certificate that they can obtain at the end of it.

## 5. DISCUSSION AND CONCLUSION

At the time of the interviews, the shared feeling among teachers was satisfaction, because the format of one of the tasks of the speaking that they had repeatedly criticised had finally been changed. As we have seen, teachers do not shy away from expressing their negative views on tasks that, from their point of view, must be improved. This straightforward attitude follows East’s (2015, p. 102) findings: “When the perceptions of the primary consumers (teachers and students) differ from the assessment developers, very strong feelings about the assessment can be evoked.” Regarding the certification of oral skills, the teachers interviewed here were quite clear in stating that these tasks must reflect communicative situations that are close to those likely to be encountered in real life, that are authentic, as argued by Messick (1996). In fact, one of the most widely shared views with regard to the previous exam format was that it lacked authenticity, but that is to be expected: in elaborating on Bachman and Palmer (2010), Green (2013, p. 41) wrote that “there are restrictions on the type of tasks [test developers] can employ which mean that test tasks can never fully reproduce a ‘real life’ experience”.

As shown in section 4.2, the effect of washback on teaching practices can take different forms. For some teachers it means focusing on training their students constantly so that they get used to the exam format. For others, it means trying to minimise the importance of the test, although this does not imply that they disregard its format altogether. Our analysis showed that washback affects the choice of activities: once the exam format was changed, some teachers contemplated replacing some of their former exam-preparatory activities with others more in keeping with the new exam. This inclination to adapt instructional materials to attune them to a standardised exam was initially attested by Mathison (1987), and Smith et al. (1990) related this decision to the pressure teachers feel to improve students’ test scores. These changes in the selection of speaking activities that derive directly from the exam format are a result of negative washback, since, in a course that is clearly framed within the CEFR teachers seem to select classroom activities according to an exam format that does not completely reflect real communicative situations.

In fact, we have confirmed that, regardless of whether teachers believe that systematic preparation for the exam format is needed, they all prepare their students for it in one way or another, with varying degrees of frequency and emphasis. This means that they find this practice beneficial for their students, which agrees with the results found by Gebril and Eid (2017). These authors concluded that teachers regarded preparing students for the tests as both useful and beneficial since it familiarised them with the item format, among other things. Similarly, Winke and Lim (2017) found that exam preparation for the listening part of a standardised exam was helpful in that it made students familiar with the exam format, which is a crucial aspect of exam preparation. Finally, we cannot disregard the students' need to become familiarised with exam format: Alderson and Hamp Lyons (1996) suggested that it was actually students that brought teachers to adopt exam-oriented materials in the classroom, and Lumley and Stoneman (2000) found that students actually expect teachers to use them.

Concerning emotional washback, previous literature has proven that it interferes in additional language learning, especially in speaking tasks, and various authors include language testing and assessment as one of the main sources of anxiety for students (Young, 2013; Zhang & Rahimi, 2014). Within language testing, one factor that increases anxiety is unfamiliar exam format, which proves once more that even a little preparation is necessary to help the students' emotional state. Although teachers stated that they are aware of their students' anxiety, none of them explicitly addresses or asks them how they feel, which should be at the base of a successful teaching approach that is centred on students' needs. As stated by previous research on anxiety, teachers' efforts to create an environment where anxiety is reduced both in class and during tests can prove really beneficial for students and will positively impact their learning and performance (Smyth et al., 2021). On the other hand, the analysis has shown that teachers can also be affected by test anxiety. However, the majority of the teachers that we interviewed simply recognise the importance of the exam and try to prepare their students accordingly, mostly by working with activities and resources with a format that resembles that of the exam.

This study aims to widen the reach of the research on the multi-faceted nature of washback (see Cheng 2013 for a review of the challenges for future research) by focusing on a specific skill and context: speaking in a FAL B1 course that ends with a standardised exam. In this sense, the in-depth qualitative interviews with eleven experimented teachers allow for a detailed exploration and description of the phenomenon. However, the specificity of the participants and the context become, in turn, a limitation, since they provide results for one additional language (French), one specific level (B1) and one geographical area (Catalonia). Despite this limitation, we believe that the conclusions drawn from the meaningful interactions with these teachers can pave the way for future research on washback, since they uncovered distinct areas of concern for teachers that can contribute to a focalisation on the exploration of teachers' perceptions.

We have shown that the relationship between testing and teaching is overly complex and far from causal (Ali & Hamid, 2020), as our results expose once more the existing interrelations among educational ecosystems, and the role they each play in the washback that can result from official examinations. Moreover, in this specific case the influence of tests on teaching is not unidirectional (Green, 2013), as teachers' feedback about the exam leads to significant changes in its design, which has now improved its *communicativeness*.

Still, the influence that the final official exam has on teachers' practices seems to be higher than intended by test developers. If they favour some activities over others due to their inclusion in the exam, we must wonder whether other fundamental aspects of speaking and spoken interaction are left out because they are not "relevant". This discovery can prove useful for education policy-makers in charge of official exams, since it points to the need to offer continuous training to the teachers who implement the corresponding courses. Not only would this help to prepare teachers emotionally, by helping them to reflect on their own beliefs and practices, but it might also reduce the differences between what developers expect teachers to do and what teachers actually do following their own beliefs. This training would undoubtedly supplement their assessment literacy, a concept that has been gaining prominence in language teaching (Coombe et al., 2020).

## 6. REFERENCES

- Alderson, C., & Hamp Lyons, L. (1996). TOEFL preparation courses: A study of washback. *Language Testing*, 13(3), 280–297. <http://doi.org/10.1177/026553229601300304>
- Alderson, C., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14(2), 115–129. <https://doi.org/10.1093/applin/14.2.115>
- Ali, M., & Hamid, O. (2020). Teaching English to the test: Why does negative washback exist within secondary education in Bangladesh? *Language Assessment Quarterly*, 17(2), 129–146. <https://doi.org/10.1080/15434303.2020.1717495>
- Bachman, L., & Palmer, A. (2010). *Language assessment in practice*. Oxford University Press.
- Barnes, M. (2017). Washback: Exploring what constitutes "good" teaching practices. *Journal of English for Academic Purposes*, 30, 1–12.
- Brekhus, W. H., Galliher, J. F., & Gubrium, J. F. (2005). The need for thin description. *Qualitative Inquiry*, 16(6), 1–19. <https://doi.org/10.1177/1077800405280663>
- Brinkmann, S. (2007). Could interviews be epistemic? An alternative to qualitative opinion polling. *Qualitative Inquiry*, 13, 1116–1138. <http://doi.org/10.1177/1077800407308222>
- Burns, A. (2012). A holistic approach to teaching speaking in the language classroom. In *Symposium 2012 - Lärarrollen i svenska som andraspråk*, Stockholm, October 2012, 165–178. Available at [https://andrasprak.su.se/polopoly\\_fs/1.204517.1411636356!/menu/standard/file/Anne\\_Burns.pdf](https://andrasprak.su.se/polopoly_fs/1.204517.1411636356!/menu/standard/file/Anne_Burns.pdf).
- [CEFR] Council of Europe (2001). *Common European framework of reference for languages: learning, teaching, assessment*. Press Syndicate of the University of Cambridge.
- Cheng, L. (2013). Consequences, impact, and washback. In A. J. Kunnan (Ed.), *The companion to language assessment* (pp. 1–14). John Wiley & Sons. <https://doi.org/10.1002/9781118411360.wbcla071>
- Cheng, L., Watanabe, Y., & Curtis, A. (Eds.) (2004). *Washback in language testing: Research contexts and methods*. Lawrence Erlbaum.
- Coombe, C., Vafadar, H., & Mohebbi, H. (2020). Language assessment literacy: what do we need to learn, unlearn, and relearn? *Language Testing in Asia*, 10(3). <https://doi.org/10.1186/s40468-020-00101-6>
- Dewaele, J. M., MacIntyre, P. D., Boudreau, C., & Dewaele, L. (2016). Do girls have all the fun? Anxiety and enjoyment in the foreign language classroom. *Theory and Practice of Second Language Acquisition*, 2(1), 41–63.

- Dörnyei, Z., & Ryan, S. (2015). *The psychology of the second language learner revisited*. Routledge.
- East, M. (2015). Coming to terms with innovative high-stakes assessment practice: Teachers' viewpoints on assessment reform. *Language Testing*, 32(1), 101–120. <https://doi.org/10.1177/0265532214544393>
- Elder, C. (2018). Test review: Certifying French competency: The DELF tout public (B2). *Language Testing*, 35(4), 615–623. <https://doi.org/10.1177/0265532218781627>
- Fulcher, G. (2015). Assessing second language speaking. *Language Teaching*, 48(2), 198–216. <http://doi.org/10.1017/S0261444814000391>
- Gebril, A., & Eid, M. (2017). Test preparation beliefs and practices in a high-stakes context: A teacher's perspective. *Language assessment Quarterly*, 14(4), 360–379. <https://doi.org/10.1017/S0261444814000391>
- Geranpayeh, A., & Taylor, L. (2013). *Examining listening*. Cambridge University Press.
- Gkonou, C. (2017). Towards an Ecological Understanding of Language Anxiety. In C. Gkonou, M. Daubney & J.-M. Dewaele (Eds.), *New Insights into Language Anxiety. Theory, Research and Educational Implications* (pp. 135–155). De Gruyter. <https://doi.org/10.21832/9781783097722-009>
- Goh, C. C. M., & Burns, A. (2012). *Teaching speaking: A holistic approach*. University Press.
- Green, A. (2013). Washback in language assessment. *International Journal of English Studies*, 13(2), 39–51. <http://doi.org/10.6018/ijes.13.2.185891>
- Hewitt, E., & Stephenson, J. (2012). Foreign Language Anxiety and Oral Exam Performance: A Replication of Phillips's *MLJ* Study. *The Modern Language Journal*, 96(2), 153–289. <https://doi.org/10.1111/j.1540-4781.2011.01174.x>
- Holter, M. T., Johansen, A. B., Ness, O., Brinkmann, S., Høybye, M. T., & Brendryen, H. (2019). Qualitative interview studies of working mechanisms in electronic health: Tools to enhance study quality. *Journal of Medical Internet Research*, 21(5), e10354. <https://doi.org/10.2196/10354>
- Horwitz, E. K. (2001). Language anxiety and achievement. *Annual Review of Applied Linguistics*, 21, 112–126. <https://doi.org/10.1017/S0267190501000071>
- Horwitz, E. K. (2010). Foreign and second language anxiety. *Language Teaching*, 43(2), 154–167. <https://doi.org/10.1017/S026144480999036X>
- Horwitz, E. K. (2017). On the misreading of Horwitz, Horwitz and Cope (1986) and the need to balance anxiety research and the experiences of anxious language learners. In C. Gkonou, M. Daubney, & J. M. Dewaele (Eds.), *New Insights into Language Anxiety: Theory, Research and Educational Implications* (pp. 31–47). Multilingual Matters.
- Krippendorff, K. (2018). *Content analysis. An introduction to its methodology* (4<sup>th</sup> ed.). Sage.
- Lumley, T., & Stoneman, B. (2000). Conflicting perspectives on the role of test preparation in relation to learning? *Hong Kong Journal of Applied Linguistics*, 5(1), 50–80.
- Mathison, S. M. (1987). The perceived effects of standardised testing on teaching and curriculum. Unpublished dissertation, University of Illinois at Urbana-Champaign, Urbana.
- Merriam, S. M. (2009). *Qualitative research: A guide to design and implementation*. Jossey-Bass.
- Messick, S. (1996). Validity and washback in language testing. *Language Testing*, 13(3), 241–256. <http://doi.org/10.1177/026553229601300302>
- Rubio Alcalá, F. (2004). *La ansiedad en el aprendizaje de idiomas*. Servicio de publicaciones de la Universidad de Huelva.
- Runté, R. (1998). The impact of centralized examinations on teacher professionalism. *Canadian Journal of Education / Revue canadienne de l'éducation*, 23(2), 166–181. <http://doi.org/10.2307/1585978>

- Salehi, M., & Marefat, F. (2014). The effects of foreign language anxiety and test anxiety on foreign language test performance. *Theory and Practice in Language Studies*, 4(5), 931–940. <http://doi.org/10.4304/tpls.4.5.931-940>
- Smith, M. L., Edelsky, C., Draper, K., Rottenberg, C., & Cherland, M. (1990). The role of testing in elementary schools. *CSE Technical Report 321*. CRESST. University of California.
- Smyth, A. M., García Manzanares, N., & Fernández-Muñoz, J. J. (2021). Anxiety and Personality as Indicators of Academic Performance in University Foreign Language Classrooms. *Porta Linguarum*, 36, 27–42. <http://doi.org/10.30827/portalin.v0i36.15376>
- Spratt, M. (2005). Washback and the classroom: the implications for teaching and learning of studies of washback from exams. *Language Teaching Research*, 9(1), 5–29. <https://doi.org/10.1191/1362168805lr152oa>
- Talmy, S. (2010). Qualitative in applied linguistics: From research instrument to social practice. *Annual Review of Applied Linguistics*, 30, 128–148. <https://doi:10.1017/S0267190510000085>
- Taylor, L. (2005). Washback and impact. *ELT Journal*, 59(2), 154–155.
- Vandergrift, L. (2015). The DELF in Canada: Perceptions of students, teachers, and parents. *The Canadian Modern Language Review*, 71(1), 52–74. <https://doi.org/10.3138/cmlr.1752>
- Wall, D. (2013). Washback in language assessment. In C. A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics*. <http://doi.org/10.1002/9781405198431.wbeal1274>
- Winke, P. (2011). Evaluating the validity of a high-stakes ESL test: why teachers' perceptions matter. *TESOL Quarterly*, 45(4), 628–660. <http://doi.org/10.5054/tq.2011.268063>
- Winke, P., & Lim, H. (2017). The effects of test preparation on second language listening test performance. *Language Assessment Quarterly*, 14(4), 280–297. <http://doi.org/10.1080/15434303.2017.1399396>
- Young, D. J. (2013). Affective Factors and Second Language Spanish. In K. L. Geeslin (Ed.), *The Handbook of Spanish Second Language Acquisition* (pp. 369–385). <https://doi.org/10.1002/9781118584347.ch21>
- Zhang, L. J., & Rahimi, M. (2014). EFL learners' anxiety level and their beliefs about corrective feedback in oral communication classes. *System*, 42, 429–439. <http://doi.org/10.1016/j.system.2014.01.012>