

## Comparing candidates' beliefs and exam performance in speaking tests

**Cristina Pérez-Guillot**

[mcperez@upv.es](mailto:mcperez@upv.es)

**Julia Zabala-Delgado**

[juzadel@upv.es](mailto:juzadel@upv.es)

**Universitat Politècnica de València, Spain**

### ABSTRACT

The development of a language exam is not a linear process but rather a round cycle in which, by using the test, we obtain information that will in turn be applied to improve each of the steps in the cycle. The goal of our study was to analyse students' beliefs about their performance in the speaking section of a language proficiency exam and compare them with their actual results in the exam, in order to determine whether their beliefs were based on their actual level of competence or if they were based on other factors, such as anxiety or stress due to the particular characteristics of this section of the exam.

**Key words:** *perceptions, competence, candidates, proficiency exam, oral skills, exam development*

### I. INTRODUCTION

The development of a language test is a process that involves several stages, from designing the test and endowing it with adequate contents, to administering the test and analysing the results obtained. However, developing a test is not a linear process but instead a round cycle in which, by using the test, we obtain information that will in turn be applied to improve each of the steps in the cycle.

The goal of our study was to analyse students' beliefs about their performance in the speaking section of a language proficiency exam in relation to other skills, and compare them with their actual results in the exam in order to determine whether their beliefs were based on their actual level of competence or if they were based on other factors, such as anxiety or stress due to the particular characteristics of this section of the exam. *Since the examination of reliability depends upon our ability to distinguish the effects (on test scores) of the abilities we want to measure from the effects of other factors*

(Bachman 1990, p.163), being able to differentiate external factors from the actual level of competence of the candidate would necessarily improve test reliability.

Consequently, determining the basis for their beliefs, either factual or self-perceived, would allow us to determine which aspects of the process could be modified to improve the reliability and thus the quality of our exam.

## **II. STATE OF THE ART**

The process of foreign language acquisition has been examined from different points of view – cognitive, psychological, linguistic, pragmatic and cultural, to mention just a few – and the exact nature of the process is still unknown.

Traditional language learning theories focus mainly on the study of what is learned and what is not learned in a language, explaining both processes by means of the strategies used to acquire knowledge and the reasons for success or failure in acquisition. This approach focuses on learning itself, on specific objectives and on the means used to achieve them and the results obtained. However, it fails to pay attention to the factors surrounding this process, the factors that add complexity to the process and that include not only objective components, but also subjective or external components, which will largely contribute to the final outcome. The starting point is therefore to consider language learning as a broad field in which external factors play an important role and, amongst them, those characteristics that are individual to each student and make their learning unique.

Examining the process of second language acquisition from this broad point of view, there are a number of subjective variables that belong to the students' individual field and that have a significant effect on their learning. This explains the different degrees of success in learning a language achieved by different subjects who follow the same programme and have a comparable intellectual ability. However, it is worth mentioning here that, although social and affective strategies are mentioned in the literature (Dörnyei 1994; Gardner and Lambert 1959; Hardison, 2014; Horwitz 1995; Jee 2014; Sparks et al. 2011), many of the authors reviewed (Bachman 1990; Bachman and Palmer 1996; Cohen 2003; O'Malley and Chamot 1990) focus primarily on cognitive and

metacognitive variables and the relationships between them, considering social and affective factors as crucial but difficult to quantify.

However, some studies (Baddeley 2007; Carroll and Sapon 1959; Conway et al. 2007; Pimsleur 1966) attribute this difference in the degree of success to the students' ability, on the cognitive aspect, leaving aside the emotional aspect, i.e. their attitude, motivation and beliefs about their own learning process. However, "if we were to devise theories of second language acquisition or teaching methodologies that were based only on cognitive considerations, we would be omitting the most fundamental side of human behaviour" (Brown 2000: 142). The affective domain is difficult to describe scientifically since it refers to emotion or feeling, yet the emotional side of human behaviour is intrinsically related to the cognitive side and needs to be taken into consideration.

Sustained by developments in the field of foreign language teaching towards student-centred learning, the study of these factors has become increasingly important together with, more specifically, the study of how students' perceptions and beliefs are a fundamental aspect on their path to learning a language. In fact, Foreign Language Anxiety (FLA) has been defined as a particular type of anxiety occurring specifically in foreign language learning situations,

"a distinct complex construct of self-perceptions, beliefs, feelings and behaviours related to classroom language learning arising from the uniqueness of the language learning process",  
"a phenomenon related to but distinguishable from other specific anxieties" (Horwitz and Cope 1986: 128-129)

and it is made up of three principal components: (a) communication apprehension, (b) test anxiety, and (c) fear of negative evaluation.

Motivation has also been considered influential in the degree of success of foreign language students, and this includes both instrumental motivation – the desire to obtain something from studying a second language – and integrative motivation – the desire to integrate into the culture of the second language (Gardner and Lambert 1959). In fact, and although both types of motivation contribute to second language learning success, students who are the most successful are those who are interested in the culture of origin and native speakers and have a desire to integrate into the society in which the language is used (Falk 1978).

Bachman and Palmer (1996) observed two types of variability in students' performance in language tests: (1) variability due to differences between individuals in terms of the language skills, strategies and processes used, as well as personal characteristics such as cultural and emotional differences, etc., and (2) variability due to the different characteristics of the method or tasks used in the test, such as the assessment modes or types of tasks used. According to Dornyei (2009) individual differences should be considered as higher level amalgams or constellations of cognition, affect and motivation that act as "wholes".

As a consequence of this approach, the subjective variables of the acquisition process mentioned above also play an important role in language testing, since the design of a test needs to take into consideration not only the characteristics of the tasks – test format, input provided, time allotted – but also the individual characteristics of the users – the positive or negative emotions or feelings they may have about their learning process, the examiner, the subject or context, the presence or absence of excessive anxiety when faced with the task, their motivation, etc. Accordingly, two aspects need to be taken into consideration simultaneously: (1) the characteristics of the task, which need to reflect the construct of the test and mirror target language use, and (2) the individual characteristics of the learner, which will affect their learning process and therefore their performance in a test situation.

As can be seen from the aforementioned arguments, although it would be desirable that the primary factor in the outcome of a language test were the ability of the test-taker or the adequacy of the test construct and structure of the tasks, in actual fact there are many other variables coming into play, ranging from the context to the individual characteristics of each test-taker.

Such factors become even more relevant in assessing speaking, since speaking in a foreign language is perhaps the most difficult skill to master as it involves a complex process of constructing meaning (Celce-Murcia and Olshtain 2000) which is performed at the same time as the act of speaking and therefore requires the planning and simultaneous monitoring of utterances. In fact, the ability to express oneself orally in a foreign language is a fundamental part of mastering language use and, as mentioned by

Luoma (2004), reflects not only our personality but also our self-image and ability to reason:

Speaking is also the most difficult language skill to assess reliably. A person's speaking ability is usually judged during a face-to-face interaction, in real time, between an interlocutor and a candidate. The assessor has to make instantaneous judgements about a range of aspects of what is being said, as it is being said. This means that the assessment might depend not only upon which particular features of speech (e.g. pronunciation, accuracy, fluency) the interlocutor pays attention to at any point in time, but upon a host of other factors such as the language level, gender, and status of the interlocutor and the personal characteristics of the interlocutor and candidate. (Luoma 2004: ix).

Furthermore, speaking a language is especially difficult for foreign language learners because effective oral communication requires the ability to use the language appropriately in social interactions (Fulcher 2003).

Traditionally, most students sitting official exams show high levels of stress when dealing with the speaking section of the test and explain their reaction by expressing their doubts about their own speaking ability (Phillips 1992; Stephenson and Hewitt 2001). However, and in light of the above, we believe that the fact that the speaking section of the test causes more stress in students is, in many cases, not because of their ability or lack of it, but because of the construct of speaking mentioned. As Bandura (1997: 37) states: “perceived self-efficacy is not a measure of the skills that one possesses, rather it is a belief about what one can do in the future, and under different conditions, with the skills that one has”. Consequently, perceived self-efficacy, the extent of one's belief in one's own ability to reach goals, will probably influence the way people will react in the face of difficulties and, therefore, a more positive perception of one's skills will influence performance in the real world. Real performance in the real world is in turn what performance in an exam situation should be expected to mirror and what exam tasks and context should be expected to elicit.

### **III. METHODOLOGY**

#### **III.1. Participants**

The participants in our survey<sup>1</sup> were the candidates sitting the CertACLES exam at the Universitat Politècnica de València (UPV) in June 2014. There were a total of 324 candidates for the three examinations carried out in June 2014: 101 for the B1 examination, 186 for the B2 examination, and 37 for the C1 examination. Our survey obtained 201 answers, that is, 62% of the candidates voluntarily took part in the study. It is interesting to note that the higher the level of language exam, the higher the participation of candidates: 55% of B1 candidates participated in the study, compared to 58% of B2 candidates and 62% of C1 candidates.

CertACLES exams are proficiency exams developed by the Language Centre of the UPV in accordance with the model developed by the Spanish Association of Language Centres in Higher Education (ACLES 2011a, b). ACLES introduced a model for a language examination – the CertACLES model (ACLES 2011b) – that would be followed by all higher education institutions belonging to the organisation and that was intended to allow for the assessment of communicative competence with a standard and comparable framework which all member institutions needed to adhere to. This framework is solid enough to provide for a standard tool for measuring language ability while allowing each individual university to adjust their exam to meet the needs of their environment. Each university is therefore in charge of designing its own individual exams, which have to comply with the framework but have to take into consideration each particular context, not only in terms of test construct and specifications, but also in terms of administration dates and frequencies. CertACLES exams measure the four skills – reading, writing, listening and speaking – and give equal weight to each section. They were officially recognised by the Spanish Conference of Rectors in 2011 (CRUE, 2011) and by the Regional Government in Valencia in 2013 (DOGV, 2013).

The profile of the candidates was expected to consist mainly of a student population, although students and staff from other universities in the area who do not offer their own language proficiency tests were also expected to take part. To further specify the

profile of the candidates, information was requested as part of our survey and the results were as follows<sup>1</sup>:

### III.1.1. Age range

Fifty-seven per cent of participants were in the 18 to 30 age range, which would correspond to an examination developed by a higher education institution. Interestingly enough, there is a high percentage of participants in the study who were over 30 years of age, which indicated that there was a high number of participants who were either university staff, alumni or external candidates that wanted to sit an official language exam.

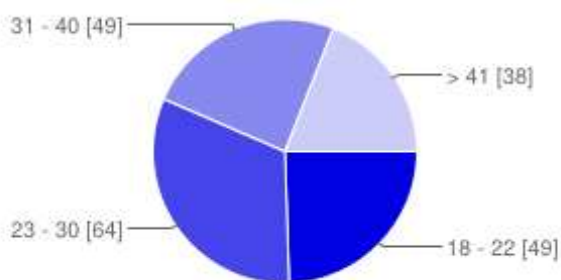


Figure 1. Age range

Table 1. Age range

Age	No. of candidates within the age range	Percentage of candidates within the age range
18 - 22	49	25%
23 - 30	64	32%
31 - 40	49	25%
> 41	38	18%

### III.1.2. Education

The results were as expected given the type of examination and the examining body (a higher education institution), the large majority of participants hold a university degree, and 15% of them have doctoral degrees.

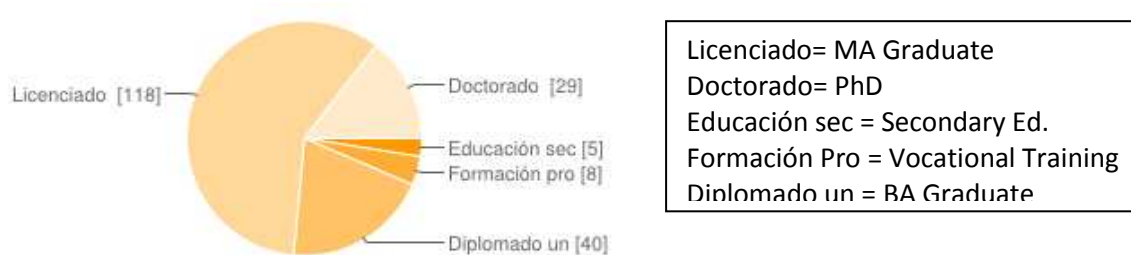


Figure 2. Education

Table 2. Education

Education	No. of candidates	Percentage of candidates
Secondary education	5	3%
Vocational training	8	4%
University degree (BA)	40	20%
University degree (MA)	118	59%
PhD	29	15%

### III.1.3. Motivation for taking a language exam

Since we were analysing students' perceptions and FLA when confronting a speaking test, we were interested in knowing their reasons for taking the exam. Their reasons would indicate the orientation of their motivation, either integrative or instrumental



(Gardner and Lambert, 1959), and would thus help predict their degree of success. As we can see from the results, illustrated in Table 3, in 75% of the cases the motivation for taking the exam was instrumental. Only 25% of the candidates showed an integrative motivation and stated that the reason for taking the exam was personal satisfaction, which was assumed to mean travelling to other countries and meeting native-speaking people as well as learning the culture of native-speaking countries.

Table 3. Motivation

<b>Reasons for taking an exam</b>	<b>No. of candidates</b>	<b>Percentage of candidates</b>
Personal satisfaction	50	25%
Mobility grant	18	9%
Graduation requirement	38	19%
Professional projection	130	65%
Other	10	5%

### **III.2. Materials**

The main goal of our study, as stated in our introduction, was to analyse students' beliefs about their performance in the speaking section of a language proficiency exam and compare them with their actual results in the exam, in order to determine whether their beliefs were based on their actual level of competence or if they were based on other factors, such as anxiety or stress due to the particular characteristics of this section of the exam. In order to do this, we needed to examine, on the one hand, their feelings with respect to the different sections of the exam in terms of perceived difficulty and candidate anxiety and, on the other hand, the results obtained by the candidates in the actual examination. By looking at the results obtained in the speaking section of the

examination and comparing these results with those obtained by the same candidate in the other sections of the exam, we would be able to compare candidates' self-perception with their actual performance, and thus devise strategies that could be introduced into the design of the examination to reduce the negative influence of factors external to the candidates' language competence.

Accordingly, our study would initially be divided into two different steps: (1) analysing students' perceptions of their performance in the exam, and (2) analysing students' actual results in the exam.

### ***III.2.1. Analysing students' perceptions of their performance in the exam:***

For the sake of practicality, we decided to use the free tool for generating surveys provided by Google, *Google Forms*, to create our survey. This tool allowed us to design a relatively simple survey with automatic data processing and charting, but with a compatible table in Excel format to allow further modification or alternative processing of the data obtained. Likewise, the system also allowed the creation of a link to the survey that could be sent to the students' email addresses from the Language Centre's email account. The fact that the tool involved no additional costs and that it was user-friendly, only requiring a few minutes to be able to start using it, was also a key factor in our decision. Google Forms requires the individual who is designing and administering the survey to have a gmail account. This email account does not need to be the one used to send the survey to the participants – which was a question of concern for us since we did not want to use an account not belonging to the university – but it will be visible in the link sent and therefore needed to have some appearance of reputability. To achieve this, we set up a gmail account for the language centre in which not only the name of the language centre was specified, but also the initials of the university, to make the sender easily identifiable.

Before designing our survey, we had to take into consideration the characteristics of the tool we were going to use. Google Forms provides different layouts, allowing us to send respondents in different directions depending on the answer and allowing different types of question formats. The question formats provided are as follows: *Text* – open questions with short answers; *Paragraph Text* – open questions with longer answers;

*Multiple Choice* – controlled answers where one option is chosen from among several; *Checkboxes* – controlled answers where users select as many options as they like; *Choose from a list* – controlled answers in which users select one option from a dropdown menu; *Scale* – controlled answer in which users rank something on a scale of numbers; *Grid* – controlled answer in which users select a point from a two-dimensional grid; *Date* – controlled answer in which users pick a date on a calendar; *Time* – controlled answer in which users select a time of day or a length of time. Our initial intention was to use either the *Text* format or the *Paragraph Text* format, preferring the short-answer questions for the sake of conciseness. However, we also wanted to favour easy processing of the information and we realised that using this type of format would not allow the data to be processed automatically. In the end and after much consideration, we decided to use a Multiple Choice format since it limited the respondents' production and allowed for easier processing by automatically generating charts and summaries of results. In fact, Google Forms can be connected to spreadsheets in Google Sheets, and if a spreadsheet is linked to the form, responses will automatically be sent to the spreadsheet from where information is taken and automatically summarised and presented in a summary of results. For those questions in which we intended to measure a level (level of difficulty, anxiety, etc.), we used the *Scale* format, since the processing of results was similar to that of the Multiple choice format.

### ***III.2.2. Analysing students' actual results in the exam:***

The candidate's marks that were analysed belonged to the speaking exam of the CertACLES Certification paper administered in July 2014. This exam aims to evaluate the communicative competence of the candidates and the contents and construct of the exam and the marking criteria are based on the CEFR descriptors. To that end, the exam evaluates the four main communicative macro-skills, i.e. speaking, listening, writing and reading, each with a specific weight of 25% of the total score of the exam.

A candidate is considered to have reached the corresponding language level if the final mark is equal to or higher than 60% of the total possible points, provided that a

minimum of 50% of the possible mark has been attained in each skill. The marks are awarded on a scale of 0 to 10 points (100%) expressed to one decimal place:

- Between 6.0 and 6.9 points (60%-69% of total marks possible) = PASS
- Between 7.0 and 8.9 points (70%-89% of total marks possible) = MERIT
- Between 9.0 and 10 points (90%-100% of total marks possible) = DISTINCTION

The speaking test is conducted by two oral examiners, an interlocutor and an assessor, with paired candidates. The reason for choosing this task format had to do with our goal of mirroring real-life communication, while minimising anxiety and tension in the candidate. As Heaton (1988) states, interviews are adequate attempts to assess oral skills but students are not placed in “natural” speech situations and they are therefore subject to psychological tensions which will necessarily affect their performances. CertACLES exams attempt to minimise this effect by having an interlocutor and an examiner present in the interview to allow the interlocutor to focus on candidates while they speak and avoid interruptions that would occur while the interlocutor takes notes. In this way the interlocutor is responsible for conducting the interview and for giving a global impression of the overall communicative ability of the candidate, but it is the assessor who is responsible for providing an analytical assessment of each candidate’s performance. The assessor does not take part in the interaction with the candidates and is thus able to apply a detailed analytical scale with four criteria: (1) grammar, which refers to appropriate use of grammatical forms; (2) vocabulary, which measures the accuracy and the use of lexical forms; (3) Discourse management, which focuses on relevant discourse and coherence; and (4) pronunciation and interactive communication, where the focus is not only on the ability to be understandable but also the candidate’s ability to take an active part in the development of the discourse. Moreover, having two candidates taking the exam together allows for equal interaction where there is no power relationship (interlocutor/candidate), but instead a conversation between two members of the same peer group.

Since the exam aims to obtain different types of oral production in a single interview, the interview is divided into three parts:

**Part One.** Conversation between the interlocutor and each candidate. There is a set of standard questions on personal details and preferences grouped by topic (country of

origin, studies, hobbies and interests, education, travel, technology, etc.). The interlocutor can decide which questions to ask based on the responses obtained in order to elicit enough performance by the candidate for assessment within the time frame available.

**Part two.** Simple standardised rubric with minimal language input. The candidates are each given one or two photographs (B1 candidates have one picture to describe and B2 and C1 candidates are given two pictures to allow them to use more complex vocabulary for comparison and contrast). The objective is therefore to compare and contrast during an individual long turn. After each candidate has spoken, their partner is asked one question related to the topic.

**Part Three.** Conversation between candidates. The interlocutor gives some pictures to the candidates. They are asked to speak for a set amount of time and justify their opinions, speculate, express preferences and draw conclusions within the target language use defined for each level of examination. At the end of the interaction the interlocutor may ask the candidates further questions on the topic.

### **III.3. Procedure**

#### ***III.3.1. Analysing students' perceptions of their performance in the exam:***

A survey was designed with multiple choice questions on the candidates' profile and their opinion on the difficulty of the different sections (from 0 to 5, 0 being the easiest and 5 being the most difficult). Once the survey had been designed and implemented in Google Forms, we generated the link and sent it out to all the candidates participating in the June 2014 exam sessions (B1, B2 and C1 candidates). The email was sent after they had taken the examination so that their opinions were based on the same exam from which their marks were going to be analysed in our second step. Moreover, and to avoid bias in their responses, the link was sent before results were published and a deadline was established for the collection of responses, no responses being accepted if received after the publication of exam results.

### III.3.2. Analysing students' actual results in the exam:

An excel spreadsheet was designed to introduce the candidates' results for the different parts of the exam, that is, listening, reading, writing and speaking. This would facilitate the analysis of the results, and allow for an analysis of the weaknesses and strengths of the different candidates.

A spreadsheet was designed for each of the examinations (B1, B2, C1) and the structure was as follows:

**Candidate number** is the number assigned to each candidate for easier identification; **ID, Name, Surname**, are fields needed to issue the official accreditation certificate; **listening mark, reading mark, writing mark, speaking mark** are individual marks per skill, and **overall mark** is the mark obtained from the weighting of the different skills. Finally, a **register number** is provided for the certificate issued.

## IV. RESULTS

After collecting the data from the survey and analysing the results obtained by the students in the different parts of the exam that they had rated as regards difficulty, the results were as follows.

As we can see in Figure 4, for the speaking section, most of the candidates gave a rank of 3 or higher, indicating higher difficulty; in fact, 45% of the candidates ranked the level of difficulty of the speaking section as 4 or 5.

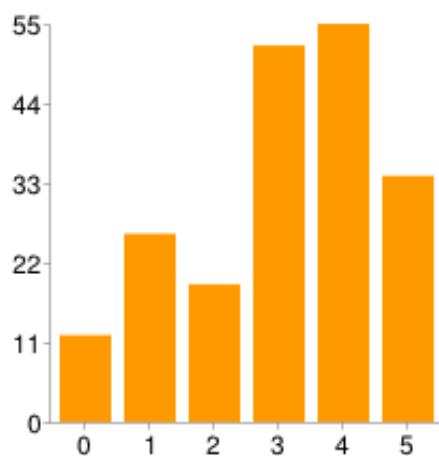


Figure 3. Perceived difficulty of the speaking section

Figure 5 shows the level of difficulty of the listening section, which was ranked higher, although only slightly so; in fact, 53% of the candidates ranked it as having a level of difficulty of 4 or 5.

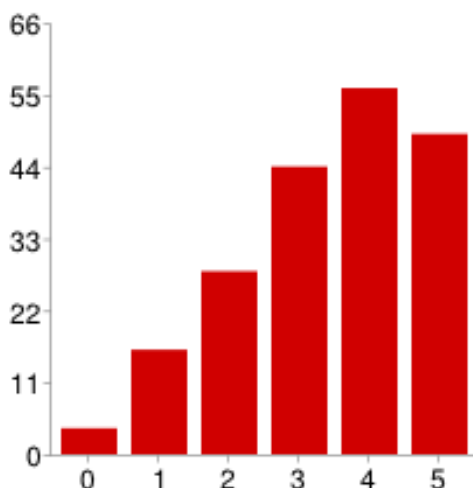


Figure 4. Perceived difficulty of the listening section

Reading was ranked the easiest, as we can see in Figure 6. 72% of the candidates ranked it between 0 and 3 on the scale of difficulty and 13% of those indicated the level of difficulty as non-existent.

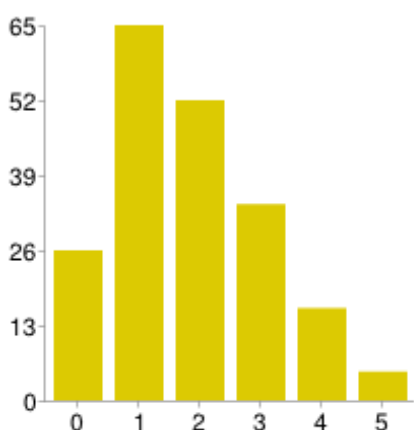


Figure 5. Perceived difficulty of the reading section

As for writing, as shown in Figure 7, it was considered a medium-difficulty section, with only 30% of the candidates ranking the difficulty of the exam above 3.

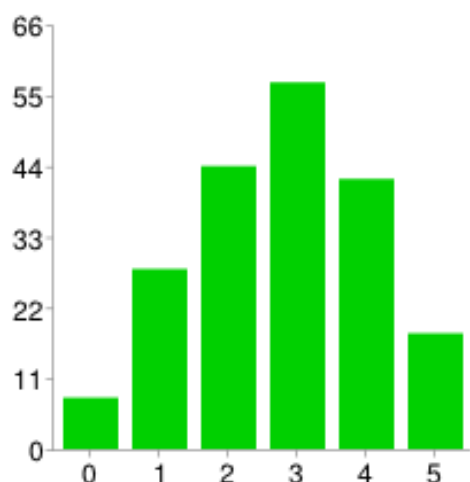


Figure 6. Perceived difficulty of the writing section

In light of these results, candidates considered the listening section to be the most difficult, closely followed by the speaking section, and by the writing and reading sections, which were far behind in terms of perceived difficulty. This contradicted our initial beliefs, since with their own reactions in the classroom and their reluctance to complete speaking and writing tasks, our students usually express more anxiety towards productive skills in the classroom and there is a higher demand for writing and speaking preparation courses, leaving reading and listening as areas that are not specifically prepared by students but are learnt or practised in general English courses.

As for the candidates' marks in the examination, which are translated in the table below, the results obtained were as follows:

As we can see in Figure 8, the vast majority of the candidates (71%) obtained a mark that allowed them to pass the exam.



### No. of pass/fail marks

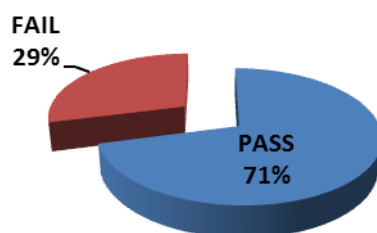


Figure 7. No. of pass/fail marks

It is also important to highlight that, as we can see in Figure 9, 210 candidates passed the examination, 122 of them obtaining either a merit or a distinction.

### No. of candidates per mark achieved

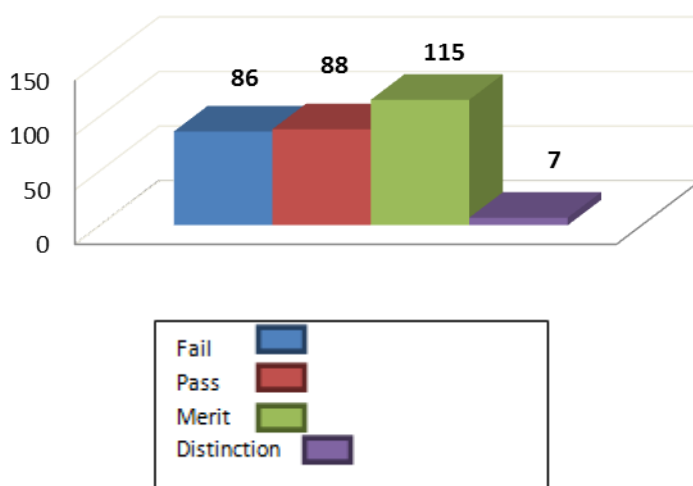


Figure 8. No. of candidates per mark achieved

Figure 10 shows the number of candidates who failed the examination per skill and level of examination. The figure shows three clusters of results indicating the candidates at levels B1, B2 and C1. As we can see, at level B1 a higher percentage of candidates failed the speaking section. In contrast, B2 candidates have higher failing rates in the

writing section, while speaking has the second best results after reading. For the C1 examination, the results are similar to B2, speaking having the second best results of all the sections of the exam.

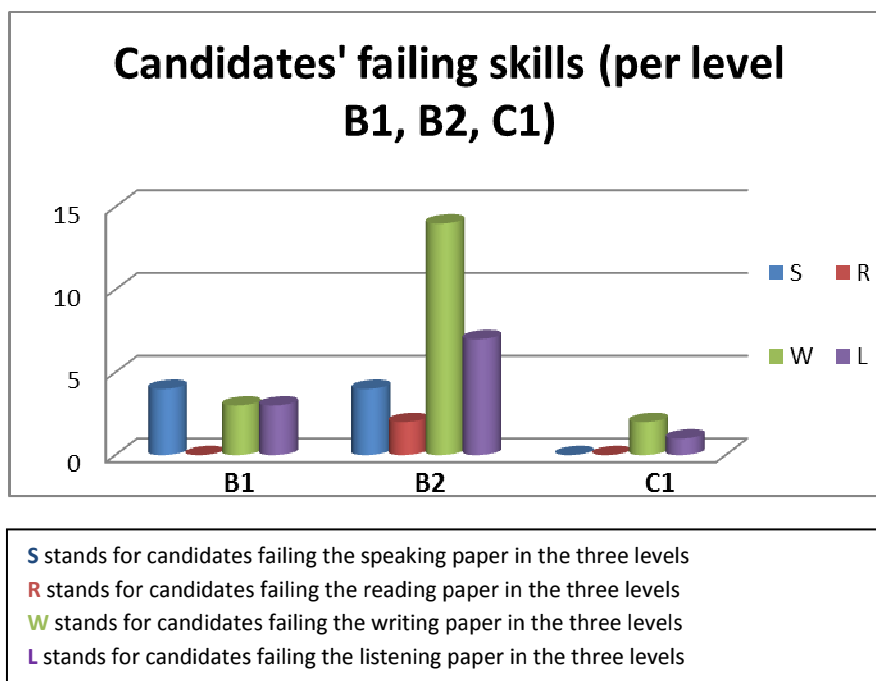


Figure 9. Candidates with only one failed skill

To further illustrate Table 10, in Table 5 we can see the number of candidates with one or more failed skills and a specification of the skills failed. It can also be observed how the results further indicate that candidates predominantly fail because of the writing paper, particularly at higher levels. The lower figures at C1 are due to the small number of candidates who failed (only 5 candidates failed the C1 examination).

Table 4. Illustrative table of candidates failed per level

	B1	B2	C1	Total number of candidates
<b>S</b>	<b>4</b>	<b>4</b>	<b>0</b>	<b>8</b>
S-W	1	8	0	9
S-R	1	1	0	2
S-L	1	2	1	4
<b>R</b>	<b>0</b>	<b>2</b>	<b>0</b>	<b>2</b>
R-L	2	1	0	3
R-W	0	0	0	0
<b>W</b>	<b>3</b>	<b>14</b>	<b>2</b>	<b>19</b>
L-W	2	3	1	6
<b>L</b>	<b>3</b>	<b>7</b>	<b>1</b>	<b>11</b>
W-R-L	1	2	0	3
S-W-L	1	2	1	4
S-W-R	1	1	0	2
S-R-L	2	0	0	2
<b>ALL</b>	<b>2</b>	<b>0</b>	<b>0</b>	<b>2</b>

## **V. CONCLUSIONS**

The results of our study show a mismatch between self-perceived efficacy and actual performance, particularly for higher levels of proficiency. In fact, although according to our survey most candidates ranked speaking as the second most difficult skill, the results of the exam show that the number of candidates who fail the exam because of the speaking section is comparatively lower. This is not the case, however, for B1 candidates, who showed a more accurate level of self-perceived efficacy, as seen in the results in Table 10. This is in line with Bandura's (1997) statement about the perception of self-efficacy, in that candidates produce less accurate assessments as they progress through higher levels of language study.

However, as stated in our results, candidates' perceived efficacy is accurate in the case of reading and listening, since they are both the easiest and the most difficult sections and the candidates' mark reflects this as being so. The greatest mismatch is therefore in the productive skills, since writing is considered a medium-difficulty skill and it is in fact the skill in which the candidates' performance is ranked lower. Speaking is perceived as a high-difficulty skill but this difficulty is not reflected in the candidates' results, as few of them fail because of the speaking section. Therefore, candidates' perception of their efficacy in the speaking section does not seem to correspond to their actual ability, which would indicate that their perceptions are indeed influenced by factors that are external to their actual performance. FLA comes into play and, consequently, modifications in the test process should be arranged to reduce anxiety for candidates. Some of the modifications suggested would be the following:

- Organising exams with paired interviews whenever possible in order to avoid relationships of power with the examiner and thus reduce stress.
- Facilitate the presence of an assessor whenever possible in order to allow the examiner to act only as the interlocutor.
- Individual arrangements for candidates to facilitate schedules and allow them to choose the time of day at which they would feel more comfortable taking the test.
- Flexible examination dates, to eliminate stress in candidates who have conflicting commitments (academic, professional, family-related, etc.).

- Preparation time and warm-up questions to allow them to feel more at ease with the topic, as well as get to know both the interviewer and the other candidate (in cases when the interview is paired).
- Additional prompts to facilitate discussion topics during the exam and prevent candidates from relying on their resourcefulness or imagination.
- Start and finish the interview on a positive note to improve confidence and self-image, which could then be mirrored in real-life performance.

We consider that the results of our study call for further research on factors outside the content of the exam, factors related to administration and organisation, as well as those related to individual characteristics of the candidate (personality, background, etc.), which will undoubtedly explain the difference between perceptions and actual results.

## Notes

- <sup>1</sup> The survey was carried out in Spanish to allow all participants to fully understand the questions; the titles and legends in the graphs are therefore in Spanish. Under each graph there is a representation of the information in table format and translated into English.

## REFERENCES

- ACLES** 2011a. Asociación de Centro de Lenguas de Educación Superior. 20 April 2014 <<http://www.acles.es/>>
- ACLES** 2011b. Modelo de acreditación de exámenes de acles. 20 April 2014 <[http://www.acles.es/multimedia/enlaces/9/files/fichero\\_29.pdf](http://www.acles.es/multimedia/enlaces/9/files/fichero_29.pdf)>
- Bachman, L.F. and Palmer, A.S.** 1996 *Language testing in Practice*. Oxford: Oxford University Press.
- Bachman, L.F.** 1990. *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Baddeley, A.** 2007. *Working memory thought, and action*. Oxford: Oxford University Press.

- Bandura, A.** 1997. *Self-Efficacy: The Exercise of Control*. New York: Freeman.
- Brown, D.H.** 2000. *Principles of language learning & teaching*. New York: Longman.
- Carroll, J.B., and S.M. Sapon.** 1959. *Modern language aptitude test*. New York: The Psychological Corporation/Harcourt Brace Jovanovich.
- Celce-Murcia, M., and E. Olshtain,** 2000. *Discourse and context in language teaching: A guide for language teachers*. Cambridge University Press.
- Cohen** 2003. "The learner's side of foreign language learning: Where do styles, strategies, and tasks meet?" *International Review of Applied Linguistics in Language Teaching*, 41(4), 279-291.
- Conferencia De Rectores De Las Universidades Españolas (CRUE).** 2011 (8 de septiembre). Propuestas sobre la acreditación de idiomas. 25 April 2014 <[http://www.acreditacion.crue.org/export/sites/Acreditacion/doc/Propuestas\\_acreditacion\\_de\\_Idiomas-AG8-09-2011.pdf](http://www.acreditacion.crue.org/export/sites/Acreditacion/doc/Propuestas_acreditacion_de_Idiomas-AG8-09-2011.pdf)>
- Conway, A., C. Jarrold, M. Kane, A. Miyake, and J. Towse,** (Eds.). 2007. *Variation in working memory*. Oxford: Oxford University Press
- Diari Oficial De La Comunitat Valenciana (DOCV).** 2013. Orden 93/2013, de 11 de noviembre, de la Consellería de Educación, Cultura y Deporte, por la que se modifica el anexo único denominado certificados y diplomas que acreditan la competencia en lenguas extranjeras del Decreto 61/2013, de 17 de mayo, por el que se establece un sistema de reconocimiento de la competencia en lenguas extranjeras en la Comunitat Valenciana y se crea la Comisión de Acreditación de Niveles de Competencia en Lenguas Extranjeras. 17 May 2014 <[http://www.docv.gva.es/datos/2013/11/13/pdf/2013\\_10793.pdf](http://www.docv.gva.es/datos/2013/11/13/pdf/2013_10793.pdf)>
- Dörnyei, Zoltán.** 2009. "Individual differences: Interplay of learner characteristics and learning environment". *Language Learning*, 59.s1, 230-248.
- Fulcher, G.** 2003. *Testing second language speaking*. New York: Pearson Longman.
- Gardner, R.C., and Lambert, W.E.** 1959. "Motivational variables in second-language acquisition". *Canadian Journal of Psychology/Revue canadienne de psychologie*, 13 (4), 266-272.

- Hardison, D.M.** 2014. Changes in Second-Language Learners' Oral Skills and Socio-Affective Profiles Following Study Abroad: A Mixed-Methods Approach. *Canadian Modern Language Review*, 70 (4), 415-444.
- Heaton, J.B.** 1988. *Writing English language tests* (2<sup>nd</sup> ed.). Harlow: Longman.
- Horwitz, E.K., Horwitz, M.B., and Cope, J.** 1986. "Foreign language classroom anxiety". *The Modern Language Journal*, 70 (2), 125-132.  
<http://dx.doi.org/10.1017/CBO9780511733017>
- Jee, M.J.** 2014. "Affective factors in Korean as a Foreign Language: anxiety and beliefs". *Language, Culture and Curriculum*, 27 (2), 182-195.
- Luoma, S.** 2004. *Assessing speaking*. Cambridge University Press.
- O'Malley, J.M. and Chamot, A.U.** 1990. *Learning Strategies in Second Language Acquisition*. Cambridge: Cambridge University Press.
- Pimsleur, P.** 1966. *Pimsleur language aptitude battery (PLAB)*. New York: The Psychological Corporation.
- Phillips, E.M.** 1992. "The effects on language anxiety on students' oral test performance and attitudes". *The Modern Language Journal*. 76 (1), 14-26.
- Sparks, R.L., Humbach, N., Patton, J.O.N., and Ganschow, L.** 2011. "Subcomponents of second language aptitude and second language proficiency". *The Modern Language Journal*, 95 (2), 253-273.
- Stephenson, J., and Hewitt, E.** 2011. Los desafíos de realizar la investigación en el aula: el caso de un estudio réplica sobre la ansiedad en el aprendizaje del inglés. *Dedica. Revista de educação e humanidades*, 1, 365-382

Received: 3 December 2014

Accepted: 22 July 2015

*Cite this article as:*

**Pérez-Guillot, C. & Zabala-Delgado, J.** 2015. "Comparing candidates' beliefs and exam performance in speaking tests". *Language Value* 7, 22-44. Jaume I University ePress: Castelló, Spain. <http://www.e-revistas.uji.es/languagevalue>.

DOI: <http://dx.doi.org/10.6035/LanguageV.2015.7.3>

**ISSN 1989-7103**

**Articles are copyrighted by their respective authors**