



UNIVERSITAT DE VIC
UNIVERSITAT CENTRAL
DE CATALUNYA

Final Dissertation

A Deep Analysis of the English Core Competences Test in Catalonia: What
Do Teachers Think About it?

Laia VILAR SABATA

2017-2018

Tutor: Mireia Canals

Infant Education and Primary Education, minor in English

Facultat d'Educació, Traducció i Ciències Humanes

Universitat de Vic-Universitat Central de Catalunya

Vic, May 2018

A Deep Analysis of the English Core Competences Test in Catalonia: What Do Teachers Think About it?

Laia Vilar Sabata

Abstract

English has emerged as a shared language used to establish communication in a global world. Thus, schools around the world have faced the need of teaching this language. In Catalonia, within the core competence tests, an English test is carried out by the end of Primary School to acknowledge the students' level of English at a regional level. Thus, the current study is aimed to deeply analyze how the English Core Competence Test has evolved since the first edition in 2009. In order to do that, the four skills are examined. Moreover, a key aspect that this article aims to analyze is what teachers think about these tests and what they would improve. The study closes with conclusions and future lines of research.

Keywords: assessment, core competences, foreign language teaching, Primary Education

Resum

L'anglès s'ha convertit en una llengua compartida utilitzada per comunicar-se en un món global. És per això que les escoles arreu del món s'han trobat amb la necessitat d'ensenyar aquesta llengua. A Catalunya, dins les proves de Competències Bàsiques, es duu a terme un test d'anglès a final d'Educació Primària per tal de conèixer el nivell d'anglès dels estudiants a nivell regional. El present estudi, doncs, pretén analitzar profundament la manera com el test d'anglès de Competències Bàsiques ha evolucionat des de la seva primera edició al 2009. Per tal de fer-ho, s'ha examinat com es tracten les 4 habilitats lingüístiques. A més, un altre aspecte clau que aquest article pretén analitzar és què pensen diverses mestres d'anglès sobre aquests tests i què millorarien. Per acabar, l'estudi es tanca amb conclusions i futures línies de recerca.

Paraules clau: avaluació, Competències Bàsiques, ensenyament de llengua estrangera, Educació Primària

Resumen

El inglés se ha convertido en la lengua compartida utilizada para comunicarse en un mundo global. Es por ello que las escuelas en todo el mundo se han encontrado con la necesidad de enseñar esta lengua. En Cataluña, dentro de las pruebas de Competencias Básicas, se lleva a cabo un test de inglés a final de Educación Primaria para conocer el nivel de inglés de los estudiantes a nivel regional. El presente estudio, pues, pretende analizar profundamente la forma en que el test de inglés de Competencias Básicas ha evolucionado desde su primera edición en 2009. Para ello, se ha

examinado cómo se tratan las 4 habilidades lingüísticas. Además, otro aspecto clave que este artículo pretende analizar es qué piensan varias maestras de inglés sobre estos tests y que mejorarían. Por último, el estudio se cierra con conclusiones y futuras líneas de investigación.

Palabras clave: evaluación, Competencias Básicas, enseñanza de lengua extranjera, Educación Primària

1. Introduction

In recent years, in a world in command of globalization, English has become a shared language used to establish relationships among the citizens all over the globe. Since nowadays, the need for a global language is particularly appreciated by the international academic and business communities, the adoption of a single lingua franca has been most in evidence (Crystal, 2003). This is the reason why, English as a second or foreign language is so important in countries where English is not the mother tongue, and this is also why effective English teaching has become a priority for most educative systems (Albaigés and Ferrer-Esteban, 2012). Actually, as Seidlhofer (2005) underpins, despite being welcomed by some and deplored by others, it cannot be denied that English functions as a shared language and as a global lingua franca. Thus, in today's world, bilingualism and multilingualism are the norm rather than the exception (Harris and McGhee-Nelson, 1992). In this context, where English as a lingua franca is so important, schools have noticed the importance of teaching English since the first moment children enter school so as to widen their opportunities for the future. Families have also been putting a strong pressure on schools for what refers to English language, and thus, schools have had to define themselves towards language teaching and offer a good English project to respond to this general world needs. In this global world, where learning languages has become a central issue, the countries in Europe have notified the need to have competent citizens in English language, and therefore, they have developed learning plans and methods as well as assessment or acknowledgement tools to know their students' English level. Regarding Catalonia, it is one of the territorial areas in which the learning of a foreign language throughout the entire primary school is guaranteed (Albaigés and Ferrer-Esteban, 2012). However, taking into account a comparative analysis on the learning of foreign languages in the educational system, in Catalonia we observe the weakness of English language knowledge from a comparative perspective: only 32,1% of the population can understand English language (Albaigés and Ferrer-Esteban, 2012).

Nevertheless, regarding language assessment, it is important to highlight that it was not until the late nineteenth century when this need emerged and when assessment appeared as a domain of interest for researchers and educators. And it was about the twentieth century when standardized tests were initiated (Gould, 1996). From then on, many standardized tests to know children's English level have been developed but since they are relatively new, many of them still need improvement as well as the way the language is introduced. Actually, as Albaigés and Ferrer-Esteban (2012) state, teaching foreign languages is a challenge for educational systems that often do not have the right infrastructure to respond to the demands it represents.

In the Catalan educational system, it was not until 2009 when the first standardized tests were initiated by Consell Superior d'Avaluació. Since the tests were first implemented, by the end of Primary Education, children need to pass several exams on the main subjects they have been taught since they started 1st of Primary: Catalan, Spanish, English and Maths. Nowadays, a total amount of 2246 schools in Catalonia take part in these tests, which evaluate the competences and the basic understanding that children need to have acquired by the end of Primary to be able to follow with complete normality the curriculum planned for Secondary. Nevertheless, the results of these tests do not take effect neither in the final evaluation of the students nor in the decisions to pass to the following course.

These tests are standardized tests, which, according to Nazari (2012), are identified as the regularization of procedures and instruments and the statistical analysis of results. Thus, with these tests, the Educational Department and every Primary Education institution can have accurate information available about how students are doing in a very important moment of their schooling, since the data is provided by an homogeneous measuring instrument.

Nowadays, after several modifications, the English test consists of 5 exercises. Children are asked to listen to two texts, read two other texts and answer multiple choice questions for all of them. Finally, they are also asked to write a text, in which the topic is given. On the first page of the aforementioned test, children have the instructions to carry it out, which are in Catalan. In these instructions, they are told how to answer the questions and the aspects that they need to take into account when answering them. The first part of the test consists of two listenings. They are listened 3 times and children can have a look at the questions they need to answer (for 45 seconds) before listening to them. Actually, Vandergrift (2007) states that listening comprehension lies at the heart

of language learning, but it is the least understood and least researched skill. The second part consists of two readings, for which children also need to answer some multiple choice questions. Both texts that children need to read are one page long. Finally, the third and last part of the test consists of a writing, in which children need to write between 40 and 50 words about a given topic. In order to make it easier for them to start writing, they have clues on what to write and also on what to write in each part of the text.

Since this article addresses English language assessment in Catalonia, it aims to deeply analyze these standardized tests since its first implementation and to know and analyze the opinion of several teachers on what these tests evaluate and how they do it. Thus, the article presents what are the core competences, as well as the method used to carry out the research, and the results of both the analysis of the test and the analysis of the interviews to the Primary English teachers.

2. Assessing language through competences

In recent years, the acquisition of competences has become a central issue in Primary education. Historically, competences have emerged in education as an alternative to address the shortcomings of traditional pedagogical models and approaches, such as behaviorism, cognitivism and constructivism, although they rely on some of its theoretical and methodological approaches; however, this is done with a new perspective, with a change in logic, moving from the logic of content to the logic of action (Tobón, Pimienta, García, 2010). In this sense, as Westera underpins, “the traditional emphasis on factual knowledge as such, does no longer meet the requirements of a changing society” (2001:1), and thus we need to talk about something that goes beyond the skills: the competences. As Westera (2001) claims, a competence is a highly valued qualification that accounts for the effective use of one’s knowledge and skills in a specific, usually difficult or complex context. In the same line, Entwistle quoted in Gordon et al. (2009) points out that what we learn depends on how we learn and why we have to learn it. Zabala and Arnau (2007) consider that evaluating through competences lies in evaluating the processes followed to confront and solve certain problem situations. Therefore, As Gordon, Halasz, Crawczyk, Leney, Michel, Pepper, Putkiewicz and Wisniewky (2009) state, the focus is now on acquiring learning outcomes for professional and private life placed in a perspective of lifelong and life-wide learning.

Actually, as a matter of fact, assessment is becoming a big challenge for anyone engaged in the field of teaching. In the late twentieth century it has assumed a singular importance throughout the world. Focusing specifically on language teaching, Yip and Matthews (2006) argue that when assessing students, language dominance must be related to underlying competence (i.e., knowledge of language) to be compatible with a competence-based theory of transfer. Moreover, recently, the concept of dynamic assessment has emerged with plenty of strength. As Nazari (2012) points out, there is a huge difference between evaluating outcomes only and using assessment as part of instruction, as a dynamic process. In this sense, many authors stand up for the idea of an assessment based on competences rather than on contents, an assessment that helps the student in his learning process and which aim is not to give a mark but to guide the student and allow their self-reflection and self-regulation.

On the other hand, referring specifically to language tests, as Bachman (2001) claims, the primary consideration in developing a language test is the use for which it is intended. Thus, for the author, it would make sense to ask the question “What makes a given test useful for its intended purpose?” or “What are the qualities that we need to consider when evaluating the usefulness of a given test?”. Bachman (1991:681) also adds that

“in order to make inferences about levels or profiles of ability, or predictions about capacity for using language to perform future tasks in non test language use contexts, we need to demonstrate two kinds of correspondences: (a) that the language abilities measured by our language tests correspond in specifiable ways to the language abilities involved in non test language use, and (b) that the characteristics of the test tasks correspond to the features of a target language use context.”

Moreover, Shohamy, Or and May (2017) add that the field of language testing is traditionally viewed as consisting of two major components: one focusing on the “ what,” referring to the constructs that need to be assessed, and the other component pertaining to the “ how”, which addresses the specific procedures and strategies used for assessing the “ what.” Thus it is not only important the purpose of the evaluation but also the way how it is carried out. Therefore, it can be claimed that what the research suggest is that when assessing the language level of students, it has to be taken into account what is the objective, how is it going to be tested, and make sure that it responds to real needs in real contexts, and that what children are learning is going to be useful in their life.

3. Method

The educative paradigm in which the research has been developed is the interpretative one which, as Starman (2013) states, emphasizes subjective experiences and the meanings they have for an individual. In the current research, the aim was to understand and analyze how the English Core Competence Test has evolved since it was introduced nine years ago, taking into account both the test and the opinions, points of view and meanings that several English teachers have. Because it was the critical point of focus in the study, it is also a qualitative research.

In order to carry out the study and achieve the objectives, the English Core Competence Test from 2009 to 2017 has been deeply analyzed. Moreover, since teachers' opinion about these test was also an aspect to be analyzed, some interviews have been done to four teachers from four different schools in Osona.

Table 1: Teachers participating in the research

Teacher	School	Teaching grades
Teacher 1	Escola Xuriguier, public school (Centelles)	1st-6th grades
Teacher 2	Escola l'Era de dalt, public school (Tona)	1st-6th grades
Teacher 3	Escola Ildefons Cerdà, public school (Centelles)	4th-6th grades
Teacher 4	Escola Vedruna, state-funded school (Tona)	4th-6th grades

Thus, to do so, the different instruments that have been used are the tests and the interviews to the four English teachers. These instruments have permitted to analyze the evolution that the English Core Competence Test of 6th of Primary has followed since the first test in 2009 and also to know some English teachers' point of view.

3.1 Instruments

To carry out the research and gather the appropriate data, an interview has been created, the same for all teachers that participate in the study.

Table 2: General questions

Do you think that the time given to answer all the questions is appropriate?
What do you think about the amount of activities that students need to solve? Are them enough, too much...?
Do you think the explanations of the tasks should be written in Catalan or in English?
Are the instructions clear enough?
Do you think that the four skills are given the same importance?
Does each activity only require the use of one skill or more than one?
Would you include speaking tasks in these tests? how?

Table 3: Listening

What do you think about the listenings length?
Do you think that listening to the text three times is appropriate? How many times would you make them listen to it?
Is it important for you that students are able to have a look at the questions before listening to the speech?
Do you think that answering multiple choice questions correctly guarantees children's understanding of the whole speech in listenings?
What do you think that is more tested, memory skills or comprehension?
What kind of questions would you do?
What kind of listenings would you pose? (dialogues, descriptions, songs...)
Do you teach children about how to listen?

Table 4: Reading

What do you think about the length of the texts?
Do you think that multiple choice questions guarantee children's understanding of the whole text? Why?
What kind of questions would you do?
What kind of texts would you propose?
Do you think the vocabulary is accessible for all children?

Table 5: Writing

What would you evaluate in writing?
What importance do you give to orthography?
Do you think children should choose between different options of writing? Why?
Is it important for you giving an example of writing to the children? Why?
Do you think children should be told what to write in each part of the writing? Why?

In order to later analyze the data gathered, the interviews were recorded. The main aim of the interviews was to see what teachers think about the tests in general and more specifically, about how the different skills are assessed in these tests. In order to analyze the interviews, a table of analysis, in which the different aspects on what teachers were asked to give their opinion are put together and compared, has been designed. In this way, some comparisons among all the teachers words have been established and some conclusions have ben drawn. Finally, another table of analysis to know the evolution of the tests has also been created. In this way, every item has been compared and some conclusions have emerged.

4. Results

The results of the current study have been organized in two different blocks. First, the structure of the test is deeply examined and taken into account. After, several teachers' point of view expressed through an interview is analyzed and compared.

4.1 English Core Competence Test

General aspects of the test

General aspects of the tests have not changed much since the first test in 2010. However, there are some aspects that have evolved. The first two years, the test was 45 minutes long but as years went by, they incremented the length (60 minutes in 2013). However, the last three years, the time to carry out the test has not been specified in the instructions, which have always been in Catalan. Regarding the skills tested, at the beginning only listening and reading were tested. It was in 2013

when writing was added. However, the number of texts in the listening and in the reading parts, have always been the same, two.

Listening

The length of each listening has been different every year: it has never exceeded 2 minutes and never has been less than 1 minute long. The kind of texts have always been very different, and different kinds of texts in the same test have been suggested (dialogues, instructions, phone calls, radio programs...). Regarding their instructions, at the beginning, the two first years of the test, they were in Catalan. Nevertheless, in 2012, it changed and they started to be in English. The same happened with the times given to listen the texts, which the first year were only two but from then on, one more time to listen was given, so children could listen to the text 3 times.

Children have always been given some time to read the questions before listening to the texts. Nevertheless, the first year it was not specified in the test, the second one children had 24 seconds, and from then on, they had 45 seconds. The number of questions asked in both listenings has always been between 14 and 15, except the first year, when there were only 7 questions. Nevertheless, the number of options in the multiple choice has always been the same, three. One aspect that has been detected is the fact that almost half of the questions asked in this listening part are literal, so children do not need to understand to be able to answer, they can answer only using memory and recalling exactly what has been said. On the other hand, the first two years, more visual aids were given in the multiple choice questions. However, some years after, they decreased and in the most recent tests there have been only one or two per listening, or even none. Finally, as for the written answers asked in the listenings, there is not a pattern followed along the years, indeed, some years there was one and some others not.

Reading

As far as the reading is concerned, its length has varied through the years. When the test was first implanted, the length of each reading was of half of a page. However, from the third year of the test on, every reading has been one page long. In this sense, the length of the readings has followed an increase. Besides, as in the case of the listenings, the kind of texts have always been very different, and different kinds of texts have been proposed (descriptive texts, emails, informative pamphlets,

advert...). The number of questions asked in each reading has always swung between 8 and 12, although almost always there have been 12 questions. Related to this, the number of options to choose in the multiple choice questions has changed over the years, since the two first years there were 4 options and from then on there started to be 3. The kind of questions asked, as in the listenings, have been found to be, almost half of them, literal, so that children do not need to understand the text to be able to answer them. On the other hand, in general, children have not been given visual aids, although in two of the tests you can find one and three questions with a visual aid. Moreover, children have not usually been asked to write answers, but only to circle the correct option. However, there were three tests in which they were asked to do so, even though it was a short answer that could be done only with a word. Finally, it is important to highlight that during the first three years of these tests, there was no vocabulary support, but from the fourth year on, children started to be given a bank of difficult words (with the translation into Catalan) to help their understanding.

Writing

Writing was introduced after 3 years, so there was no writing in these tests until 2013. At the beginning, only between 30 and 45 words were asked to write, and if there were less, students were penalized. Then, after 2015, children started to be asked to write between 40 and 50 words, which is also the amount of words that they need to write nowadays. Regarding the importance given to orthography, the first year and also the two last years of the test, it was said that it should be avoided making orthography errors. However, the years in between, it was not specified in the test, so children were not warned about any orthographic requisite to pass the test. On the other hand, in all the tests students were given a context to start writing. However, they did not have any example of writing similar of what they were being asked to write but they had clues about what to write and clues on what to write in each part of the test, which could help them starting to build their ideas and do not experiment the fear of the blank paper. Nevertheless, it is important to highlight that they only had clues on what to write in each part of the test the last three years, not the first two ones.

Valuation

In 2013 children were asked to answer two questions: they needed to value the difficulty of this test and say if they had liked doing English activities. This was a kind of valuation that they needed to

do in order to give feedback to the people in charge of the design of the test. In 2014 children only needed to value the difficulty of the test and all the other years, children were not asked to value the test.

4.2 English teachers' advice

Analyzing and comparing the four interviews that have been done to the teachers, it can be claimed that, in general, they all feel that the tests are too difficult and too long. Deeply analyzing the interviews it can be claimed that the teachers have different points of view on the different aspects they were asked about.

First, examining what they said in the interviews, and more specifically what concerns to the length of the test and the amount of activities that children need to do, in general they agree in that it is too long and that less activities would proof children's level. However, regarding the time given to carry the whole test out, they all think think that it is enough, except from one teacher who thinks that it is too little. As far as the instructions are concerned, teachers have different opinions on the language they think the instructions should be. Two of them agree in that they should be in English if what we want to test is English, but there are two who think that if children need to understand what they need to do, it is better to have them in Catalan. However, they all think that the instructions are clear enough to be understood, but there is one of the teachers who thinks that are too long. As for the different skills that are tested, they agree on that there is not given the same importance to the four skills (listening, speaking, reading and writing), and in general, all of them except one, agree on that more importance should be given to oral and aural skills. However, there is a teacher who thinks that speaking should be tested later, because she considers that children are too young to be able to produce output. On the other hand, all of them agree on that in each activity, more than one skill is tested.

Listening

Regarding the listening part in the test, teachers do not share the same opinion on what concerns to the length. Two teachers think that the length is fine, one thinks that the listenings are too long and another one thinks that what really matters is not the length but the topics that these listenings are about, so they should be appropriated for children's age. Nevertheless, they all agree in that

listening to them 3 times is a good number, so it is enough and besides, they add that having the chance to look at the questions in advance is basic.

In reference to the type of activities, which are multiple choice questions, two teachers point that comprehension cannot be guaranteed with this kind of questions, while two think that they are good as long as they are not too literal. Similarly, Haney and Evans (1999) state, multiple choice tests may or may not be dynamic depending on the way they are designed. Nevertheless, there is one teacher who strongly claims that multiple choice can sometimes be confusing and make children have doubts and thus, that there should be different kinds of exercises so that every child had some in which they could perform well and feel comfortable with. Taking that into account, most of the teachers think that what is tested in these exams are both comprehension and memory (if the questions are literal), but there is one who claims that it is, most of all, comprehension. This is the reason why, all of them except one (who would keep posing multiple choice questions) would pose different kinds of activities in which children would need to perform differently. On the other hand, the kind of listenings suggested, they all think that they should be appealing for children and since there are two, very different between them. Finally, regarding their performance in the classroom, they all said that they work on listenings in the classroom but that they find difficult to teach strategies, so it is hard for them to teach how to listen.

Reading

Most of the teachers think that the readings are too long. However, there is one teacher who points out that the length is not the problem, but the fact that children need to read two texts, and that one would be enough to see if they understand. Moreover, like in the listening, there are two teachers who claim that comprehension can not be tested with multiple choice questions, while two others stick with the idea that it depends on how the questions are designed. Like in the listening, also, all teachers except from one (who would keep the multiple choice), also would like to have different kinds of activities and not only multiple choice ones, so as to see how children perform in different ways. As for the topics of the texts, two teachers think that the topics proposed are good enough. However, two of them maintain that the topics are not appropriate for children, so they should be attractive for the students and take into account their age. Related to this, they think that the vocabulary of the texts is sometimes accessible and sometimes too difficult, so it depends on the year.

Writing

Regarding writing, there is one teacher who strongly claims that it should not be evaluated in sixth grade, since children are still too young and what they should work most are listening and speaking. However, all the other teachers think that writing can be included in these tests. Moreover, they suggest that different things should be taken into account when correcting the writings: coherence, cohesion, adaptation to the context, grammar... In general, in Primary, they do not give a lot of importance to orthography but there is one teacher who strongly defends that there have to be some limits, some minimum. Furthermore, all the teachers agree in that it would be good to have different options of writing, so that children could choose in which one they feel better. They also agree in that it would also be good to have an example of a writing like the one children need to do, so they could have a model to take as a good example. Finally, the teachers also think that it would be really positive that children had clues on what to write in each part of the text, so that they would not have the fear of the blank paper.

5. Discussion

Torrance and Pryor (1998) claim that teachers come into their classrooms without the knowledge of how to develop useful testing instruments to monitor student progress, instead they are only able to analyze outcomes and not the process. However, this is not surprising if we take into account that there is very little attention paid to assessment in most teacher training programs (Edelenbos and Kubanek-German, 2004). Moreover, specific diagnostic tests to check children's abilities are not easy to design. Actually, Alderson et al. (1995) quoted in Alderson (2005), state that it is difficult to diagnose precisely strengths and weaknesses in the complexities of language ability. Nevertheless, many teachers, consider this kind of tests, so final assessment tests, the framework for their assessment (Nazari, 2012).

The English competence test in Catalonia has substantially evolved since its implementation in 2009, which is something that the teachers who have been interviewed have proved. As it has been observed, through the years, there has been an increase of the length of the test, so more minutes are now given to take the exam. This is the reason why most of the teachers interviewed think that the time given is enough. However, they still consider that even though there has been an increase of

the time given to ask all the questions, the test is too long in terms of number of activities and that less exercises of each skill would be enough to test if children have acquired the competences or not. As for the instructions of the test, they have always been in Catalan, which is something that two teachers totally disapprove and that two that think that is correct, so as to make sure that children understand what they have to do. However, as Bachman and Palmer (1996) claim, it is good to have general instructions, so an explicit statement in which the parts of the test are summarized, as well as the task types, the number of tasks and the time allocation with a clear indication of the relative value of each.

The skills that are tested in these exams, are only listening, reading and during the last years there has also been writing. The teachers interviewed agree on that all the skills are not given the same importance in the test and three of them also agree on that listening and speaking should be given most of the importance. The listening texts in the test have never exceeded 2 minutes, which is something that two of the teachers approve. However, there are two who find them rather long. The two kind of texts suggested to listen have all the years been very different between them, which is something that all the teachers value, a part from suggesting topics meaningful for the children.

Regarding the instructions of the listenings, at the beginning they were in Catalan, but then, they changed and started to be in English, which most of the teachers value. The times given to listen the texts also changed through the years, so at the beginning children could only listen to them twice and from then on, three times were given. Listening to the texts three times is something that all teachers think that is appropriate. They all value the fact of having time to take a look at the questions before starting to listen, which is something that since the first test, although the seconds given to do it have changed, has always been like this.

The type of activities suggested have always been multiple choice questions, which have been around 14 and 15, except for the first year of the test. So referring to the multiple choice activities, whilst there are two the teachers who consider that comprehension is not ensured and that many questions can be answered without having understood the meaning of the text, there are two other teachers. Fulcher and Davidson (2007) state that when we pose multiple choice questions, perhaps we presume our students are accustomed to it from prior instruction; perhaps we can make no such presumption. This means that maybe, children who have been trained to solve multiple choice

activities may find easier to answer some questions even though their foreign language level is not higher than the one other students have.

Referring to reading, the two readings' length has not always been the same: there has been an increase of their length, which recently has usually been of a whole page. This is something that most of the teacher disapprove, so they think that a whole page reading is too long. There is one teacher who strongly claims that maybe it is not a length problem but the fact that children need to read two texts, and that is too much. Regarding the kind of texts suggested in these tests, in general they have always been very varied and covering different topics. In this case, two of the teachers interviewed think that the texts proposed are good enough while two consider that the topics should be more appropriated to children's age, so they think that there is a lack of meaningfulness. The activities asked to solve, as well as in the listening, have always been multiple choice questions. Taking that into account, all teachers except from one would like that children could have a more varied range of activities, and not only multiple choice questions which, in their opinion, many times do not guarantee comprehension since they are too literal and they do not require the understanding of the text. Regarding the vocabulary, the last years of the test there has been a bank of words supporting primary language learners' reading, and this is something that teachers value, since they think that the vocabulary of the texts has not always been accessible.

To what writing concerns, in general, teachers have found that writing is one of the most well designed activities in these tests, even though they think it is not the most important skill to be tested at this age. Actually, Nola (2002) describes testing writing as an undeniably difficult procedure. After a few years since writing was first introduced in the test, children were asked to write between 40 and 50 words, which most of the teachers found rather too few. Moreover, the beginning, a lot of importance was given to orthography, and this is something that all the teachers disapprove: they think many other aspects should be taken into account. On the other hand, teachers found really good the fact that children have a given context and clues on what to write, so that they do not have the fear of the blank paper. In this sense, the test had characteristics and features that the teachers appreciated.

6. Conclusions

The English standardized tests established by Departament d'Ensenyament have evolved since its first implementation in 2009. However, all the changes made in the design of the test have been substantial and only some aspects have been reformulated. Thus, the most important change since its implementation, has been the addition of the writing skill, which was not tested at the beginning.

Along the 9 years of the test, teachers have developed a strong opinion about it. This study was aimed to follow the evolution of the test and to discover these teachers' opinion towards it. The results displayed by the study suggest that although there are some aspects of the test that teachers approve, there are many others they think that would need improvement. Thus, it can be claimed that most of the teachers that have been interviewed have a dynamic approach about assessment, and they find that multiple choice activities cannot always guarantee comprehension, so other kind of activities should also be included in these standardized tests. Moreover, teachers realize that the four skills are not given the same importance and that actually, listening and reading are the most valued ones. So, although the test is set out in this way, it has been found that teachers would rather include speaking and give more importance to the oral and aural skills, so the ones that are more used in real life. It can also be concluded that another aspect that teachers value a lot are the topics of the texts: they would like the topics of the texts, both the texts of the listening and the texts of the reading to be meaningful for the students and appropriate for their age.

Finally, a line for further research could be to investigate how the results of these tests have evolved and make a study about the criteria followed to design the test so as to see if there is a relationship between these two factors. Besides, it could also be interesting to study if there has been an increase of the level demanded and how this level is situated towards the Common European Framework (CEFR).

7. Bibliography

Albeigés, B., and Ferrer-Esteban, G. (2012). *L'èxit educatiu a Catalunya. Indicadors dels sistemes educatius: L'estat del'educació a Catalunya. Anuari 2011*. Barcelona: Fundació Jaume Bofill.

Alderson, C. (2005). *Diagnosing foreign language proficiency: the interface between learning and assessment*. London: Continuum.

Bachman, L. (1991). What does language testing have to offer?. *TESOL quarterly*, 25, 671-704.

Bachman, L. (2001). Designing and developing useful language tests. In Elder, C., Brown, A., Grove, F., Hill, K., Iwashita, N., Lumley, T., ... O'Loughlin, K. (ed.), *Studies in language testing (1st ed.)*, p. 109-123. Edinburgh: Cambridge University Press.

Bachman, L., and Palmer, A. (1996). *Language testing in practice*. London: Oxford University Press.

Consell Superior d'Avaluació del sistema educatiu. (2018). Avaluació de sisè d'educació primària. Recuperat de http://csda.gencat.cat/ca/arees_d_actuacio/avaluacions-consell/avaluacio-sise-primaria/prova_2018/que-es/.

Crystal, D. (2003). *English as a global language (2nd ed.)*. Cambridge: Cambridge University Press.

Edelenbos, P., and Kubanek-German, A. (2004). Teacher assessment: the concept of diagnostic competence. *Language testing*, 21, 259-283.

Flucher, G. and Davidson, F. (2007). *Language testing and assessment: an advanced resource book*. Abingdon: Routledge.

Gordon, J., Halasz, G., Krawczyk, M., Leney, T., Michel, A., Pepper, D., ... Wisniewski, J. (2009). *Key competences in Europe: opening doors for lifelong learners across the school curriculum and teacher education*. Warsaw: Case network reports.

- Gould, S. (1996). *The mismeasure of man. Revised and expanded*. New York: W.W. Norton.
- Haney, M. and Evans, J. (1999). National survey of school *psychologists* regarding use of dynamic assessment and other nontraditional assessment techniques. *Psychology in the Schools*, 36, 295-304.
- Harris, R., and McGhee-Nelson, E. (1992). *Bilingualism: Not the exception any more*. In R. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 3–14). Amsterdam: Elsevier.
- Nazari, B. (2012). Teach-to-test instruction of dynamic assessment: a critical overview. *Bellaterra Journal of Teaching & Learning Language & Literature*, 5, 56-68.
- Nola, N. (2002). Testing writing in the EFL classroom: student expectations. *English teaching forum*, 14-19.
- Shohamy, E., Or, I., and May, S. (2017). *Language testing and assessment (3rd ed)*. Cham: Springer.
- Starman, A. (2013). The case study as a type of qualitative research. *Journal of contemporary educational studies*, 1, 28-43.
- Seidlhofer, B. (2005). English as a lingua franca. *ELT Journal*, 59, 339-341.
- Tobón, S., Pimienta, J., and García, J. (2010). *Secuencias didácticas: aprendizaje y evaluación de competencias*. Mexico: Pearson education.
- Torrance, H., and Pryor, J. (1998). *Investigating formative assessment: Teaching, learning and assessment in the classroom*. Buckingham: Open University Press.
- Westera, W. (2001). Competences in education: a confusion of tongues. *Journal of curriculum studies*, 33, 75-88.

Vandergrift, L. (2007). Recent developments in second and foreign language listening comprehension research. *Language teaching*, 40, 191-210.

Yip, V., and Matthews, S. (2006). Assessing language dominance in bilingual acquisition: a case for mean length utterance differentials. *Language assessment quarterly*, 3, 97-116.

Zabala, A., and Arnau, L. (2007). La enseñanza de las competencias. *Aula de innovación educativa*, 161, 40-46.