

Bilingual and Non-Bilingual Classes at Primary School in Italy Results from a Standardised National Test

Francesca Costa

Università Cattolica del Sacro Cuore, Italia

Cristina Mariotti

Università degli Studi di Pavia, Italia

Abstract This study aims at determining the effectiveness of the Bilingual Education Italy (BEI) project as opposed to monolingual programmes in primary education in Italy, a country with low exposure to English as a foreign language. The BEI project originated from a memorandum of understanding between the Directorate-General for Educational Systems and School Autonomy of the Ministry of Education, the Ministry for Universities and Research, and the British Council. The research questions focus on the difference in the performance in English, oral mathematics and oral Italian between BEI and non-BEI students. To answer them, a statistical analysis of the English language competence state test in two primary schools involved in the BEI project was carried out. A descriptive analysis of the scores obtained by BEI and non-BEI students was performed, followed by the analysis of the correlation between the scores in English, Italian and mathematics, and a linear regression considering all the relevant variables. The results point to an advantage of BEI students over non-BEI ones.

Keywords BEI (Bilingual Education Italy). Primary School. Standardised National Test. CLIL (Content and Language Integrated Learning).

Summary 1 Introduction. – 2 Overview of Studies on the Effectiveness of BE Programmes in Primary Education. – 3 The Bilingual Education Italy (BEI) Project. – 4 Methodology. – 4.1 Research Questions. – 4.2 Sampling. – 5 Results. – 5.1 First Type of Analysis; – 5.2 Second Type of Analysis; – 5.3 Third Type of Analysis. – 6 Discussion. – 7 Pedagogical Implications and Conclusions.



Peer review

Submitted 2022-09-08
Accepted 2023-02-14
Published 2023-05-18

Open access

© 2023 Costa, Mariotti | © 4.0



Citation Costa, F.; Mariotti, C. (2023). "Bilingual and Non-Bilingual Classes at Primary School in Italy. Results from a Standardised National Test". *EL.LE*, 12(1), -166.

DOI 10.30687/ELLE/2280-6792/2023/01/006

1 Introduction

Research suggests that people who can think in more than one language have cognitive advantages over monolinguals which translate into valuable transversal competences such as problem-solving and the ability to communicate effectively (Glasbeek 2017). With respect to brain organisation and neurocognitive processing, an enriched language environment from an early school age seems to have a positive effect on learners' cognitive abilities that goes beyond learning the L2 (van de Craen, Ceuleers, Mondt 2007).

Being able to think in a language implies higher level cognitive thinking than is required by merely displaying knowledge; furthermore, educational practices that combine the formal learning of a language with learning a language through meaningful content can foster the development of the kind of higher order cognitive processing that leads to successful language learning and thus to thinking in that language (Pavón, Diaz 2020, 3). This is the kind of blend offered by bilingual programmes. These programmes, which vary as far as the amount of exposure to the language and starting age are concerned (Costa 2021), imply that two (or more) languages are used as a medium of instruction to teach non-language subject matter rather than just the languages themselves (Cummins 2013; Bialystok 2018). For the purposes of this study, bilingual education (henceforth, BE) is conceptualized as any school programme where non-linguistic subjects are taught in at least one additional language, in this case English, thus including programmes such as CLIL.

Despite the acknowledged benefits of BE, Bruton (2011) argued that these types of programmes may not be as beneficial as expected because on one hand they might lead to the oversimplification of content, and on the other hand the better results obtained by students enrolled in experimental classes may not be reliable due to a selection bias. As a matter of fact, it has been observed that high-performing students are more likely to be enrolled in these classes as opposed to classes where the foreign language is taught in a traditional way. Moreover, Faubert (2009) identified two difficulties in measuring the effect of bilingual education: the availability of reliable data and the difficulty of measuring the impact of the programmes that have been implemented without a control and treatment group. It is therefore important that research becomes increasingly oriented towards examining data obtained through official, standardised instruments in order to provide policymakers with meaningful data as suggested by Perez-Cañado (2011) and Dalton-Puffer and Smit (2013).

To address these concerns, the present study aims at determining the impact of the BEI (Bilingual Education Italy) project on the development of English language skills in a country with low exposure to English, by analysing performance measures of pupils' com-

petence. At the end of primary education in Italy (i.e. at age 10), students are assessed by means of a test designed and validated by the National Institute for the Evaluation of the Education and Training System (INVALSI),¹ a public research institution supervised by the Ministry of Education, Universities and Research. This test assesses the level of achievement of key competences commonly associated with mainstream school subjects.

The rationale for this article is to fill a gap in the literature as to date, in Italy, there is a paucity of studies that have conducted a systematic analysis of experimental and control groups (BE and non-BE) in primary education, and no studies have so far investigated official institutional data provided by the INVALSI.

2 Overview of Studies on the Effectiveness of BE Programmes in Primary Education

From the onset of BE programmes in Canada in the 1960s, research has investigated the cognitive advantage of this type of education for learners. Studies carried out in North America and in Europe claim that BE-type programmes lead to increased L2 proficiency (Swain, Lapkin 1982; Admiraal, Westhoff, Bot 2006; Dalton-Puffer 2007; Casal, Moore 2009), they are not detrimental to the learning of the subject matter taught in the additional language (Lambert, Tucker 1972; Swain, Lapkin 1982; Genesee 1987; Jäppinen 2005; Seikkula-Leino 2007), and can even be beneficial for the formal learning of the students' L1 (Nikolov, Mihaljević Djigunović 2006; van de Craen, Ceuleers, Mondt 2007). A meta-analysis of 17 investigations supports the view that teaching/learning a subject matter in an additional language can lead to even better results with respect to using just one language (Rolstad, Mahoney, Glas 2005). In this line of research, Costa (2021) carried out a narrative review of 25 studies investigating forms of BE across the world at all educational levels and showed that BE does not seem to be detrimental either to the acquisition of content (subject-matter) or to the development of the L1.

However, in her reflections on CLIL research Sylven (2013) pointed out that findings in this area should be interpreted bearing in mind that each country has its own profile defined by nation-specific contextual factors which include, but are not limited to, language policy framework, teacher training and education, the age when the programme is implemented, and extramural exposure to English. The wide variety of variables and the specific profile of each nation can

¹ National Institute for the Evaluation of the Education and Training System, cf. d.lgs. No. 286/2004.

account for the differences in the quality of experiences and results obtained in measuring the effects of BE programmes.

Against this backdrop, we will now focus on the literature analysing BE in the primary cycle of education which is based on results obtained by comparing the performances of experimental and control groups. In order to address the geographical bias pointed out by Sylven (2013), the scope of this review will be limited to southern Europe primarily because countries in this area typically share similarities in terms of low extramural exposure to English and lack of a governmental policy in the provision of teacher training programmes. The selection bias mentioned by Bruton (2011) will only be addressed where the authors make specific reference to the sorting criteria adopted by schools.

In Italy, Infante (2010) surveyed 298 primary school students for over two years in seven state schools where the subjects taught in English were art, science, history and technology. Students were pre- and post-tested on their English (on both receptive and productive skills), and they were post-tested on their Italian and the subject matters taught in English. Findings show no significant differences in the performance scores either for subject-matter content or language in the two cohorts.

In Spain, over the last two decades, BE programmes in the form of CLIL have proliferated especially in primary education due to governmental policies seeking to comply with the request of the European Union to achieve multilingualism in Europe (Casal, Moore 2009). In Andalusia, Lorenzo et al. (2010) carried out an extensive study relying on a sample of 403 primary and secondary schools where the additional language was either English, French, or German. Findings show that pupils in all BE classes significantly outperformed non-BE pupils as regards L2 language learning in the four skills. It is important to note that the better performance of pupils in BE classes in this study should not be ascribed to pre-selection, as the authors explicitly address the topic declaring that students had been randomly assigned to the programme.

Jiménez Catalán and Ruiz de Zarobe (2009) investigated L2 competence by comparing the receptive vocabulary of 65 non-CLIL and 65 female CLIL students. The results point to a significantly better performance by CLIL students.

Fernández-Sanjurjo et al. (2017) investigated the repercussions of English-medium instruction on content acquisition by analysing the competence in Science of 709 6th year primary students in CLIL and non-CLIL streams. The authors found that the cohort studying in the L1 performed slightly better than the bilingual one, thus contrasting results from the literature (see for instance Serra 2007 for primary schools in Switzerland). However, it should be pointed out that all standards were reached by both groups and all students were test-

ed on content in the L1 even if they had studied in English, since this could have affected their performance.

Finally, Coral et al. (2018) analysed the results achieved by sixth grade Catalan students in the English language competence state test during the period 2009-12 using an extensive dataset comprising 85 primary schools. The authors chose a reference school where a type of BE was successfully implemented with the teaching of Physical Education in English and where specific strategies for teaching subject matter content in English were applied. They compared the results obtained in the foreign language test assessing reading and listening comprehension 1) in schools where English was taught in a traditional way, 2) in schools where some form of non-CLIL experimental English language teaching was carried out, 3) in schools applying the CLIL approach, and 4) in the reference school. The results show that English competence improved more in schools implementing CLIL and even more so in the reference school, where pupils were taught PE in English for 2.3 hours per week as opposed to 1 or 1.3 hours per week in the other schools offering CLIL programmes. The authors point out that since most of the literature has argued that increasing time exposure is not enough to improve L2 competence (De Graaff et al. 2007), the results obtained in this school were probably due to the fact that substantial methodological changes were made in the teaching of PE in English such as combining language teaching techniques and specific PE teaching styles stimulating high-order thinking skills, cooperative learning, and language scaffolding.

Based on the reviewed studies, it is possible to affirm that BE programmes carried out in primary schools in Southern Europe point towards an improvement for L2 acquisition especially in receptive skills and reveal partially negative results as far as subject matter knowledge is concerned.

3 The Bilingual Education Italy (BEI) Project

This article looks at two Italian schools in Northern Italy taking part in the BEI project, which was developed as a pilot project in six schools in 2009. BEI stems from a memorandum of understanding between the Directorate-General for Educational Systems and School Autonomy of the Ministry of Education, the Ministry for Universities and Research, and the British Council (Costa 2023). The project was set up in February 2010 and is very similar to one previously carried out in Spain (Dobson, Smit 2010). Bilingual Education is carried out through literacy activities from the start of primary school, with some subjects being taught in English. The main subjects chosen by the schools are science, geography, art and literacy. Right from the beginning students learn to read and write in English through sys-

tematic synthetic phonics. Cavalieri and Stermieri (2016) indicate role-playing games, storytelling, TPR, and Synthetics Phonics (Costa 2019; Costa, Pladevall-Ballester 2020; Costa, Mair 2022) as typical educational strategies used in this project, which is based on an inclusive vision and aims at being a resource for all schools. The following requirements are necessary to participate in the project: teachers should have a B2 level of English and a willingness to participate in project design meetings, training seminars, and conferences; at least 50% of the classes in each school must be involved; the project must run continuously for all five years of primary school; and 25% of the weekly teaching hours must be dedicated to English.

After the first four years of experimentation, a monitoring of the project was carried out by the University of Modena and Reggio Emilia (Bondi et al. 2014) using questionnaires given to the teachers, semi-structured interviews of the six headmasters and teachers at the schools involved, a focus group made up of children and parents, and a common task involving 10 recordings (per school) of a picture description and 20 written assignments (per school) to evaluate the results. The evaluation found that the headmasters, teachers and students were satisfied while the parents, though appreciating the project, had some doubts about the study of grammar and the way they were tested in the subject matters studied. The parents also expressed a desire for an extension of the BEI project to the lower secondary school level. Regarding the language assessment, students were found to have an A2+ level along with other positive aspects such as language creativity and pronunciation accuracy. After the first five years of experimentation in primary school, the project was extended to lower secondary school and called the CLIL Excellence programme.

4 Methodology

This study is based on a quantitative research paradigm as its aims are to analyse the results of the standardised national test of English language competence (reading and writing) in both BEI (experimental group) and non-BEI classes (control group) in two schools located in Northern Italy longitudinally during the period 2017-19. The period 2017-19 was chosen because it was the first two years in which the INVALSI test for English was administered to schools and because by taking into consideration datasets gathered before the COVID-19 pandemic, we can exclude that the results are affected by variations in the teaching mode. To this end, the performance in a national standardised English test (INVALSI) for the BEI and non-BEI classes was statistically compared. Given the wealth of data provided by the INVALSI together with the English standardised test performance, the correlation between performance in English and per-

formance in Italian and mathematics was also compared in the two groups. Finally, a linear regression analysis was carried out to see whether, taking into account all other variables provided by the INVALSI, it could be concluded that there were significant performance differences in English in one of the two groups.

4.1 Research Questions

The following research questions were designed:

RQ1. Is there any difference in the performance in English (in the INVALSI test), in oral mathematics (with marks obtained at the end of primary school) and in oral Italian (with marks obtained at the end of primary school) between BEI and non-BEI students? First type of analysis.

RQ2. Is there a correlation between the INVALSI score in English and the mark in oral Italian on the one hand and the INVALSI score in English and the mark in oral mathematics on the other? Second type of analysis.

RQ3. If, based on the results of the analysis carried out to answer RQ1 and RQ2, BEI students perform significantly better than non-BEI ones, is this result only ascribable to the fact that they were in a BEI class or should differences in student performance in the English INVALSI be ascribed to other variables according to a linear regression model? Third type of analysis.

4.2 Sampling

As described above, the two schools participating in this study both had BEI classes and non-BEI classes as part of the initial project. Data relating to the schools and the test results were requested directly from the INVALSI using a procedure which safeguards data and guarantees anonymity of the candidates. All school heads in the BEI schools (6) were asked for their consent in order to request data from the INVALSI. Two of them complied and so the request was sent to the INVALSI. The sample was composed of 898 students from two schools where the BEI project is carried out in some of the classes.

The datasets received by the INVALSI concerned the performance in the INVALSI tests for English carried out in the fifth grade for the following school years: 2017-18 and 2018-19. The original 4 data sets (school 1: 2017-18 and 2018-19; school 2: 2017-18 and 2018-19) were combined into a single data set, the two schools and the two school

years were considered homogeneous. 898 students were involved with 21 variables being part of the dataset, of which only those relevant for this study have been taken into account, i.e. type of BEI/non-BEI class; gender; country of origin; mother's educational level; oral grade in Italian, oral grade in mathematics; raw score in English; adjusted (after correction factor) score in English;² score in English as a percentage; adjusted (after correction factor) percentage score in English. There were 547 pupils in BEI classes (60.9%) and 351 students (39.1%) in non-BEI classes. 415 students were female (46.2%) and 483 male (53.8%).

5 Results

The following section will provide a first type of analysis of the performance in English, Italian, and mathematics of both BEI and non-BEI students. A second type of analysis is then undertaken to determine whether there is a correlation between performance in English and performance in Italian and mathematics in both groups. Finally, a third type of analysis in the form of a linear regression model will be proposed taking into account all the variables.

5.1 First Type of Analysis

This section describes the performance in English, Italian, and mathematics of both BEI and non-BEI students.

5.1.1 Performance in English for Students in the BEI and non-BEI Classes

In order to fully evaluate the English performance of the two groups, the whole range of scores available from INVALSI were initially taken into account: 'raw scores', 'scores adjusted for the correction factor', 'raw percentages (of correct answers)', and 'percentages (of correct answers) adjusted for the correction factor'. The distribution of 'raw scores' and 'raw percentages' from the English test are graphically assessed and compared for BEI and non-BEI students. It is clear from the graphs that the results of the non-BEI classes are very heterogeneous: the distribution is very spread both in terms of 'raw scores' and 'raw percentages' [fig. 1]. The median performance in English is

² The methodology allows us to estimate the probability of cheating for each class of students. On the basis of this probability, a correction factor is constructed that 'penalises' the average class scores as the degree of suspicion of cheating increases.

much lower for the non-BEI classes than for the BEI classes; this attitude is confirmed comparing the mean values in the two groups using the Welch Two Sample t-test.

In particular, the mean 'raw score' for BEI students (22.56) is significantly higher ($p\text{-value} < 2.2\text{e-}16$) than the mean 'raw score' for non-BEI students (19.46). In addition, the mean 'raw percentage' achieved by BEI students (89.94) is significantly higher ($p\text{-value} < 2.2\text{e-}16$) than the mean 'raw percentage' achieved by non-BEI students (78.26).

Comparison of Performance in English for BEI and Non-BEI Classes

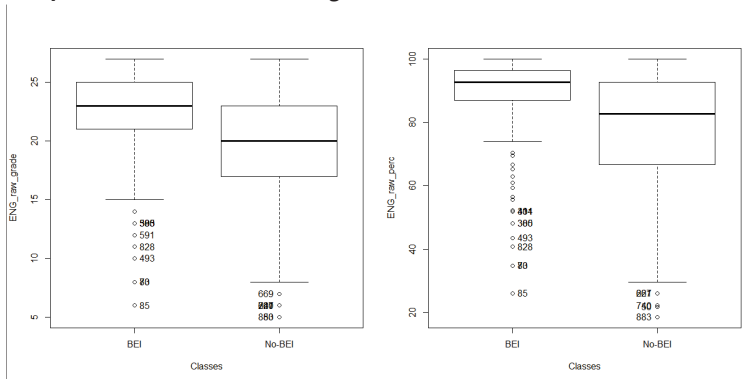


Figure 1 Comparison of the distribution of the raw scores and the raw percentages on the INVALSI English test for BEI and non-BEI students

After applying the correction factor, the same conclusions are obtained, as shown in figure 2. The performances achieved in English for BEI and non-BEI classes continue to differ even when we compare the 'scores adjusted for the correction factor' and the 'percentages adjusted for the correction factor' instead of the raw measures. By performing the Welch Two Sample t-test we can see that the mean 'score adjusted for the correction factor' for BEI students (21.29) is significantly higher ($p\text{-value} = 6.7\text{e-}10$) than the mean 'score adjusted for the correction factor' obtained by non-BEI students (19.43). In addition, the mean 'percentage for BEI students adjusted for the correction factor' (84.91) is significantly higher ($p\text{-value} = 3.9\text{e-}9$) than the 'mean percentage adjusted for the correction factor' obtained by non-BEI students (78.15).

Comparison of Performance in English for BEI and Non-BEI Classes After the Correction Factor

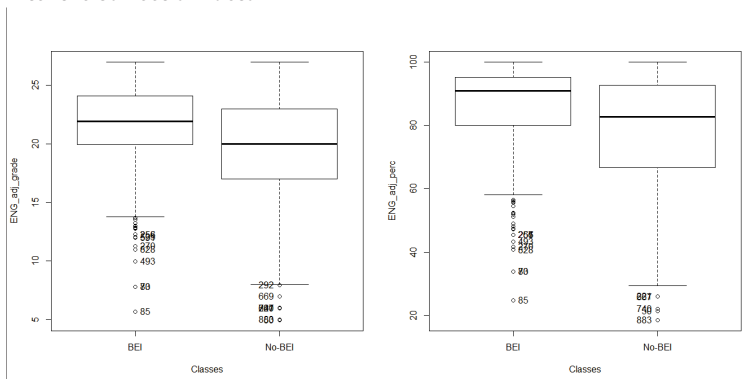


Figure 2 Comparison of the distribution of the scores and the percentages for BEI and non-BEI students, adjusted for the correction factor

5.1.2 Performance in Italian for BEI and non-BEI Classes

The means of the 'grades in oral Italian' are equal to 7.97 and 7.83 respectively for BEI and non-BEI students, performing the Welch Two Sample t-test. Thus, we can conclude that BEI students seem to perform significantly better than non-BEI students ($p\text{-value} = 0.017$) but only at a level of 0.05. This conclusion can be interpreted in two ways: we could conclude that those who are more predisposed to learning Italian are likely to strengthen their knowledge of English as well, but it is also possible that the improvement in English due to the enrolment in BEI classes will lead to greater competence in Italian too.

5.1.3 Performance in Mathematics for BEI and non-BEI Classes

With reference to the 'grade in oral mathematics', the same analysis used for grades in Italian was carried out. It revealed that statistically BEI students do not, on average, perform significantly better in mathematics than non-BEI students at any commonly used level of significance ($p\text{-value} = 0.1012$).

5.2 Second Type of Analysis

This section describes some of the correlations between performance in English Italian and mathematics for both the BEI and non-BEI groups.

5.2.1 Is There a Correlation Between Performance in English and Performance in Italian? Does This Apply to Both Groups?

There is empirical evidence of a linear correlation between the variables 'raw score in English' and 'oral grade in Italian'. In fact, the Pearson linear correlation coefficient is 0.346. Therefore, a high score in English in the INVALSI test tends to correspond to a high grade in Italian. We then assess whether the same type of relationship is observed in BEI and non-BEI classes. In particular, in BEI classes the Pearson coefficient is 0.391, while for non-BEI classes, the index is 0.3, so there seems to be a greater correlation in BEI classes. However, in both of classes there is a significant linear relationship, since the p-values of the suitable tests are equal to $< 2.2e-16$ and $1.107e-8$, respectively, for BEI and non-BEI classes.

5.2.2 Is there a Correlation between Performance in English and Performance in Mathematics? Does this Apply to both Groups?

The same analysis was undertaken regarding performance in mathematics. The results below indicate a linear correlation between 'raw score in English' and 'grades in oral mathematics' for all students, i.e. both those in BEI classes and those in non-BEI classes. The Pearson linear correlation coefficient is 0.366 and it is significantly different from 0 ($p\text{-value} < 2.2e-16$), thus demonstrating that the good students perform well in all areas. As in the previous case concerning the relation between the performance in English and in Italian, we observe a lower value of the Pearson index in non-BEI classes than in the others. In BEI classes, the Pearson linear correlation coefficient is 0.488 while in non-BEI classes it is 0.265; the two values provide evidence in favor of the presence of a significant positive relation (the p-values are $< 2.2e-16$ and $5.3e-7$ respectively, in BEI and non-BEI classes).

The analysis confirms that students tend to perform in a similar way in English, Italian and mathematics, but in particular, in the BEI classes, English and mathematics seem to go more 'hand in hand' than do English and Italian. All the relationships are less marked in non-BEI classes. Similarly, the results highlight that the students in BEI classes seem to perform better on average in Italian and mathematics compared to students from other classes.

5.3 Third Type of Analysis

Having performed an exploratory analysis of the different composition between BEI and non-BEI classes, our data show that in the BEI classes there are more children with Italian parents and a mother with a college degree than in the non-BEI classes. However, what also emerges from these data is that while the mother's educational level seems to affect performance in English, the parents' country of origin is less significant. It has also been shown that BEI classes have a better performance in Italian and this variable strongly correlates with the performance in English; it is worth checking these variables to assess the actual surplus attributable to the BEI class. This analysis is presented below.

To assess the effect on the English performance for students in BEI and non-BEI classes (after controlling for the effects of the other relevant variables mentioned above), a linear regression model was carried out using the 'scores adjusted for the correction factor' such as the response variable, since this seems to be the variable that best satisfies the hypothesis of normal distribution.

The multiple linear regression model was constructed taking into account the previously examined explanatory variables: BEI, country of origin, mother's educational level and gender in addition to quantitative variables: oral grade in Italian and mathematics.

The results are summarised in Table 1.

Table 1 The table shows the estimated coefficients and their significance: **** 0.001 *** 0.01 ** 0.05 * 0.1 ' ' 1

Linear Regression Model	
Variable	Estimated coefficient
BEI	-1.79***
Country of origin (baseline: foreign)	1.05**
Mother's educational level (baseline: college degree or post-high school diploma)	-0.63*
Mother's educational level (baseline: Professional qualification or lower)	-0.87*
Gender (baseline: male)	-0.26
Oral grade in Italian	0.24
Oral grade in mathematics	1.21***

The model is globally significant (as highlighted by the p-value of the F test: p-value < 2.2e-16), although it has a limited goodness of fit to the data as shown by the R-squared index: 0.1663, which indicates that performance in English can be further modelled using other factors (which, however, cannot be quantified as they are largely latent). The results undoubtedly confirm the effectiveness of BEI

classes. In particular, in the model which includes the explanatory variables listed in Table 1, we have:

- The 'oral Italian grade' and 'gender' variables are not significant, which means they do not constitute relevant reasons to properly evaluate the performance in English in this model, which includes the other explanatory variables, all of which are highly significant. In fact, by including other variables such as non-BEI and BEI classes, they become so effective at explaining the performance in English that they make the relationship we had observed between performance in Italian and performance in English weaker. This might be a consequence of the particular composition of the BEI classes where we tend to find the best performers.
- Holding constant gender, mother's educational level, country of origin, and grades in oral mathematics and oral Italian, being in a BEI class significantly improves students' grade in English. This variation is higher than that associated with the other variables, and therefore it appears that type of class is a more relevant explanation of English grades than the other variables.

The results of the analysis are summarised in Table 2.

Table 2 Summary of results

	Element of analysis	Results
First type of analysis (RQ1)	Explorative performance in English (INVALSI test)	BEI significantly higher on average
	Explorative performance in Italian (oral grade)	BEI significantly higher on average
	Explorative performance in mathematics (oral grade)	No differences on average detected between BEI and non-BEI classes
Second type of analysis (RQ2)	Relation between performance in English (INVALSI test) and oral Italian grade	Both the BEI and non-BEI classes are correlated, but for the BEI class we observe a stronger correlation
	Relation between performance in English (INVALSI test) and oral mathematics grade	Both the BEI and non-BEI classes are correlated, but for the BEI class we observe a stronger correlation
Third type of analysis – linear regression model (RQ3)	Explanation of the English performance by using the type of class, gender, mother's educational level, country of origin, oral Italian grade and oral mathematics grade as variables	BEI leads to a higher performance in English, having held fixed the levels of gender, mother's educational level, country of origin, oral Italian grade and oral mathematics grade

6 Discussion

This study compared the BEI (experimental group) and the non-BEI (control group) classes using the INVALSI tests for English as the focus of the analysis. The sample was made up of 898 primary school students in Italy in the 2017-18 and 2018-19 school years. The results show that BEI classes have a significant advantage in many areas, not only in English, as was to be expected given the qualitatively and quantitatively different input.

Having said this, we should bear in mind that 74% of students whose parents have college degrees are in the BEI classes, so there could be an underlying element of preselection (Bruton 2011), but the linear regression analysis demonstrates that being in a BEI class as opposed to a non-BEI class is a stronger predictor of the performance in English than the mother's educational level. It is true, however, that at the time of enrolment in Italian public schools, parents may opt for one section as opposed to another, although it is the school personnel that form the classes and do not have to abide by this choice.

As mentioned above, BEI provides students with more hours in English, but what is important to note is that these additional hours involve not only quantitatively different but also qualitatively different input. In addition, it should be kept in mind that any differences in performance between BEI and non-BEI classes may be attributed not only to the different inputs but also to a hypothetical greater acculturation and predisposition to learning of students in BEI classes.

The analysis (linear regression model) reveals some aspects that are not strictly related to the research questions proposed in this article, but which nonetheless deserve mentioning. The linear regression model also shows that when the type of class (BEI or non-BEI), gender, country of origin, and skills in mathematics and Italian are held constant, having a mother with a professional qualification rather than a high school diploma or a college degree significantly lowers the grade in English. With class membership (BEI or non-BEI), gender, mother's educational level and skills in mathematics and Italian held constant, being a heritage pupil significantly increases the performance in English. An attempt was made to provide a more in-depth conclusion through the interactive effect of type of class and country of origin, but the interaction was not a significant indication that, with all other variables held constant, a heritage pupil in a BEI class would not achieve a higher grade in English compared to a foreigner in a non-BEI class. This point marginally touches on the topic of the present article because it confirms that being a heritage pupil does not seem to negatively affect educational performance (in this case in English). Moreover, it is assumed that pupils of non-Italian origin are already bilingual within their family, so this result may also confirm that being bilingual does not affect the educational perfor-

mance in English (which in this case would be the third language) rather based on the above, it affects it in a positive way. In addition, with the type of class (BEI or non-BEI), gender, mother's educational level, country of origin, and competence in Italian held constant, each additional point in the oral mathematics score is associated with a significant increase on average of the score in English, showing that BEI classes have an advantage in English as well as in other subjects.

7 Pedagogical Implications and Conclusions

As already pointed out, Italy is a country with low exposure to English as an L2. Moreover, it is among the European countries where people are less likely to speak a foreign language (Eurobarometer 2012). It is therefore critical to provide policymakers with information that might help shape policies in public education more effectively, starting from primary education where the positive effect of learning subject matter using an L2 on learners' cognitive abilities can have a particularly far-reaching effect (van de Craen, Ceuleers, Mondt 2007).

The data analysed in the present study show that the best performances in English tend to be located in BEI classes, i.e. the BEI/non-BEI variable is stronger than the other ones. Thus, holding all other variables constant, being in a BEI class significantly improves students' grades in English. At the same time, though, results show that the mother's educational background seems to be playing a role, thus confirming the relevance of this parameter for academic success (Queirolo Palmas 2002; Ballarino, Checchi 2006; Ciccotti, Sabbadini 2007; Besozzi, Colombo, Santagati 2009). The results of the present study do not confirm Infante's conclusions (2010), who did not find any significant differences either in the English performance scores nor in the experimental or the control classes. Instead, they are in keeping with what was found by Jiménez Catalán and Ruiz de Zarobe (2009), Lorenzo et al. (2010), and Coral et al. (2018), who showed that pupils in the experimental classes significantly outperformed pupils in control classes as far as their receptive skills of English proficiency are concerned.

It should be noted that since the INVALSI test is not based on the syllabus for bilingual programmes, which are geared towards using the L2 to learn content, this makes the results obtained in English by BEI pupils in this test all the more relevant.

A possible limitation in the interpretation of these results might derive from the fact that our data relies on one macro-variable, namely the students' score in a written test, to assess the English competence of students, which is a complex and multi-faceted phenomenon. Moreover, the comparison between performance in English and performance in Italian and mathematics draws on two different evaluation modalities, namely the INVALSI test scores in the case of English

and primary school grades in the case of Italian and mathematics. However, it should also be pointed out that our dataset satisfies the conditions posited in the literature for effective research in BE: our measures are reliable as they are obtained through a standardised national test, and our analysis was based on a control and treatment group (Faubert 2009; Perez-Cañado 2011; Dalton-Puffer, Smit 2013).

Moreover, it might be argued that care should be taken in the interpretation of results showing a better L2 performance of students enrolled in BEI programmes since these programmes typically have more hours of instruction and more exposure to the target language, which is, in turn, expected to lead to greater language acquisition (Krashen 1982). This is of course true, but it has been observed that greater exposure to the language is not enough to improve L2 competence since the quality of the input provided also plays a role in students' achievements (De Graaf et al. 2007). We can therefore reasonably assume that in our data both factors might have concurred and that it might be useful to carry out further research into the characteristics of the input delivered in BEI programmes.

Going back to the main objective of our analysis, it seems that BEI classes have an advantage in English as well as in other subjects compared with non-BEI classes. Considering the inclusive policy of the initial project it would be logical, on the basis of the present statistical analysis, to propose BEI in all the classes at the schools in question and not just in some of them, in order to create a project that would apply to everyone without discrimination of any kind.

Bibliography

- Admiraal, W.; Westhoff, G.; de Bot, K. (2006). "Evaluation of Bilingual Secondary Education in the Netherlands: Students' Language Proficiency in English". *Educational Research and Evaluation*, 12(1), 75-93. <https://doi.org/10.1080/13803610500392160>.
- Ballarino, G.; Checchi, D. (a cura di) (2006). *Sistema scolastico e disuguaglianza sociale. Scelte individuali e vincoli strutturali*. Bologna: il Mulino.
- Besozzi, E.; Colombo, M.; Santagati, M. (2009). *Giovani stranieri, nuovi cittadini. Le strategie di una generazione 'ponte'*. Milano: Franco Angeli.
- Bialystok, E. (2018). "Bilingual Education for Young Children: Review of the Effects and Consequences". *International Journal of Bilingual Education and Bilingualism*, 21(6), 666-79. <https://doi.org/10.1080/13670050.2016.1203859>.
- Bondi, M.; Cavalieri, S.; Poppi, F.; Stermieri, A. (2014). *Il progetto IBI/BEI nella scuola primaria: sintesi del rapporto di monitoraggio*. https://www.britishcouncil.it/sites/default/files/final_sintesi.pdf.
- Bruton, A. (2011). "Is CLIL so Beneficial, or just Selective? Re-Evaluating some of the Research". *System*, 39(4), 523-32. <https://doi.org/10.1016/j.system.2011.08.002>.
- Casal, S.; Moore, P. (2009). "The Andalusian Bilingual Sections Scheme: Evaluation and Consultancy". *International CLIL Research Journal*, 1(2), 36-46.
- Cavalieri, S.; Stermieri, A. (2016). "The BEI/IBI Project: A Study on the Best Practices in Intergrating Language and Content Learning in Primary Schools". Garzone, G.; Heaney, D.; Riboni, G. (eds), *Focus on LSP Teaching: Developments and Issues*. Milan: LED, 211-44.
- Ciccotti, E.; Sabbadini, L.L. (2007). "Come cambia la vita dei bambini. Indagine statistica multiscopo sulle famiglie". *Questioni e Documenti. Quaderni del Centro Nazionale di documentazione e analisi per l'infanzia e l'adolescenza*, 42. Istituto degli Innocenti Firenze.
- Coral, J.; Lleixà, T.; Ventura, C. (2018). "Foreign Language Competence and Content and Language Integrated Learning in Multilingual Schools in Catalonia: An ex Post Facto Study Analysing the Results of State Key Competences Testing". *International Journal of Bilingual Education and Bilingualism*, 21(2), 139-50. <http://doi.org/10.1080/13670050.2016.1143445>.
- Costa, F. (2019). *Enjoy Teaching English. Insegnare inglese nella scuola primaria*. Firenze: Giunti Scuola.
- Costa, F. (2021). "What does Research Tell us about Experiences and Forms of Bilingual Education?" Maggioni, M.L.; Murphy, A. (eds), *Back to the Future*. Bern: Peter Lang, 91-112.
- Costa, F. (2023). "Mediation and Mode Continuum in Primary Bilingual (English and Italian) Education". Lopriore, L. (ed.), *Engaging Research: Transforming Practices for the English as a Foreign Language Classroom*. Alexandria: TESOL Press, 9-21.
- Costa, F.; Mair, O. (2022). "Synthetic Phonics in Italian Bilingual Schools: Teachers' Views and Pronunciation". Cortina-Pérez, B.; Andúgar, A.; Álvarez-Cofiño, A.; Corral, S.; Martínez León, N.; Otto, A. (eds), *Addressing Future Challenges in Early Language Learning and Multilingual Education*. Madrid: Dykinson, 196-204.

- Costa, F.; Pladevall-Ballester, E. (2020). "Learners' Outcomes and Effective Strategies in Early Second Language Learning". *EuroAmerican Journal of Applied Linguistics and Languages E-JournALL*, 6(2), 1-10.
- Cummins, J. (2013). "Bilingual education and Content and Language Integrated Learning (CLIL): Research and its Classroom Implications". *Revista Pa-dres y Maestros*, 349, 6-10.
- De Graaff, R.; Koopman, G.J.; Anikina, Y.; Westhoff, G. (2007). "An Observation Tool for Effective L2 Pedagogy in Content and Language Integrated Learning (CLIL)". *International Journal of Bilingual Education and Bilingualism*, 10(5), 603-24. <https://doi.org/10.2167/beb462.0>.
- Dalton-Puffer, C. (2007). *Discourse in Content and Language Integrated (CLIL) Classrooms*. Amsterdam: Benjamins.
- Dalton-Puffer, C.; Smit, U. (2013). "Content and Language Integrated Learning: A Research Agenda". *Language Teaching*, 46(4), 545-59. <https://doi.org/10.1017/S0261444813000256>.
- Dobson, A.; Pérez Murillo, M.D.; Richard, J. (2010). *Bilingual Education Project Spain Evaluation Report*. <https://www.britishcouncil.es/sites/default/files/bilingual-education-project-spain-evaluation-report-en.pdf>.
- European Commission (2012). *Special Eurobarometer 386 – Europeans and Their Languages*. <https://europa.eu/eurobarometer/surveys/detail/1049>.
- Faubert, V. (2009). "School Evaluation: Current Practices in OECD Countries and a Literature Review". *OECD Education*. Working Paper No. 42. <https://er-ic.ed.gov/?id=ED529650>.
- Fernández-Sanjurjo, J.; Fernández-Costales, A.; Arias Blanco, J. (2017). "Analysing Students' Content-Learning in Science in CLIL vs. non-CLIL Programmes: Empirical Evidence from Spain". *International Journal of Bilingual Education and Bilingualism*, 22(6), 661-74. <https://doi.org/10.1080/13670050.2017.1294142>.
- Genesee, F. (1987). *Learning through Two Languages: Studies of Immersion and Bilingual Education*. Cambridge: Newbury House Publishers.
- Glasbeek, G. (2017). *Policy Paper – The Importance of Transversal Skills and Competences for Young People in a Modern Europe*, Association des États Généraux des Étudiants de l'Europe (AGEE). <https://www.aegee.org/policy-paper-the-importance-of-transversal-skills-and-competences-for-young-people-in-a-modern-europe>.
- Infante, D. (2010). *Il Content and Language Integrated Learning (CLIL) in Italia, Modelli didattici e sperimentazioni nella scuola primaria*. Roma: Editrice Nuova Cultura.
- Jäppinen, A.K. (2005). "Thinking and Content Learning of Mathematics and Science as Cognitive Development in Content and Language Integrated Learning (CLIL): Teaching through a Foreign Language in Finland". *Language and Education*, 19(2), 148-69. <https://doi.org/10.1080/09500780508668671>.
- Jiménez Catalán, R.; Ruiz de Zarobe, Y. (2009). "The Receptive Vocabulary of EFL Learners in two Instructional Contexts: CLIL versus non-CLIL Instruction". Ruiz de Zarobe, Y.; Jiménez Catalán, R. (eds), *Content and Language Integrated Learning: Evidence from Research in Europe*. Clevedon: Multilingual Matters, 81-92.

- Krashen, S. (1982). *Principles and Practice in Second Language Acquisition*. New York: Pergamon.
- Lambert, W.; Tucker, R. (1972). *Bilingual Education of Children: The St. Lambert Experiment*. New York: Newbury House Publishers.
- Lorenzo, F.; Casal, S.; Moore, P. (2010). "The Effects of Content and Language Integrated Learning in European Education: Key Findings from the Andalusian Bilingual Settings Evaluation Project". *Applied Linguistics*, 20, 1-25.
- Nikolov, M.; Mihaljević Djigunović, J. (2006). "Recent Research on Age, Second Language Acquisition and Early Foreign Language Learning". *Annual Review of Applied Linguistics*, 26, 234-60. <https://doi.org/10.1017/S0267190506000122>.
- Pérez Cañado, M.L. (2011). "CLIL Research in Europe: Past, Present, and Future". *International Journal of Bilingual Education and Bilingualism*, 15(3), 1-27. <http://doi.org/10.1080/13670050.2011.630064>.
- Pavón, V.; Díaz, W. (2020). "The Bilingual Advantage: The Impact of Language Learning on Mind & Brain. Position Paper". *EduCluster Finland-University of Jyväskylä Group*. <https://educclusterfinland.fi/bilingual-advantage>.
- Queirolo Palmas, L. (2002). "Nuove e vecchie disuguaglianze nella scuola di massa". Ribolzi, L. (a cura di), *Formare gli insegnanti. Lineamenti di sociologia dell'educazione*. Carocci: Roma, 207-20.
- Rolstad, K.; Mahoney, K.; Glas, G. (2005). "The Big Picture: A Meta-Analysis of Program Effectiveness Research on English Language Learners". *Educational Policy*, 19(4), 572-94. <https://doi.org/10.1177/0895904805278067>.
- Seikkula-Leino, J. (2007). "CLIL Learning: Achievement Levels and Affective Factors". *Language and Education*, 21(4), 328-41. <https://doi.org/10.2167/le635.0>.
- Serra, C. (2007). "Assessing CLIL at Primary School: A Longitudinal Study". *International Journal of Bilingual Education and Bilingualism*, 10(5), 582-602. <http://doi.org/10.2167/beb461.0>.
- Swain, M.; Lapkin, S. (1982). *Evaluating Bilingual Education: A Canadian Case Study*. Clevedon: Multilingual Matters.
- Sylvén, L.K. (2013). "CLIL in Sweden – Why does It not Work? A Metaperspective on CLIL across Contexts in Europe". *International Journal of Bilingual Education and Bilingualism*, 16(3), 301-20. <https://doi.org/10.1080/13670050.2013.777387>.
- van de Craen, P.; Ceuleers, E.; Mondt, K. (2007). "Cognitive Development and Bilingualism in Primary Schools: Teaching Maths in a CLIL Environment". Marsh, D; Wolff, D. (eds), *Diverse Contexts – Converging Goals. CLIL in Europe*. Bern: Peter Lang, 185-200.

