

The Effects of CLIL on Subject Matter Learning: The Case of Science in Primary and Secondary Education



Juan de Dios Martínez Agudo

Abstract This chapter focuses on an under-researched topic in relation to Content and Language Integrated Learning (CLIL): the effects of bilingual programmes on content learning. It was carried out with a total of ten schools (in Primary and Secondary Education) and 318 students in public bilingual and charter non-bilingual schools in the autonomous community of Extremadura. Students' performance in the subjects of Science in Primary Education and Natural Science in Compulsory Secondary Education is compared across schools in order to provide data on the important issue of whether CLIL programmes are watering down subject matter learning or promoting it as successfully as in monolingual streams. Factor and discriminant analyses are performed with the type of school and educational level variables in order to determine whether CLIL is truly responsible for the differences ascertained or whether other variables account for a greater proportion of the variance. The statistical analysis confirms the positive effects of CLIL on content subject learning, showing that bilingually educated students outperform monolingually educated ones at both educational stages, although this difference is clearly significant when finishing their Secondary Education studies. The results also provide some clear-cut differences in terms of type of school and educational level, as public bilingual schools only outperform charter non-bilingual schools only at the end of Compulsory Secondary Education, which suggests the long-term effects of CLIL on subject matter learning.

1 Introduction

Based on the pioneering immersion programmes in Canada and bilingual programmes in the USA which have led to the progressive introduction of bilingual education worldwide, there has been an explosion of interest and debate in the

J. d. D. Martínez Agudo (✉)
Faculty of Education, University of Extremadura, Badajoz, Spain
e-mail: jdmtinez@unex.es

© Springer Nature Switzerland AG 2021
M. L. Pérez Cañado (eds.), *Content and Language Integrated Learning in Monolingual Settings*, Multilingual Education 38,
https://doi.org/10.1007/978-3-030-68329-0_9

last two decades about the potential of the European approach to bilingual education—CLIL (Content and Language Integrated Learning)—, not only in Europe but also around the world. CLIL has been conceived as an alternative to the Communicative Language Teaching (CLT) approach (Coyle et al., 2010), an extension of it (Dalton-Puffer, 2007; Lasagabaster & Sierra, 2010) or just a new paradigm of education (Ouazizi, 2016). While Coyle et al. (2010) do not view CLIL as ‘simply another step in language teaching or a new development in content-subject methodology’ (p. ix) but rather as ‘a major rethink of how we teach what we teach’ (pp. ix–x), Dalton-Puffer (2011, p. 195) is cautious and warns that, contrary to most expectations, CLIL is not a panacea, as evidenced by the fact that some downsides have been recently reported by CLIL research as well as the fact that there is a wide gap between what is provided in CLIL teaching and what comes out in terms of CLIL learning. Due to the idiosyncrasies of the European context and the unceasing search for improved language teaching methods aimed at increasing L2 competence at a time when European integration is idealised, what becomes clear is that CLIL has emerged as a ‘timely solution to European plurilingual education’ (Pérez Cañado, 2012, p. 315) in an increasingly globalised world where ‘(bi)multilingualism is the norm whereas monolingualism is the exception’ (Ouazizi, 2016, p. 113).

Regarding the different conceptualisations of CLIL viewed as an umbrella construct which lacks conceptual precision as well as the fact that its scope is still not clear-cut, Cenoz et al. (2014) recognise that the varied interpretations of the CLIL approach suggest that CLIL is understood in different ways, which makes it difficult to pin down its uniqueness. For example, the language and content balance as well as the intensity of the exposure to the foreign language, among other CLIL core characteristics, are understood in different ways, as Cenoz et al. (2014) point out. Based on the succinct definition by Coyle et al. (2010, p. 1), ‘Content and Language Integrated Learning (CLIL) is a dual-focused educational approach in which an additional language is used for the learning and teaching of both content and language’, what becomes clear is that the core feature of CLIL is, no doubt, its dual-focused nature, as this ‘two for one’ approach strives to promote the integration of both L2 learning and content learning. Such a duality is certainly seen as the main strength of this educational approach, as both foreign language and subject matter content should be learnt and taught in an integrated way. Needless to say, such integration, in turn, involves the major challenge facing CLIL teachers and learners; as Ruiz de Zarobe (2013, p. 235) claims, ‘the challenge remains of how to enable learners to make best use of both areas in the classroom’. In the same vein, Cenoz et al. (2014, p. 244) also make it clear that

The dual role of language and content has been understood in different ways. According to Ting (2010: 3), ‘CLIL advocates a 50:50/Content: Language CLIL-equilibrium’. However, research conducted in actual CLIL classrooms shows that it is difficult to achieve a strict balance of language and content (Dalton-Puffer, 2007; Mehisto, 2008; Pérez-Vidal & Juan-Garau, 2010).

Besides its dual nature, another essential feature of CLIL pedagogy is precisely its diversity (Cenoz, 2015; Coyle, 2010). This diversity of models or formats is a visible

trend in the European context. Coyle (2007, p. 49) considers that there is no model 'which suits all CLIL contexts'. Certainly, CLIL comprises many different variants as it has been implemented in a variety of forms since the 1990s (Cenoz et al., 2014). In relation to this diversity, Coyle (2008, p. 101) also makes it clear that 'there is a lack of cohesion around CLIL pedagogies. There is neither one CLIL approach nor one theory of CLIL'. In the same vein, Mehisto et al. (2008, p. 12) claim that CLIL implementation takes different forms as 'CLIL is an umbrella term covering a dozen or more educational approaches (e.g. immersion, bilingual education, multilingual education, language showers and enriched language programmes)'. Similarly, Ruiz de Zarobe (2013, p. 233) states that almost all EU states implement some form of CLIL with varying degrees of success, responding in different ways (Eurydice, 2006): 'Under the acronym CLIL we recognize a wide range of models, which show divergences as regards the age of implementation of the model or the intensity of the exposure to the foreign language (...), to name but a few differences'. According to Cenoz (2015, p. 21), 'There is great diversity in the implementation of CBI/CLIL programmes and these programmes are dynamic and change because they have to keep up with new challenges in society'. To be more specific, such diversity can be seen in the differences observed in teaching methodology, as some programmes are more content-oriented than others (Cenoz, 2015). In this respect, Kerstin (2013) also concludes that discrepancies in results obtained across CLIL contexts in Europe may be due to nation-specific contextual factors such as policy framework, teacher education, age of implementation and extramural exposure to English. Lastly, Cenoz et al. (2014, p. 258) also suggest that 'Rather than insisting on the uniqueness of CLIL, efforts might be better spent establishing a taxonomy of different common forms of CLIL/CBI so as to circumscribe the diverse contexts in which CLIL is found'. Needless to say, this diversity of CLIL programme formats also involves great challenges when carrying out research on CLIL, as Cenoz et al. (2014) rightly point out.

Much of what we currently know about CLIL approach comes from Applied Linguistics research (Marsh & Frigols, 2013). Specifically, Second Language Acquisition (SLA) research studies provide, as Lasagabaster and López (2015, p. 43) remind us, 'some arguments in favour of CLIL programmes on the grounds that they create conditions for naturalistic language learning, increase the time of exposure to the foreign language and provide an aim for language use in the classroom (Dalton-Puffer & Smit, 2007)'. While several theoretical arguments propose that CLIL promotes content learning, other theories, in contrast, suggest negative effects. Based on information processing theories, Piesche et al. (2016, p. 109) argue that bilingual education in general and CLIL contexts in particular are assumed to lead to 'a greater cognitive control and selective attention which prevents the working memory from being overloaded and thus leads to more effective cognitive processes'. Additionally, it is also made clear that bilingual students are expected 'to process information more deeply because they have to invest more mental effort' (Piesche et al., 2016, p. 109). On the contrary, the perspective of cognitive load theory (Sweller et al., 2011) sustains that students' working memory is overloaded by simultaneously processing new content and the foreign language.

The potential of CLIL in terms of linguistic and cognitive benefits has been fully discussed and documented within the international research literature over the last two decades (Casal & Moore, 2009; Cenoz, 2015; Coyle, 2002; Coyle et al., 2010; Dalton-Puffer, 2008; Halbach, 2008; Lasagabaster & Ruiz de Zarobe, 2010; Muñoz, 2002; Madrid & Hughes, 2011; Pérez Cañado, 2016, 2017, 2018; cf. also Chapters ‘CLIL and L1 Competence Development’, ‘The Impact of CLIL on FL Grammar and Vocabulary’, and ‘The Effects of CLIL on FL Learning: A Longitudinal Study’ in this volume). However, recent studies seem to move beyond the impact of CLIL in terms of linguistic benefits and, consequently, address its effects on the development of content subject knowledge, which has been a neglected research area so far. In relation to this, Nikula (2016) makes it clear that there seems to be a shift in emphasis in CLIL research from studies focusing exclusively on the potential of CLIL in terms of L2 learning outcomes to studies that point towards the need to adopt a truly integrated view on language and content, thus exploring the effects of CLIL on the development of content subject learning. Very little is known for certain about the effects of CLIL on the development of subject matter knowledge (Dalton-Puffer, 2011). Such effects still remain unclear, as Nikula and Mard-Miettinen (2014, p. 14) rightly state: ‘the overall image to date remains rather inconclusive, suggesting that this is an area where more research is needed’. By comparing different models of bilingual education, Cenoz et al. (2014, p. 10) also recognise that ‘there is a greater focus on language than on academic achievement in Canadian immersion research, the same can be said of research on CLIL where research on content is extremely limited’. This argument is also supported by Pérez Cañado (2012, p. 315) who claims that ‘there is still a well-documented paucity of research in this area’. Such lack of research into content outcomes may be due to the fact that ‘CLIL research is conducted by language educators rather than subject specialists, and therefore focuses almost exclusively on language, with content knowledge rarely examined or measured’ (Paran, 2013, p. 323). In the same vein, Cenoz et al. (2014, p. 257) argue that ‘Specifically, much, if not most, research on CLIL has been conducted by ESL/EFL scholars who have compared CLIL and non-CLIL groups of learners and reported higher achievement in English for CLIL learners’. Similarly, Fernández-Sanjurjo et al. (2017, p. 1) argue that ‘So far, CLIL research has focused primarily on language attainment in the L2 and the L1, but students’ achievements as regards content subjects have been largely ignored’. Accordingly, as Dallinger et al. (2016, p. 25) conclude, ‘the effects of CLIL on content learning remain an open question’. What becomes clear is future CLIL research agenda needs to address this under-researched strand in depth (Cenoz et al., 2014; Lasagabaster & Ruiz de Zarobe, 2010; Paran, 2013; Pérez Cañado, 2016, 2017) because ‘we simply do not have enough evidence’ (Paran, 2013, p. 331). The few existing research studies focusing on CLIL-effects on academic content learning to date are in fact contradictory and present mixed results, as Piesche et al. (2016, p. 109) remind us. In the European CLIL context, while most of the studies conducted so far report positive outcomes for academic content learning, other studies have recently found no differences between bilingual and non-bilingual students in terms of content knowledge and some studies have even revealed negative effects on content subject competence, as will be described below in further detail. Accordingly,

as will be examined below, CLIL research offers contradictory results which vary across European contexts.

Overall, CLIL research has provided empirical evidence on the benefits of CLIL education on content learning, concluding that bilingual learners assimilate the content of the academic subjects at the same pace or even better than their non-bilingual counterparts (Jäppinen, 2005; Murray, 2010; Madrid & Hughes, 2011; Mattheoudakis et al., 2014; Ouazizi, 2016; Pérez Cañado, 2012, 2016, 2017, 2018; Serra, 2007; Surmont et al., 2016; Ullmann, 1999; Wode, 1999; Xanthou, 2011). The study by Ullmann (1999) in the United Kingdom reveals that CLIL Secondary Education students assimilating subject-contents (in History and Geography) show enhanced subject matter learning. In Germany, Wode (1999) also reports that CLIL Secondary Education students perform better in Geography and History than their monolingual peers. The longitudinal study by Jäppinen (2005) conducted in Finland concludes that content subject learning, Maths and Science particularly, might be promoted by CLIL as a result of the stimulation of cognition/thinking processes which seem to have positive repercussions on subject matter learning. Similarly, the longitudinal study by Serra (2007) in Switzerland reveals that CLIL Primary Education students obtain higher scores in Mathematics than their non-CLIL counterparts. The cross-sectional study performed in Spain by Madrid (2011) reports that CLIL students learning History and Geography perform better than their non-CLIL peers. Identical results are also reported by Xanthou (2011), whose study with CLIL Primary Education students in Cyprus shows that assimilating academic contents (Science, specifically) through English is beneficial. In the same vein, the study by Madrid and Hughes (2011) in Spain also provides positive results for CLIL in terms of academic content learning (Science in Primary Education and Social Science in Secondary Education) by factoring in type of school as an intervening variable, as in the present study. The study by Mattheoudakis et al. (2014) in Greece, where CLIL was introduced as a pilot project in 2010, also reveals both language and content gains for CLIL students learning Geography in the context of Primary Education. Ouazizi (2016) also reports positive findings with CLIL Secondary Education students learning Mathematics in Belgium, concluding that CLIL exerts a positive influence on content knowledge due to the cognitive benefits of CLIL which seems to stimulate cognitive flexibility (Coyle et al., 2010) and/or cognitive development. Lastly, Surmont et al.'s (2016) study carried out in Belgium also reports that CLIL appears to have a positive impact on the mathematical knowledge of Secondary Education students, even after a very short period of time (three months). As shown by all these studies conducted in various European countries, subject matter knowledge is positively affected by the CLIL approach. Contrary to what might be expected, Van de Craen et al. (2007) hold that subject matter knowledge is not of less quality in CLIL than in traditional education.

A neutral position is also visible in the present discussion as different studies have reported no differences between monolingually and bilingually educated students concerning their content subject knowledge. For example, Bergroth (2006) argues that CLIL students learning Mathematics in Finland do not obtain lower results than their non-CLIL counterparts when finishing their Secondary Education studies,

and indeed perform just as well as their non-CLIL peers. In the same vein, the longitudinal study conducted in the Netherlands by Admiraal et al. (2006) also reports that no negative impact was found in CLIL Secondary Education students' content knowledge in History and Geography. Similar results are also reported by Stehler (2006) in Switzerland, who concludes that CLIL has neither a positive nor a negative influence on academic content knowledge. The quantitative and qualitative research study carried out by Alonso et al. (2008) in the Basque autonomous community in Spain relating to the effectiveness of plurilingual education through CLIL approach in Secondary schools concludes that the assimilation of academic contents taught in English is similar, if not superior, to those relating to non-CLIL students. All in all, these studies reveal that academic content knowledge is not threatened by CLIL in view of a lack of differences observed between both student cohorts (CLIL/non-CLIL).

At the opposite end of the debate are those recent studies which report the negative effects of CLIL on content subject knowledge (Anghel et al., 2016; Dallinger et al., 2016; Fernández-Sanjurjo et al., 2017; Piesche et al., 2016; Sotoca, 2014). For example, the study by Sotoca (2014) conducted in bilingual and non-bilingual public Primary Education schools in Madrid (Spain) reports statistically significant differences in favour of non-bilingual schools in Science, which may be due, according to the author, to a greater level of exigency for academic subjects in bilingual schools. The study carried out in Spain by Anghel et al. (2016) factored in parents' educational level and also reveals significantly negative effects in Natural Science knowledge for those CLIL Primary Education students of less educated parents. Another research study also conducted in Spain is that of Fernández-Sanjurjo et al. (2017), who conclude that monolingual students learning Science achieve better results than their bilingual peers. Similarly, the study performed in Germany by Piesche et al. (2016) shows that monolingually educated students outperform bilingually educated ones in learning Science, although it is also made clear that the negative effects of CLIL on students' content learning are small. In the same vein, Dallinger et al. (2016) in Germany also report a negative CLIL-effect on content learning, concluding that CLIL students progress more slowly and need to receive more input to achieve the same results in terms of content learning. In short, all these studies report a detrimental effect of CLIL education on academic achievement. Additionally, several research studies also point towards students' difficulties in expressing subject knowledge through the foreign language (Jäppinen, 2005; Piesche et al., 2016). Perhaps one convincing reason for this CLIL negative effect might be, as Marsh et al. (2000) identified, the high linguistic demands of the content areas.

Once the initial euphoria of this innovative educational approach has passed since its emergence on the European scene in 1994, a more critical attitude has recently emerged in response to the need to address some 'problematic issues of CLIL' (Paran, 2013, p. 334), calling into question certain controversial aspects or challenges. In relation to the present CLIL research scenario, Pérez Cañado (2018, p. 20) argues that

the so-called 'pendulum effect' (Pérez Cañado, 2016, p. 1) can be seen at work within the CLIL research scenario, as we have moved from an initial period of unbridled enthusiasm and 'celebratory rhetoric' (Paran, 2013, p. 334) on the effects of CLIL to a more critical moment (...) a much more pessimistic outlook on CLIL implementation.

Despite the widely recognised benefits attributed to CLIL approach, certain critical voices have recently warned about the possible drawbacks of this approach (Bruton, 2011, 2013, 2015), thus making it clear that the initial enthusiasm for CLIL should not neglect the real challenges of this new educational approach (Fernández-Sanjurjo et al., 2017; Paran, 2013; Pérez Cañado, 2016, 2017). By exclusively insisting on the uniqueness and potential of CLIL (often without substantial empirical evidence), what is true is that CLIL shortcomings have not been addressed in detail. Hence, Cenoz et al. (2014, p. 256) recognise that 'There is a need for more balanced reflection on both the strengths and shortcomings or gaps in our understanding of CLIL and its effectiveness in diverse contexts'. Given the ambiguity of CLIL, researchers like Cenoz (2013) and Cenoz et al. (2014) demand more critical research beyond exclusively analysing CLIL language gains. While it is true that research results have confirmed the benefits of CLIL in terms of L2 competence, the effects of CLIL on content knowledge, in contrast, still remain an open research question, an unexplored research terrain, as little is known for certain about its real effects on the development of the subject content knowledge (Cenoz et al., 2014; Lasagabaster & Ruiz de Zarobe, 2010). In relation to this under-researched topic, Pérez Cañado (2018, p. 20) concludes that 'the research carried out thus far presents potentially serious flaws which could compromise the validity of its outcomes'. Since CLIL research has recently pointed out the neglect of influential intervening variables which need to be examined in detail, Pérez Cañado (2012, p. 330) argues that there is a 'need of solid empirical research which builds in rigorous assessment of the variables under scrutiny (...) to determine whether the gains observed are truly ascribable to CLIL practice'. Further investigation is also needed on the way language and content are integrated into CLIL classrooms. In view of such empirical gaps in our understanding of CLIL effectiveness, Cenoz et al. (2014, pp. 256–257) point out that 'Without empirical evidence concerning these issues, we simply do not know (...) there is a need to examine more carefully if content is acquired to the same extent when taught through the medium of the L2 in comparison with students' native language'. Additionally, Cenoz et al. (2014, p. 257) also clarify that 'Although these results provide general support for CLIL (although see Bruton, 2011 for an opposing view), they do not establish a clear causal link between integrated language and content teaching and learner outcomes'. Before leaving this discussion, it is undeniable that the development of CLIL pedagogy in the European context presents both strengths and weaknesses, hence the need for a more critical classroom-based research on CLIL, as Cenoz et al. (2014, pp. 258–259) suggest,

We believe that it is time for CLIL scholars to move from celebration to a critical empirical examination of CLIL in its diverse forms to better identify its strengths and weaknesses in different learning contexts (...) In other words, research is needed that goes beyond examining simply whether teaching content in an L2 or a foreign language promotes L2 competence to examining how teaching content in an L2 works and how it can be improved.

Classroom-based research on how best to integrate language and content is necessary if we are to enhance teacher effectiveness in CLIL settings (...) However, there are many aspects of the integration of language and content instruction that require careful theoretical, empirical, and pedagogical attention.

2 Research Questions

Given the scarcity of research studies addressing the effects of CLIL approach on content subject learning in monolingual contexts in Spain (Anghel et al., 2016; Fernández-Sanjurjo et al., 2017; Madrid & Hughes, 2011; Pérez Cañado, 2018; Sotoca, 2014), this chapter aims to shed some light on this still under-researched topic, assessing whether CLIL programmes water down content subject knowledge or rather promote it as successfully as in monolingual streams. Bearing in mind the literature reviewed so far on the effects of CLIL on content subject knowledge, this chapter aims to address the following research questions:

- RQ1: Does CLIL education positively or negatively affect subject content knowledge?
- RQ2: Does CLIL education lead to equal or better subject matter knowledge than traditional education?
- RQ3: When do positive CLIL-effects become visible, in the short or long term?
- RQ4: What is the differential effect exerted on the Primary and Secondary CLIL students' Science learning outcomes by the following two intervening variables: type of school (public and charter) and educational stage (Primary and Secondary Education)?

3 Method

This study forms part of a broader research project focusing on a three-year longitudinal large-scale evaluation of CLIL programmes carried out in those Spanish monolingual communities with the least tradition in bilingual education (Andalusia, Extremadura and the Canary Islands). In view of the scarce research literature available so far on the effects of CLIL on subject matter learning which presents contradictory empirical evidence, the main emphasis of this quantitative study is on the impact of CLIL education on students' Science subject knowledge at the end of Primary (6th grade) and Secondary (4th grade) Education.

Efforts have been made to ensure the homogeneity of the experimental (CLIL) and control (non-CLIL) groups in terms of motivation, verbal intelligence and English level. Pre-, post- and delayed post-tests were administered to Primary and Secondary Education students. In view of the very limited number of research studies focused on controlling the differential effect of particular intervening variables, factor and discriminant analyses were consequently conducted to ascertain the relationship or interaction between CLIL education and the intervening variables under control in

this study (type of school—public and charter—and educational stage—Primary and Secondary Education—) which may account for the differences detected between both student cohorts. To be more specific, dependent (content subject learning results), independent (CLIL programmes) and intervening (type of school and educational level) variables have been taken into consideration in the present study so as to determine whether CLIL is truly responsible for the potential differences observed or whether the aforementioned intervening variables can account for a greater proportion of the variance. Lastly, Cohen's d was employed to measure effect sizes.

3.1 Context and Participants

The context of the present study is the monolingual autonomous community of Extremadura, which is situated in the south-west of Spain, on the border with Portugal, and which has very little tradition in bilingual education (from 2004 onwards). At the present time there are 274 CLIL schools in Extremadura at Primary and Secondary Education stages.

The sample under control comprises 318 students from 10 schools (public and charter). The control group (non-CLIL) consists of 162 learners, while the remaining 156 learners form the experimental group (CLIL). Accordingly, the achievement results of both student cohorts on the subjects of Science in Primary Education and Natural Science in Secondary Education are compared across schools, examining the impact of CLIL on the intervening contextual variables (type of school and educational level). It is noteworthy that no private school participated in the present study, so the comparison with this type of school has not been possible in Extremadura. Table 1 provides an outline of the participating sample.

3.2 Instrument

The data were gathered through an initial questionnaire aimed at collecting personal data on the participants such as age and educational stage. Science subject knowledge was measured by CLIL students' final grades provided by the participating schools out of a total score of 10, which is the highest grade in the Spanish educational system.

4 Results and Discussion

RQ1 investigates whether subject matter knowledge is positively or negatively affected by CLIL education. As can be observed in Table 2, the results of our

Table 1 The research sample

		<i>N</i>	Mean (%)
Student cohorts			
	CLIL	156	49.1
	Non-CLIL	162	50.9
Gender			
	Male	174	54.7
	Female	144	45.3
Educational stages			
	Primary Education (6th grade)	162	50.9
	Compulsory Secondary Education (4th grade)	156	49.1
Type of school			
	Public	252	79.2
	Charter	66	20.8
Setting			
	Urban	113	35.5
	Rural	205	64.5

Table 2 Mean difference scores of the experimental (CLIL) and control (non-CLIL) groups on the subject matter achievement results at both educational stages

Educational level	Group	Mean	Standard deviation	Cohen's <i>d</i>	<i>p</i> value
Primary Education	Non-CLIL	6.98	1.694	-0.110	.528
	CLIL	7.16	1.587		
Secondary Education	Non-CLIL	6.21	1.817	-0.497	.002
	CLIL	7.03	1.443		

analysis confirm the positive effects of CLIL programmes on the development of content subject knowledge by comparing the resulting data of both student cohorts (CLIL/non-CLIL) (Martínez, 2020). This result is backed up by numerous research studies which indicate the positive effects of CLIL education on content subject learning (Jäppinen, 2005; Madrid & Hughes, 2011; Mattheoudakis et al., 2014; Murray, 2010; Ouazizi, 2016; Pérez Cañado, 2012, 2018; Serra, 2007; Surmont et al., 2016; Ullmann, 1999; Wode, 1999; Xanthou, 2011).

Once this positive effect has been reported, RQ2 analyses whether CLIL education leads to equal or better subject matter knowledge than traditional educational approaches, particularly whether bilingually educated students learning Science in Primary Education and Natural Science in Secondary Education perform equally well or outperform their monolingually educated peers. Unlike Fernández-Sanjurjo et al. (2017) and Piesche et al.'s (2016) studies, which show that monolingually educated students perform slightly better than bilingually educated ones when learning subject

matter knowledge, the results of the present study confirm the opposite view, as CLIL students' learning gains are higher than their non-CLIL counterparts' at both educational stages, but especially at the end of Compulsory Secondary Education. According to the results of the present study, bilingual learners assimilate the subject matter content at more or less the same pace in Primary Education, but clearly outperform their non-bilingual peers at the end of Compulsory Secondary Education. This is in line with several research studies which show comparable or even better results between both student cohorts (CLIL/non-CLIL) regarding content subject knowledge. For example, Mattheoudakis et al. (2014) confirm that content knowledge is clearly not negatively affected by CLIL education, reporting that CLIL Greek learners score higher than their non-CLIL counterparts in Geography tests. In the same vein, Ouazizi (2016) also concludes that CLIL education leads to better subject matter knowledge than traditional learning approaches, as Belgian CLIL students obtain better scores than monolinguals in Mathematics knowledge. Such a difference in global performance between both student cohorts may be due, among other aspects, both to the prior careful selection and to the high motivation and interest on the part of the families and students involved in such bilingual programmes, as suggested by Alonso et al. (2008). As can be seen in Table 2, while no statistically significant differences emerge between the experimental (CLIL) and control (non-CLIL) groups at the end of Primary Education as Cohen's d is quite low, the differences between both cohorts are, in contrast, statistically significant when finishing their Secondary Education studies, with a higher Cohen's d .

Given that RQ3 addresses whether the impact of CLIL education on content learning becomes visible in the short or long term, the results of this study confirm that the differences in academic achievement results between the experimental group (CLIL) and the control group (non-CLIL) are higher or become more visible in the long term, particularly when finishing their Secondary Education studies, in line with other studies in the Spanish context (Alonso et al., 2008; Madrid & Hughes, 2011; Pérez Cañado, 2018). While CLIL students obtain similar scores or slightly outperform their non-CLIL peers concerning Science knowledge at the end of Primary Education, bilingually educated students clearly outstrip their monolingually educated counterparts when finishing their Secondary Education studies. This difference in achievement results seems to become more visible as time goes by. Perhaps this may be due to the influence of accumulated experience in bilingual education. In relation to the impact of such experience, the study by Piesche et al. (2016) reminds us of the possible negative effects of CLIL education on content learning for students without CLIL experience. In short, this study reveals that the positive effects of CLIL education require a longer period of time, after which they will become more visible. However, this finding is not congruent with those obtained by Van de Craen et al. (2007), who reported that subject matter learning through CLIL education seems to be boosted more significantly in Primary Education than in Secondary Education. This result is not congruent either with the longitudinal study by Surmont et al. (2016), who conclude, in contrast to their expectations, that CLIL education's positive effects become visible even after a very short period of time (three months).

Our last research question inquires into the differential effect which the intervening contextual variables under control in this study (type of school and educational stage) exert on the Primary and Secondary CLIL students' Science learning outcomes. Consequently, public bilingual and monolingual schools, as well as charter monolingual ones, were compared in this study.

Considering only public bilingual schools, bilingual students achieve better results in Science than their non-bilingual counterparts at the end of both educational stages, which corroborates the benefits of CLIL in terms of content subject learning. Such results corroborate Madrid and Hughes's (2011) and Pérez Cañado's (2018) findings, relating to the fact that bilingual strands outstrip monolingual ones at the end of Primary and Secondary Education in public schools. While no substantial differences were observed between both student cohorts at the end of Primary Education, statistically significant differences were, in contrast, found in favour of CLIL learners when finishing their Secondary Education studies, in this case, with a higher Cohen's *d*, as can be seen in Table 3.

By comparing both public bilingual schools and charter non-bilingual schools, the resulting data surprisingly reveal that charter monolingual schools obtain slightly better results in Science than public bilingual schools only at the end of Primary Education. In this respect, unsubstantial differences were detected in view of the low Cohen's *d*. However, statistically significant differences between public bilingual branches and charter non-bilingual ones were found in favour of the former at the end of Compulsory Secondary Education, with a higher Cohen's *d*. In other words, the results of this study suggest that public bilingual schools outstrip charter non-bilingual schools only at the end of Secondary Education, which confirms once again that the positive effects of CLIL on content learning are mainly witnessed or verified in the long term. However, this finding completely differs from Madrid and Hughes's (2011) study, in which charter monolingual schools obtained significantly better results at the end of Compulsory Secondary Education, thus outperforming public bilingual ones even in the long term.

Lastly, it is noted that charter non-bilingual schools outperform public non-bilingual ones at the end of Primary Education, although the differences cannot be considered statistically significant, with a relatively low Cohen's *d*. However, similar results are obtained by both cohorts in both types of schools at the end of Secondary Education, with an extremely low Cohen's *d*. Consequently, such a finding is not congruent with that obtained by Madrid and Hughes (2011), who reported that the public monolingual strands lagged behind the rest of the groups at both educational stages.

Based on the discriminant analyses performed, statistically significant differences in fact emerge between the experimental (CLIL) and control (non-CLIL) groups in terms of the different intervening variables in this study (type of school and educational level). As can be seen in the tests of equality of group means, the discriminating potential of such variables becomes visible. To be more specific, Wilks' Lambda test reports that there are differences between the mean scores of both student cohorts on the content subject results, particularly at the end of Compulsory Secondary Education. In short, such differences between the experimental (CLIL) and control

Table 3 Subject content results according to educational level and type of school

Educational level	Type of school/Group	Mean	Standard deviation	Cohen's <i>d</i>	<i>p</i> value
Primary Education	Public school/Non-CLIL	6.72	1.508	-0.281	.178
	Public school/CLIL	7.16	1.587		
	Charter school/Non-CLIL	7.35	1.898	0.114	.638
	Public school/CLIL	7.16	1.587		
	Public school/Non-CLIL	6.72	1.508	-0.375	.177
	Charter school/Non-CLIL	7.35	1.898		
Educational level	Type of school/Group	Mean	Standard deviation	Cohen's <i>d</i>	<i>p</i> value
Secondary Education	Public school/Non-CLIL	6.22	1.821	-0.497	.005
	Public school/CLIL	7.03	1.443		
	Charter school/Non-CLIL	6.17	1.855	-0.562	.035
	Public school/CLIL	7.03	1.443		
	Public school/Non-CLIL	6.22	1.821	0.027	.915
	Charter school/Non-CLIL	6.17	1.855		

(non-CLIL) groups cannot be exclusively ascribed to the impact of CLIL education, as the type of school and educational level contextual variables also have a significant influence in explaining the differences found between both cohorts, as can be seen in Tables 4 and 5.

All in all, statistical analysis allows us to conclude that CLIL education does not negatively affect subject content knowledge but rather the opposite. Additionally, the

Table 4 Test of equality of group means

Primary Education	Wilk's Lambda	<i>F</i>	df1	df2	Sig.
Type of school	.568	121.481	1	160	.000
Secondary Education	Wilk's Lambda	<i>F</i>	df1	df2	Sig.
Type of school	.882	20.546	1	154	.000
Subject content results	.941	9.596	1	154	.002

Table 5 Summary of canonical discriminant functions

Primary education				
Function	Eigenvalue	% of variance	Cumulative %	Canonical correlation
1	.759	100.0	100.0	.657
Test of function	Wilks' Lambda	Chi-square	df	Sig.
1	.568	90.100	1	.000
Secondary education				
Function	Eigenvalue	% of variance	Cumulative %	Canonical correlation
1	.194	100.0	100.0	.403
Test of function	Wilks' Lambda	Chi-square	df	Sig.
1	.838	27.129	2	.000

results also reveal that the effects of CLIL education (as an independent variable) are substantial on the content subject learning results (as the dependent variable), especially at the end of Compulsory Secondary Education; that is, positive CLIL-effects are particularly felt in the long term, which corroborates Pérez Cañado's (2018) findings. The discriminating potential of type of school and educational level (as intervening contextual variables) may also account for the differences detected between the experimental (CLIL) and control (non-CLIL) groups.

5 Conclusion

After reviewing what the research literature has revealed so far, the present study aims to shed some light on this unexplored research topic, providing updated empirical evidence on the positive effects of CLIL education on the development of subject matter knowledge in Primary and Secondary Education when compared to traditional educational approaches in a monolingual Spanish region with very little bilingual education tradition.

In response to RQ1, the results of the present study confirm that subject matter knowledge is not diminished or detrimentally affected by the impact of CLIL education, but quite the opposite. Turning now to RQ2 and in view of the data obtained, it can be concluded that CLIL education strands lead to better subject matter knowledge than traditional mainstream school programmes. While no substantial differences are found at the end of Primary Education, statistically significant differences are, in contrast, detected when finishing their Secondary Education studies. In relation to RQ3, the results suggest that positive CLIL-effects are clearly observable in the experimental group (CLIL students) at the end of Compulsory Secondary Education, which indicates that positive CLIL-effects become more noticeable in the long term. As regards the last RQ, the results allow us to state that the two intervening contextual variables (type of school and educational level) have a discriminating

potential, as bilingually educated students obtain better results than monolingually educated students at both educational stages in public schools only. However, when comparing both public bilingual and charter non-bilingual schools, the results surprisingly reveal that there are clear-cut differences in the learning achievement results of both student cohorts depending on type of school and educational stage: while charter non-bilingual strands outperform public bilingual schools at the end of Primary Education (unsubstantial differences), the bilingually educated students' learning gains, in contrast, are higher than the monolingually educated ones when finishing their Secondary Education studies (statistically significant differences). Perhaps a possible convincing reason for the higher scores of bilingually educated students in Natural Science at the end of Compulsory Secondary Education lies precisely in the value of the accumulated experience in bilingual education. To a lesser extent, the discriminating potential of the intervening variables is also observed when comparing both non-bilingual charter schools and non-bilingual public ones, since while the former outperform the latter at the end of Primary Education, similar achievement results are surprisingly found at the end of Compulsory Secondary Education in both types of schools. It must be added that no statistically significant differences are in fact detected at both educational stages in both types of schools. All in all, the results of the present study confirm the differential effect of the two intervening contextual variables of this study, which may account for the differences ascertained between both student cohorts.

In short, the findings of this study confirm the educational value and effectiveness of CLIL education in comparison with traditional educational approaches as the experimental group (CLIL) obtains better results in subject matter knowledge than the control group (non-CLIL), especially in the long term. In this chapter, two factors have been targeted as influential for explaining variation in content subject learning results: type of school and educational stage.

Since the current state of CLIL research is somewhat sparse, methodologically limited and contradictory (Cenoz et al., 2014; Dallinger et al., 2016), further longitudinal research studies are certainly needed in this direction to investigate the real impact of CLIL education on the development of subject matter knowledge. Particularly, Mattheoudakis et al. (2014) advocate the need for further investigations into the strategies CLIL learners use in order to comprehend the concepts presented in the foreign language. Future research studies need to address the impact of CLIL on subject matter knowledge, over shorter and longer periods of time, but also in different learning contexts and with different age groups, as suggested by Surmont et al. (2016). Lastly, the emotional impact of CLIL education on subject matter learning, which remains an unexplored research area to date, calls for further investigation in the future so as to be able to understand how and under what affective and contextual conditions content subject learning actually develops.

Acknowledgements This work was supported by the Spanish Ministry of Economy and Competitiveness [research grant number FFI2012-32221] and the Government of Andalusia [research grant number P12-HUM-2348] (Project: *The Effects of Content and Language Integrated Learning in Monolingual Communities: A Large-Scale Evaluation*).

References

- 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.
- Alonso, E., Grisaleña, J., & Campo, A. (2008). Plurilingual education in Secondary schools: Analysis of results. *International CLIL Research Journal, 1*(1), 36–49.
- Anghel, B., Cabrales, A., & Carro, J. M. (2016). Evaluating a bilingual education program in Spain: The impact beyond foreign language learning. *Economic Inquiry, 54*(2), 1202–1223.
- Bergroth, M. (2006). Immersion students in the matriculation examination three years after immersion. In S. Björklund, K. Mard-Miettinen, M. Bergström, & M. Södergard (Eds.), *Exploring dual-focussed education: Integrating language and content for individual and societal needs* (pp. 123–134). Vaasa: University of Vaasa, Centre for Immersion and Multilingualism.
- Bruton, A. (2011). Is CLIL so beneficial, or just selective? *Re-evaluating Some of the Research. System, 39*(4), 523–532. <https://doi.org/10.1016/j.system.2011.08.002>.
- Bruton, A. (2013). CLIL: Some of the reasons why... and why not. *System, 41*, 587–597. <https://doi.org/10.1016/j.system.2013.07.001>.
- Bruton, A. (2015). CLIL: Detail matters in the whole picture. More than a reply to J. Hüttner and U. Smit (2014). *System, 53*, 119–128. <https://doi.org/10.1016/j.system.2015.07.005>.
- Casal, S., & Moore, P. (2009). The Andalusian bilingual sections scheme: Evaluation and consultancy. *International CLIL Research Journal, 1*(2), 36–46.
- Cenoz, J. (2013). Discussion: Towards an educational perspective in CLIL language policy and pedagogical practice. *International Journal of Bilingual Education and Bilingualism, 16*, 389–394. <https://doi.org/10.1080/13670050.2013.777392>.
- Cenoz, J. (2015). Content-based instruction and Content and Language Integrated Learning: The same or different? *Language, Culture & Curriculum, 28*(1), 8–24. <https://doi.org/10.1080/07908318.2014.1000922>.
- Cenoz, J., Genesee, F., & Gorter, D. (2014). Critical analysis of CLIL: Taking stock and looking forward. *Applied Linguistics, 35*(3), 243–262. <https://doi.org/10.1093/applin/amt011>.
- Coyle, D. (2002). Relevance of CLIL to the European Commission's language learning objectives. In D. Marsh (Ed.), *CLIL / EMILE—The European dimension: Actions, trends, and foresight potential* (pp. 27–28). DG Education & Culture, European Commission, University of Jyväskylä.
- Coyle, D. (2007). The CLIL quality change. In D. Marsh & D. Wolff (Eds.), *Diverse contexts-converging goals: CLIL in Europe* (pp. 47–58). Bern: Peter Lang.
- Coyle, D. (2008). CLIL—A pedagogical approach from the European perspective. In M. Van Dusen-Scholl & N. H. Hornberger (Eds.), *Encyclopedia of language and education* (pp. 97–111). New York: Springer.
- Coyle, D. (2010). Preface. In D. Lasagabaster & Y. Ruiz de Zarobe (Eds.), *CLIL in Spain: Implementation, results and teacher training*. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Coyle, D., Hood, P., & Marsh, D. (2010). *CLIL: Content and Language Integrated Learning*. Cambridge: Cambridge University Press.
- Dallinger, S., Jonkmann, K., Hollm, J., & Fieg, C. (2016). The effects of Content and Language Integrated Learning on students' English and history competences—Killing two birds with one stone? *Learning and Instruction, 41*, 23–31.
- Dalton-Puffer, C. (2007). *Discourse in Content and Language Integrated (CLIL) classrooms*. Amsterdam: John Benjamins.
- Dalton-Puffer, C. (2008). Outcomes and processes in Content and Language Integrated Learning (CLIL): Current research from Europe. In W. Delanoy & L. Volkman (Eds.), *Future Perspectives for English Language Teaching* (pp. 139–157). Heidelberg: Carl Winter.
- Dalton-Puffer, C. (2011). Content-and-Language Integrated Learning: From practice to principles? *Annual Review of Applied Linguistics, 31*, 182–204.

- Eurydice. (2006). *Content and Language Integrated Learning (CLIL) at school in Europe*. Brussels: Eurydice.
- Fernández-Sanjurjo, J., Fernández-Costales, A., & Arias, J. M. (2017). Analysing students' content-learning in science in CLIL vs. non-CLIL programmes: Empirical evidence from Spain. *International Journal of Bilingual Education and Bilingualism*, 20, 1–14.
- Halbach, A. (2008). Bilingual methodology in Primary schools. *Revista de Educación*, 346, 455–466.
- 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), 147–168.
- Kerstin, L. (2013). CLIL in Spain—Why does it not work? A metaperspective on CLIL across contexts in Europe. *International Journal of Bilingual Education and Bilingualism*, 16(3), 301–320.
- Lasagabaster, D., & López, R. (2015). The impact of type of approach (CLIL versus EFL) and methodology (book-based versus project work) on motivation. *Porta Linguarum*, 23, 41–57.
- Lasagabaster, D., & Ruiz de Zarobe, Y. (Eds.). (2010). *CLIL in Spain. Implementation, results and teacher training*. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Lasagabaster, D., & Sierra, J. M. (2010). Immersion and CLIL in English: More differences than similarities. *ELT Journal*, 64, 376–395.
- Madrid, D. (2011). Monolingual and bilingual students' competence in social sciences. In D. Madrid & S. Hughes (Eds.), *Studies in bilingual education* (pp. 195–222). Frankfurt-am-Main: Lang.
- Madrid, D., & Hughes, S. (2011). Introduction to bilingual education. In D. Madrid & S. Hughes (Eds.), *Studies in bilingual education* (pp. 17–50). Frankfurt-am-Main: Peter Lang.
- Marsh, D., & Frigols, M. J. (2013). Content and Language Integrated Learning. In C. Chapelle (Ed.), *Blackwell encyclopedia of applied linguistics* (pp. 911–920). Oxford: Blackwell.
- Marsh, H., Hau, K.-T., & Kong, C.-K. (2000). Late immersion and language of instruction in Hong Kong High Schools: Achievement growth in language and nonlanguage subjects. *Harvard Educational Review*, 70, 302–347.
- Martínez, J. d. D. (2020). To what extent does parental educational background affect CLIL learners' content subject learning? Evidence from research in Spain. *International Journal of Bilingual Education and Bilingualism*. Advance online publication. <https://doi.org/10.1080/13670050.2020.1715916>.
- Mattheoudakis, M., Alexiou, T., & Laskaridou, C. (2014). To CLIL or not to CLIL? The case of the 3rd experimental Primary school in Evosmos. In N. Lavidas, T. Alexiou, & A. M. Sougari (Eds.), *Major trends in theoretical and applied linguistics: Selected papers from the 20th international symposium from theoretical and applied linguistics* (pp. 215–233). London: DeGruyter Versitas Publications.
- Mehisto, P., Marsh, D., & Frigols, M. J. (2008). *Uncovering CLIL, Content and Language Integrated Learning in bilingual and multilingual education*. Oxford: Macmillan.
- Muñoz, C. (2002). Relevance & potential of CLIL. In D. Marsh (Ed.), *The European dimension: Actions, trends and foresight potential* (pp. 33–36). Strasbourg, European Commission.
- Murray, D. R. (2010). Irish-medium language immersion programs' effect on mathematics education. *Journal of Mathematics Education at Teachers College*, 1, 28–32.
- Nikula, T. (2016). CLIL: A European approach to bilingual education. In N. Van Deusen-Scholl & S. May (Eds.), *Second and foreign language education* (pp. 1–14). New York: Springer.
- Nikula, T., & K. Mard-Miettinen. (2014). Language learning in immersion and CLIL classrooms. In J.-O. Östman & J. Verschueren (Eds.), *Handbook of pragmatics* (pp. 1–26). John Benjamins Publishing Group.
- Ouazizi, K. (2016). The effects of CLIL education on the subject matter (Mathematics) and the target language (English). *Latin American Journal of Content & Language Integrated Learning*, 9(1), 110–137.
- Paran, A. (2013). Content and Language Integrated Learning: Panacea or policy borrowing myth? *Applied Linguistics Review*, 4(2), 317–342. <https://doi.org/10.1515/applirev-2013-0014>.

- Pérez Cañado, M. L. (2012). CLIL research in Europe: Past, present, and future. *International Journal of Bilingual Education and Bilingualism*, 15, 315–341. <https://doi.org/10.1080/13670050.2011.630064>.
- Pérez Cañado, M. L. (2016). From the CLIL craze to the CLIL conundrum: Addressing the current CLIL controversy. *Bellaterra Journal of Teaching Language and Literature*, 9(1), 9–31. <https://doi.org/10.5565/rev/jtl3.667>.
- Pérez Cañado, M. L. (2017). Stopping the “pendulum effect” in CLIL research: Finding the balance between Pollyanna and Scrooge. *Applied Linguistics Review*, 8(1), 79–99. <https://doi.org/10.1515/applirev-2016-2001>.
- Pérez Cañado, M. L. (2018). The effects of CLIL on L1 and content learning: Updated empirical evidence from monolingual contexts. *Learning and Instruction*, 57, 18–33. <https://doi.org/10.1016/j.learninstruc.2017.12.002>.
- Piesche, N., Jonkmann, K., Fiege, C., & KeBler, J.-U. (2016). CLIL for all? A randomised controlled field experiment with sixthgrade students on the effects of content and language integrated science learning. *Learning and Instruction*, 44, 108–116.
- Ruiz de Zarobe, Y. (2013). CLIL implementation: From policy-makers to individual initiatives. *International Journal of Bilingual Education and Bilingualism*, 16(3), 231–243.
- Serra, C. (2007). Assessing CLIL at Primary school: A longitudinal study. *International Journal of Bilingual Education and Bilingualism*, 10(5), 582–602. <https://doi.org/10.2167/beb461.0>.
- Sotoca, E. (2014). La repercusión del bilingüismo en el rendimiento académico en alumnos de colegios públicos de la Comunidad de Madrid. *Revista Complutense de Educación*, 25(2), 481–500.
- Stehler, U. (2006). The acquisition of knowledge in bilingual learning: An empirical study on the role of language in content learning. *Vienna English Working Papers*, 15(3): 41–46.
- Surmont, J., Struys, E., Van Den Noort, M., & Van De Craen, P. (2016). The effects of CLIL on mathematical content learning: A longitudinal study. *Studies in Second Language Learning and Teaching*, 6(2), 319–337.
- Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. Berlin: Springer.
- Ullmann, M. (1999). History and geography through French: CLIL in a UK Secondary school. In J. Masih (Ed.), *Learning through a foreign language: Models, methods and outcomes* (pp. 96–105). London: Centre for Information on Language Teaching and Research.
- Van de Craen, P., Ceuleers, E., & Mondt, K. (2007). Cognitive development and bilingualism in Primary schools: Teaching Maths in a CLIL environment. In D. Marsh & D. Wolff (Eds.), *Diverse contexts-converging goals: CLIL in Europe* (pp. 185–200). Frankfurt am Main: Peter Lang.
- Wode, H. (1999). Language learning in European immersion classes. In J. Masih (Ed.), *Learning through a foreign language: Models, methods and outcomes* (pp. 16–25). London: Centre for Information on Language Teaching and Research.
- Xanthou, M. (2011). The impact of CLIL on L2 vocabulary development and content knowledge. *English Teaching: Practique and Critique*, 10(4), 116–126.