



OPEN ACCESS

EDITED BY

Máire Ní Riordáin,
University College Cork, Ireland

REVIEWED BY

Janina Kahn-Horwitz,
Oranim Academic College, Israel
María-Carmen Sánchez-Vizcaino,
University of Economics in Bratislava, Slovakia

*CORRESPONDENCE

Evdokia Pittas
✉ pitta.ev@unic.ac.cy

RECEIVED 11 June 2024

ACCEPTED 12 August 2024

PUBLISHED 13 September 2024

CITATION

Pittas E and Tompkins L (2024) A systematic review of student learning outcomes in CLIL in LOTE.
Front. Educ. 9:1447270.
doi: 10.3389/educ.2024.1447270

COPYRIGHT

© 2024 Pittas and Tompkins. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

A systematic review of student learning outcomes in CLIL in LOTE

Evdokia Pittas^{1*} and Leah Tompkins²

¹Department of Education, School of Education, University of Nicosia, Nicosia, Cyprus, ²Department of English Philology, Autonomous University of Madrid, Madrid, Spain

Introduction: This paper aims to provide a first systematic research overview of student learning outcomes in programs teaching school subjects through languages other than English (LOTE) which are not the mother tongue of the students, according to school- or researcher-administered assessments and stakeholder perspectives, following the PRISMA statement. For brevity, we shall refer to these types of programs as CLIL in LOTE, though we have also included programs which use other labels, such as bilingual education or immersion, due to their similarities with those labeled “content and language integrated learning” (CLIL).

Methods: The selected studies, published between November 1994 and December 2023, were identified through the search of SCOPUS and EBSCO. In determining which studies to include in the review, we employed the following selection criteria: (1) articles focusing on children and youth (ages 5–17 years), (2) articles focusing on CLIL programs in LOTE, (3) articles focusing on student achievement, (4) articles focusing on studies that have collected primary data, and (5) studies that used school-/researcher-administered assessments (objective) or self/ hetero-reported measures (subjective). The screening of titles, abstracts and keywords left a final sample of $n = 29$ scientific papers, which were then read exhaustively and assessed for methodological quality.

Results: Most studies (26 of 29) addressed academic and/or linguistic outcomes, with some studies additionally addressing social/cultural outcomes, behavioral/affective outcomes, and/or (meta) cognitive outcomes. Of the learning outcomes reported, 25 (53%) were positive, five (11%) were negative, four (9%) were neutral, eight (17%) were mixed and four (9%) identified factors influencing outcomes.

Discussion: Theoretically, the study contributes to establishing more general theories about the specific role of CLIL in LOTE in students' learning. Empirically, the study outlines pathways for future research on CLIL in LOTE. In practice, the study presents challenges identified by stakeholders to suggest pathways forward in CLIL teaching/learning.

Systematic review registration: Open Science Framework (OSF): <https://osf.io/mc9uj>.

KEYWORDS

CLIL, LOTE, bilingual education, immersion, multilingual education, minorised linguistic varieties, educational outcomes

1 Introduction

The teaching of school subjects through a language different from the mainstream language of instruction has expanded rapidly over the last three decades, in response to both migration patterns and the demands of a global job market and knowledge-based economy, which increasingly require multilingual workers. Programs offering such educational

experiences go by many names (e.g., Content and Language Integrated Learning (CLIL), English as a Medium of Instruction (EMI), bilingual education, dual-language education, language-enriched education, immersion), which are sometimes used interchangeably and other times distinguished by their geographical location or pedagogical approach. However, all share the dual goal of increasing exposure to an additional language and teaching discipline-specific content. This type of schooling has gained special attention in the European context, where CLIL has become “normalized as a mainstream part of [primary and secondary] school curricula” (Hüttner and Smit, 2023, p. 125), typically with English as the medium of instruction due to its perceived importance for students’ academic and professional careers.

The growing presence of CLIL in public schools in Europe has sparked extensive research on student learning outcomes, profiles and affective experiences, among other topics. Like the schools themselves, however, the resulting literature has overwhelmingly focused on programs teaching subjects through English. Indeed, most of the first wave of studies on learning outcomes in CLIL investigated students’ English proficiency, as can be seen in overviews such as those by Dalton-Puffer (2011) and Pérez-Cañado (2012). These studies reported gains for CLIL students over their non-CLIL peers in areas such as reading and listening comprehension (e.g., Brevik and Moe, 2012), writing (e.g., Brevik and Moe, 2012; Ruiz de Zarobe, 2010), vocabulary (e.g., Jexenflicker and Dalton-Puffer, 2010; Lo and Murphy, 2010), and spontaneous oral production (e.g., Admiraal et al., 2006; Lasagabaster, 2008; Ruiz de Zarobe, 2008). As has been pointed out by other scholars (e.g., Dallinger et al., 2016; Dalton-Puffer, 2011), such advantages are expected because CLIL students typically receive greater exposure to the target language than their non-CLIL counterparts, following their regular foreign language program in addition to their CLIL lessons. In an atypical study where exposure was kept constant, CLIL students’ English proficiency was on par with that of their non-CLIL peers, and in fact lower in the area of listening (Pladevall-Ballester and Vallbona, 2016).

In terms of content outcomes, research has been more limited and results more mixed, with some researchers reporting that outcomes for CLIL students were more positive than those of their non-CLIL peers (e.g., Lorenzo et al., 2021; Pérez Cañado, 2018b), others that they were more negative (e.g., Anghel et al., 2016; Fernández Sanjurjo et al., 2018), and still others that they were neutral (e.g., Admiraal et al., 2006; Jäppinen, 2005) or neutral when additional content instruction was provided to CLIL students (Dallinger et al., 2016). A criticism of research into learning outcomes in CLIL has been many studies’ failure to control for potentially confounding variables (see, e.g., Bruton, 2011), such as students’ social backgrounds or previous academic results. These observations have led researchers to redirect some attention to these areas in the last decade, whether by controlling for them in studies on academic and linguistic outcomes (e.g., Dallinger et al., 2016; Pérez Cañado, 2018b) or investigating potential differences in CLIL and non-CLIL student profiles (e.g., Broca, 2016; Mediavilla et al., 2019; Van Mensel et al., 2020).

Another set of studies have addressed students’ affective experiences in CLIL. Stakeholders who witnessed the first years of CLIL implementation in Spain contended that the approach brought gains such as greater motivation, “more willingness to work collaboratively,” “higher personal confidence,” and greater “ability to confront challenges” (Llinares and Dafouz, 2010, p. 97). More recent studies show that students’ motivation is a key factor influencing their

achievement in CLIL (Lo, 2024; Pavón Vázquez, 2018) and that motivation varies across groups of students, with high-performance groups more motivated than average-performance groups (Riera Ventura, 2021), high-exposure (to CLIL) groups more motivated than low-exposure groups (Somers and Llinares, 2021), and urban students more motivated than rural ones (Pavón Vázquez, 2018).

An exhaustive review of research on student outcomes in CLIL goes beyond the scope of this paper, but the above paragraphs have sought to present its main foci as relate to the topic of this paper. To summarize, CLIL programs have been shown to bring language gains to enrolled students, but the extent to which these derive from the CLIL approach itself versus their additional exposure to the target language has yet to be determined (e.g., Dalton-Puffer, 2011; Pladevall-Ballester and Vallbona, 2016). Furthermore, CLILs effects on content learning remain unclear, and the increases in student motivation reported by stakeholders (Llinares and Dafouz, 2010) may not be felt equally by all groups of students (Pavón Vázquez, 2018; Riera Ventura, 2021; Somers and Llinares, 2021). Several questions remain, then, about student outcomes in CLIL generally. Moreover, the overwhelming focus on programs teaching through English means that little is known about how students fare in programs teaching through other languages, as these are less common and their outcomes receive comparatively less attention. We believe that this scenario justifies a systematic review of the research carried out to date on student outcomes in such programs.

In this paper we present a systematic review of extant literature on student outcomes in programs teaching school subjects through additional languages other than English (LOTE), as reported in both stakeholder perspectives and results from school- or researcher-administered assessments, in order to paint a comprehensive overview of the topic. In light of the similarities between programs offering this educational experience (Dalton-Puffer, 2011), and seeking to include studies beyond Europe - where the term “CLIL” is less common -, we have included findings from all contexts in which the language of instruction in several academic subjects is different from the one used in mainstream classrooms and is not English, regardless of the specific label used. This includes programs seeking to teach an additional language, programs seeking to preserve the students’ minoritized heritage language, and programs with both goals, i.e., where both minority-language and majority-language students are enrolled. However, for brevity, we will henceforth refer to all such programs as CLIL in LOTE.

2 Method

The research question that this paper sets out to examine is the following: “What is presently known about student learning outcomes in CLIL in LOTE, according to school- or researcher-administered assessments and stakeholder perspectives?” The present pre-registered systematic review focused on learning outcomes reported from both objective measures, i.e., school- or researcher-administered assessments, and subjective measures, i.e., self- or hetero-reported measures. Self-reported measures gauge students’ perceptions of their own learning, while hetero-reported measures gauge parents’, teachers’ and administrators’ perspectives on students’ learning. Both measures typically employ questionnaire, interview or focus group methods. To illustrate the difference between self- and hetero-reported measures, let us consider two studies which addressed students’ cognitive outcomes. Coyle

(2013) conducted “respectful discussions” with students (a self-reported measure), while Ní Dhiorbháin et al. (2023) interviewed parents (a hetero-reported measure). Both identified positive outcomes: a student in Coyle (2013, p. 247) stated that “I think I learnt probably more because I had to listen much more,” and a parent in Ní Dhiorbháin et al. (2023, p. 9) that “The ability to switch between languages and what it gives them of additional learning, [...] is really beneficial.” As these examples demonstrate, the two measures can offer different insights into the same outcome: while adults completing hetero-reported measures have more cognitive maturity to reflect on students’ learning outcomes, the students themselves are the ones who experience these outcomes.

As mentioned above, the present paper includes studies using objective measures and/or subjective measures (self-reported, hetero-reported), which we consider complementary. Although the literature on CLIL in LOTE is rather scant in general, there were enough studies using these two types of measures to warrant a systematic review. By synthesizing their findings, we hope to provide a holistic view of the state-of-the-art and to help identify any emerging patterns. This systematic review follows the PRISMA statement (Page et al., 2021) for achieving transparency in synthesizing the empirical evidence from the studies selected. To our knowledge, no systematic review on this topic has been carried out or registered to date.

2.1 Search strategy

The electronic bibliographic databases used for conducting a literature search were SCOPUS and EBSCO. We avoided using other databases because of the considerable overlap in published papers. During mid-November and late-December 2023, these databases were searched for articles published between November 16, 1994 and December 31, 2023. After conducting an initial search to identify relevant terms used by different researchers, we proceeded to cross-search two different sets of terms using Boolean logic (OR/AND) within the title, abstract and keyword fields. The following keywords were used: [(“CLIL” OR “Content and Language Integrated Learning” OR “bilingual education” OR “immersion” OR “dual-language education” OR “multilingual education” OR “language-enriched education”) AND (“LOTE” OR “languages other than English” OR “minorised linguistic varieties” OR “minority language*”). During this initial search, we obtained 8,702 articles. After restricting the search to the works published between 1994 and 2023 (as CLIL was established in 1994), to only academic full text journal articles, books and books series, to the subject area of Social Sciences in SCOPUS and to only papers published in English, the search yielded a total of 1,140 articles. These include both final articles and articles in press up to the initial date we concluded our search: December 31, 2023. We further limited the search results using SCOPUS “Filter by Keyword” box.

2.2 Selection criteria

The resulting articles were screened by abstract, title and keywords to determine whether they met the following set of inclusion criteria: (1) articles focusing on children and youth (ages 5–17 years), (2) articles focusing on CLIL programs in LOTE, as defined above, (3) articles addressing student achievement, (4) articles reporting on studies that collected primary data, and (5) articles reporting on

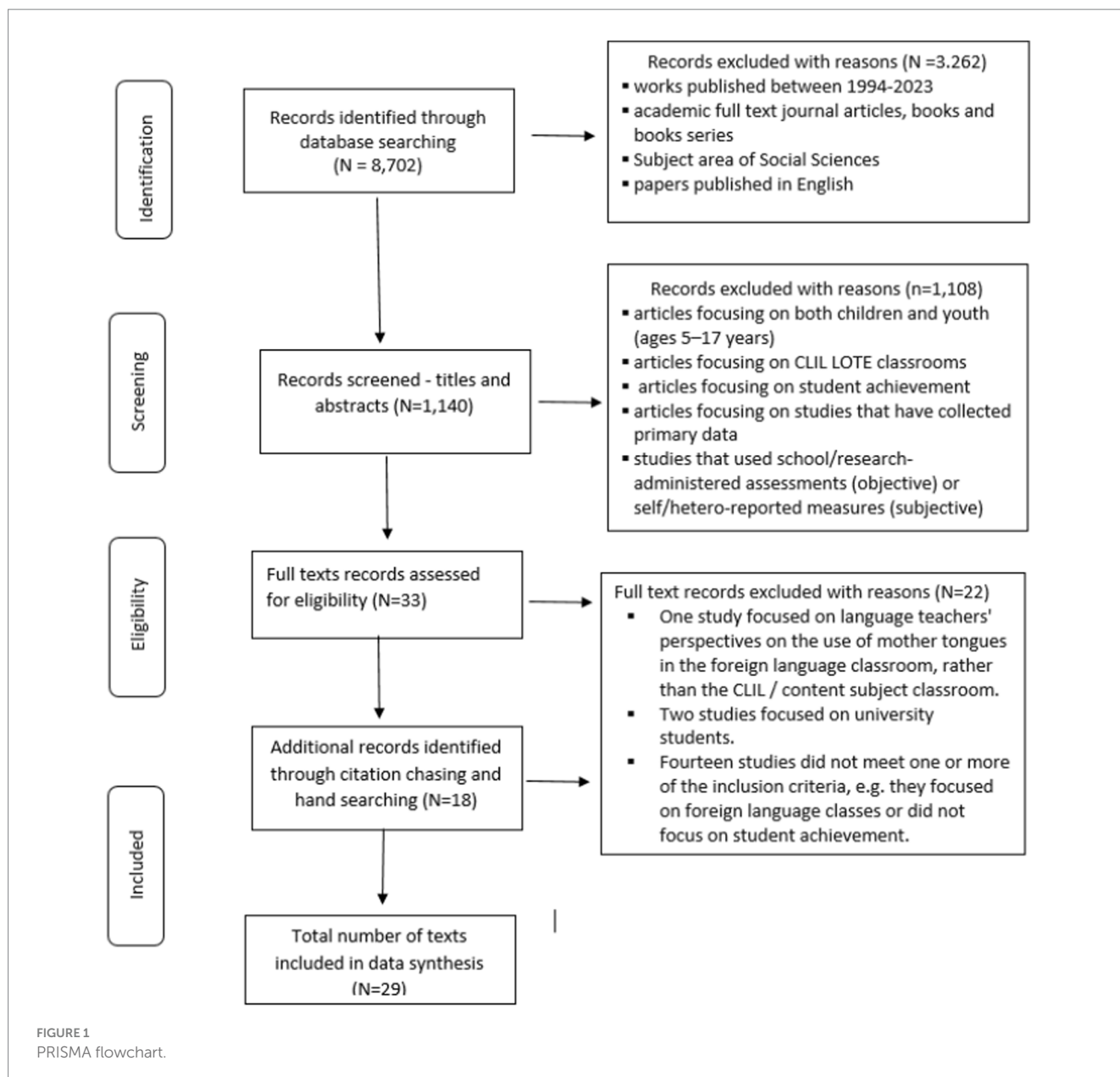
studies that used school-/researcher-administered assessments (objective) or self/hetero-reported measures (subjective). Therefore, any studies not reporting original research, such as reviews, editorials, reports, retrospective studies or opinion papers, were excluded. Conference papers were also excluded. In cases where the authors were uncertain on any of the studies’ eligibility, they screened the studies together for a second time. There were no disagreements among the reviewers. This initial selection yielded 33 studies for potential inclusion in the review. Both authors then made a more thorough reading of the identified articles to confirm their inclusion. Fourteen full text records were excluded with reasons, and an additional 20 that met inclusion criteria were identified through backward reference searching and OSF registries. The number of records identified, screened, included and excluded at every stage of the process, together with reasons for exclusion, are reported using a PRISMA flowchart (Figure 1). The total number of papers included in our final sample is 29.

2.3 Risk of bias assessment

For assessing the methodological quality (risk of bias) of the selected studies, we have used the [National Heart, Lung, and Blood Institute \(NHLBI\) Quality Assessment tools \(2023\)](#). We decided to apply the NHLBI Assessment tool because it addresses methodological issues pertinent to empirical research across disciplines, such as sampling procedures, timing of data collection, and definitions of variables, and it has been successfully used in other systematic reviews (e.g., [La Valle et al., 2022](#); [Pittas and Papanastasiou, 2023](#)). The tool lists criteria for methodological quality in the form of 14 questions to be answered with either “Yes,” “No,” or “Other,” where “other” signifies “cannot determine,” “not applicable,” or “not reported.” An affirmative response indicates a lower risk of obtaining biased results through the methodology employed. Since this tool was designed for use in the medical field, we found it necessary to make minor adaptations for its use in education research, where four criteria (Q8, Q10, Q12, Q13) are less relevant. They received an “other” answer for all studies and were consequently omitted from the list.

Each study was evaluated independently by each researcher. To determine the interrater reliability of our evaluations, we first compared our answers to each item for each article and assigned a score of “1” to those we agreed on and a score of “0” to those we disagreed on. Next, we summed the agreement scores and then divided them by the total number of criteria ($N=10$) to calculate interrater reliability per article. Finally, we summed the interrater reliability scores for all articles, divided this figure by the total number of articles, and multiplied the result by 100 to obtain an overall score. The overall inter-rater reliability, measured as percent agreement, equaled 80% (Table 1).

According to the final evaluations on all criteria, all the studies received a rating of “good,” but some studies were found to be of higher methodological quality than others (see [Supplementary Data Sheet 1](#)). Many of the articles (41%) reported on case studies which did not involve independent or dependent variables ([Bower, 2019, 2020](#); [Coyle, 2013](#); [Davis et al., 2019](#); [Méndez García and Pavón Vázquez, 2012](#); [Gartziarena et al., 2024](#); [Hunt, 2011](#); [Macleod, 2014](#); [Murtagh and Seoighe, 2022](#); [Ní Dhiorbháin et al., 2023](#); [Ozfidan, 2014](#); [Potowski, 2004](#)). These studies thus received a response



of “No” to Q5, about sample size justifications, and a response of “Not applicable” to Q6-Q14, which asked about the nature of the independent and dependent variables used in the study. As (Gorard 2013, p. 39) notes, case studies can provide “exploratory descriptive preparation for subsequent studies,” but their findings cannot offer insight into the *effects* of CLIL in LOTE on student learning outcomes because so few of the elements of research design are present, making it impossible to establish causal or even correlational links. However, given that research in CLIL in LOTE is still in its infancy, we expected to find studies describing, rather than explaining, the current state of learning outcomes in these programs as the first part of the research cycle (Gorard, 2013). We thus deem their methodological quality acceptable at this stage and have included them in our systematic literature review.

For the rest of the studies, which included at least one design element (e.g., comparator groups, time), the largest risk of bias comes from failing to control for other potential confounding variables. To

our knowledge, only three studies took measures to account for their influence (Cape et al., 2021 on cognitive outcomes; Mady, 2015 on linguistic outcomes; and Marian et al., 2013 on academic outcomes). We thus advise that readers treat findings with caution, bearing in mind that they may be influenced by factors other than the effectiveness of CLIL in LOTE. By synthesizing the findings of all extant research on the topic of CLIL in LOTE, i.e., studies with and without design elements, we hope to pinpoint recurring themes to be investigated by future explanatory studies which successfully account for potentially confounding variables.

2.4 Synthesis

We synthesized the studies by pulling together information related to the research aim of this paper (Supplementary Data Sheet 2). The following information was compiled in a spreadsheet for each article:

TABLE 1 Inter-rater reliability of the ratings (max: 10) measured as percent agreement.

Manuscript	Agreement
Gartziarena et al. (2024)	1
Méndez García and Pavón Vázquez (2012)	0.8
Coyle (2013)	1
Ní Dhiorbháin et al. (2023)	1
Bower (2020)	0.9
Bower (2019)	0.9
Davis et al. (2019)	1
Hunt (2011)	0.8
Cross and Gearon (2013)	0.9
Birnie (2022)	0.8
Murtagh and Seoighe (2022)	1
Ozfidan (2014)	0.8
Rehamo and Harrell (2020)	1
Wright and Scullion (2007)	0.8
Macleod (2014)	0.9
Bermingham (2021)	0.8
O'Hanlon et al. (2010)	0.8
Murray (2007)	0.9
Potowski (2004)	1
Manterola et al. (2013)	0.8
Bulon et al. (2017)	0.9
Pérez et al. (2016)	1
Mady (2015)	0.8
Marian et al. (2013)	0.8
Stewart (2011)	0.8
Hermanto et al. (2012)	0.9
Surmont et al. (2016)	0.8
Baten et al. (2020)	0.9
Cape et al. (2021)	0.8
Percent agreement	80% (0.8)

(a) Author/date, (b) Country, (c) Type of school/setting, (d) Language of instruction: type (e.g., foreign, minoritized) and name, (e) Language of students (target, mainstream, either, other) (f) Participants, (g) Type of study (e.g., case, cross-sectional), (h) Type of measure, (i) Data collection tools, and (o) Outcome(s) reported. Since many of the 29 articles addressed different but related topics within the same article, together they yielded 47 different outcomes which were then organized thematically into the five categories described further below.

3 Results

3.1 Demographic information

Across the selected studies, participants (school leaders, administrators, teachers, students, and parents or caregivers) were

involved in programs teaching several school subjects through non-mainstream languages of instruction which were not English. The following labels were used to refer to these programs: content and language integrated learning (CLIL), language-medium (e.g., Irish-medium, Maori-medium), bilingual education, and immersion (e.g., one-way, two-way, total). Fifteen studies were conducted with students only, six studies with teachers/administrators only, one study with parents only and seven studies with several groups of participants. The number of participants in the samples ranged from 4 to 29,479. The samples were drawn from primary and secondary schools. The papers included case studies ($N = 12$), in which one episode of data collection is carried out with one group of cases, cross-sectional studies ($N = 12$), in which one episode of data collection is carried out with two or more groups of cases which are created based on pre-existing exposure status to an independent variable, longitudinal studies ($N = 3$), in which two or more episodes of data collection are carried out with one group of cases, a quasi-experimental study ($N = 1$), in which two or more episodes of data collection (e.g., pre-test and post-test) are carried out with two or more groups of cases created according to their exposure status to an independent variable, and a cohort study ($N = 1$), in which different groups of cases created based on exposure status to an independent variable are followed through time to gauge, the effects of exposure on relevant dependent variables (Gorard, 2013). With reference to the type of measures used in each study, four studies used self-reported measures only, eight studies used hetero-reported measures only, 11 studies used objective measures only and six studies used a combination of measures. In terms of location, the studies were carried out in 10 countries: Australia, Belgium, Canada, China, England, Ireland, New Zealand, Scotland, Spain and USA.

3.2 Student learning outcomes in CLIL in LOTE, according to school-or researcher-administered assessments and stakeholder perspectives

In this section we present the data following narrative synthesis procedures, identifying recurring themes among the studies examined. The findings of our review suggest that the 47 student learning outcomes identified can be classified into five categories: (1) academic outcomes, (2) linguistic outcomes, (3) social and cultural outcomes, (4) behavioral and affective outcomes, and (5) cognitive and metacognitive outcomes. Some categories have been more researched than others, with most studies focusing on either linguistic outcomes (17), academic outcomes (7) or both (2) and then addressing one or more of the other categories as additional issues to take into account. Only three studies addressed these secondary categories without reference to academic or linguistic outcomes (Cape et al., 2021; Méndez García and Pavón Vázquez, 2012; Ní Dhiorbháin et al., 2023).

3.2.1 Academic outcomes

By “academic outcomes,” we refer to students’ performance on assessment tasks for content subjects, or their perceived performance on such tasks. Nine studies reported academic outcomes in CLIL in LOTE. Of these, seven used objective measures and two used subjective measures, either hetero-reported or self-reported. Four of the nine studies reported positive outcomes, two

reported neutral outcomes and three reported negative outcomes for enrolled students.

With reference to the studies using objective measures, three took place in contexts where primarily majority-language students were learning through an additional (foreign or minoritized) language of instruction (Bower, 2020; O'Hanlon et al., 2010; Surmont et al., 2016). Surmont et al. (2016), using a Maths test designed by the researchers, compared the scores of Flemish lower secondary students studying through French ($N = 35$) with those of their non-CLIL peers ($N = 72$) at 0 months, 3 months and 10 months of instruction. Bower (2020), referring to schools' internal monitoring of students' performance, reported the achievement of English year 8/9 students studying through French ($N =$ unspecified) in comparison with their non-CLIL peers ($N =$ unspecified). O'Hanlon et al. (2010), using multi-subject national standardized tests in Scotland, compared the scores of Gaelic-medium primary students ($N = 308$) with those of their English-medium peers ($N = 15,460$) at years 3, 5 and 7. The findings of these studies were positive or neutral academic outcomes for CLIL students. Surmont et al. (2016) found that CLIL students outperformed their non-CLIL peers on the Maths test after 3 and 10 months of instruction, with the greatest improvement being seen between months 0 and 3. In Bower's (2020) study, two schools reported that CLIL students attained higher grades across the curriculum than their non-CLIL peers. O'Hanlon et al. (2010) showed that Gaelic-medium students' achievement in Reading and Science was not harmed by learning through an additional language: their attainment was on par with that of their English-medium peers.

An additional four studies focused on the outcomes of minority-language students learning through their heritage language and the mainstream language of instruction, English, either together with majority-language students (two-way immersion) or separately (Collier and Thomas, 2004; Marian et al., 2013; Murray, 2007; Stewart, 2011). The studies in the United States looked at primary students' scores on national standardized tests. Collier and Thomas (2004) used the first- and second-language Reading scores of Spanish-speaking ($N = 29,319$) and French-speaking ($N = 160$) students in primary 1–5, in order to compare the effectiveness of several bilingual and immersion models in each language. A similar study was carried out by Marian et al. (2013), who analyzed the Reading and Maths scores of minority- and majority-language students in years 3–5 ($N = 2,009$) to gauge the effectiveness of two-way immersion against its alternatives. The studies in New Zealand looked at year 11, 12 and 13 students' scores on national standardized tests in several subjects. Murray (2007) compared the scores of Maori candidates at bilingual schools ($N = 844$) with those of their Maori peers in English-medium programs ($N =$ unspecified). Stewart (2011) compared the performance of Maori-medium students ($N = 1,317$), mainstream Maori candidates ($N =$ unspecified) and non-Maori students ($N =$ unspecified). The findings of these studies varied by context. In the USA, Collier and Thomas (2004) demonstrated that one- and two-way immersion programs enhanced test results for minority-language students over 5 years of schooling and helped close the achievement gap between children with Spanish and English as their home languages. Marian et al. (2013) also identified positive effects of two-way immersion: immersion students who spoke Spanish at home outperformed their counterparts enrolled in transitional programs of instruction, while immersion students who spoke English at home outperformed their counterparts enrolled in monolingual programs.

In New Zealand, Murray (2007) demonstrated that the achievement of Maori secondary students at Maori-medium schools was on par with that of their mainstream Maori peers for Maori language, English and Maths, but they struggled in Science. A few years later, Stewart (2011) found that Maori-medium students underperformed mainstream Maori students in both Science and Maths, and this group in turn underperformed non-Maori students.

The two studies using subjective measures took place in Ireland, where primary students were learning through Irish (Murtagh and Seoighe, 2022; Wright and Scullion, 2007). Murtagh and Seoighe (2022) focused on schools in Irish-speaking communities and used semi-structured interviews to gather primary teachers' ($N = 11$) perspectives on educational psychological services and their relationship with students' performance on standardized tests. Wright and Scullion (2007), using questionnaires to elicit students' self-evaluations of their perceived success in the classroom, compared the responses of Irish-medium ($N = 218$) and English-medium ($N = 205$) students at the end of primary 4 and 7. In terms of findings, Irish-medium students had similar self-perceptions of their success in the classroom to those of their English-medium peers (Wright and Scullion, 2007). As for Irish-medium teachers, they believed that standardized tests written in English produced skewed results for students who spoke Irish at home, which could misinform subsequent educational decisions and negatively impact the children's education (Murtagh and Seoighe, 2022).

3.2.2 Linguistic outcomes

By "linguistic outcomes," we refer to students' reported or documented level of proficiency in any area of the target language, including listening, speaking, reading, writing, vocabulary and grammar. Nineteen studies reported linguistic outcomes in CLIL in LOTE. Of these, eight used objective measures, such as students' performance on exams or in activities, while 11 used subjective measures: hetero-reported (5) or self-reported (6). Eleven of the 19 studies reported positive linguistic outcomes for students in CLIL programs, three reported mixed outcomes (gains in some areas and losses in others), one reported negative outcomes, and four identified factors influencing linguistic outcomes.

The studies using objective measures either compared CLIL students' performance with that of their non-CLIL peers (Bulon et al., 2017; Pérez et al., 2016), tracked students' progress within the same program (Baten et al., 2020; Hermanto et al., 2012) or compared different groups of students within the same program (Birmie, 2022; Mady, 2015; Manterola et al., 2013; Potowski, 2004). They all addressed majority-language students' linguistic outcomes in a foreign or minoritized language, but three of the studies which compared groups of students also referred to the outcomes of heritage speakers of the target language. In terms of studies that compared CLIL and non-CLIL programs, Pérez et al. (2016) used university entrance exams scores to compare the French proficiency of Spanish secondary students enrolled in CLIL programs ($N =$ unspecified) against those of their non-CLIL counterparts ($N =$ unspecified). Bulon et al. (2017), conducting a linguistic analysis of texts written in the CLIL language (Dutch or English) and the first language (French), compared the syntactic and lexical complexity of writing by Belgian students in different programs: Dutch learners in CLIL ($N = 132$) and non-CLIL ($N = 100$), and English learners in CLIL ($N = 90$) and non-CLIL ($N = 90$). The findings of these studies were positive linguistic

outcomes for CLIL students: when compared to their non-CLIL counterparts, CLIL students in Spain earned higher scores on their French proficiency tests (Pérez et al., 2016) and CLIL students in French-speaking Belgium wrote more syntactically and lexically complex texts in Dutch, though this was not the case on all measures for the English texts (Bulon et al., 2017).

In terms of studies that tracked students' progress, Baten et al. (2020) used vocabulary tests in English and French to compare Dutch-speaking Flemish secondary students' ($N=75$) attainment before and after receiving 3 months of CLIL instruction in each language. Hermanto et al. (2012), using linguistic and metalinguistic tests in Canadian primary students' first language (English) and the CLIL language (French), compared their skills in each language in years 2 ($N=50$) and 5 ($N=33$). The findings of these studies were also positive linguistic outcomes for CLIL students. CLIL students in Flanders developed receptive and productive vocabulary at similar rates in both English and French (Baten et al., 2020), and French-immersion students in Canada displayed patterns of development in line with those observed in fully bilingual students: their metalinguistic skills developed at similar rates in French and English, while their formal linguistic skills in French lagged behind (Hermanto et al., 2012).

In terms of studies that compared groups of students, Manterola et al. (2013), conducting a linguistic analysis of a story retelling done by one group of Basque-immersion students at ages 5 and 8, compared the discursive skills of those who spoke Basque as a first language ($N=24$) or second language ($N=37$). Potowski (2004), using classroom observation, student questionnaires and parent interviews in a dual immersion context in the USA, identified factors influencing the Spanish use of year 5 students ($N=4$) with different language backgrounds. Mady (2015), using international language proficiency tests and questionnaires with year 6 French-immersion students in Canada, compared the results of Canadian-born ($N=60$) and foreign-born ($N=30$) students. Birnie (2022), using language proficiency tests and parental questionnaires, compared Gaelic-medium year 1 students' ($N=8$) scores before and after the Covid-19 school closures in Scotland and identified characteristics of those who were more and less successful. The findings of these studies help pinpoint factors which may influence linguistic outcomes. Manterola et al. (2013) showed that the first language of primary students (Basque or Spanish) bore no relationship with their development of discursive skills, which occurred at similar rates in both groups. Similarly, dual immersion students' first language seemed to have less of an influence on their Spanish use in Potowski's (2004) study than affective factors, such as their investments in being perceived positively and in receiving praise. In contrast, French-immersion students' first language did influence their linguistic outcomes in Mady's (2015) study, where immigrant children who spoke a third language at home outperformed their Canadian-born peers on a French proficiency test. Finally, Birnie (2022) showed that engagement with homework was crucial to children's progress in Gaelic during the Covid-19 school closures: children who routinely engaged with Gaelic learning activities at home enjoyed higher levels of language attainment than those who were less engaged.

The studies using subjective measures either referred to teachers' and administrators' perceptions of students' linguistic outcomes or to students' own perceptions. Five focused on foreign language outcomes in programs enrolling (primarily) majority-language students, three on minoritized language outcomes in programs enrolling both

majority-language speakers and heritage speakers of the target language, one on minoritized language outcomes of heritage speakers, and one on minoritized language outcomes of speakers of a third language. We will begin with the studies reporting teachers' and administrators' perspectives. Bower (2020), using semi-structured interviews with school leaders ($N=12$) at English lower secondary schools offering CLIL in French, probed their perspectives on learners' linguistic progress. Hunt (2011), using questionnaires with secondary school subject teachers ($N=17$) in the UK who had prepared and conducted a CLIL lesson in French, German or Spanish as part of the study, gauged their perceptions of several outcomes, including linguistic attainment. Gartziaarena et al. (2024), using online questionnaires ($N=418$) and interviews ($N=20$) with primary teachers in the Basque Country in Spain, investigated their beliefs about the influence of several factors on students' linguistic attainment in Basque. Ozfidan (2014), also using questionnaires with teachers ($N=26$) at all educational levels in the Basque Country, asked about the effectiveness of different bilingual models for enhancing students' linguistic attainment in Basque. Davis et al. (2019), using questionnaires and semi-structured interviews with parents ($N=23$) and teachers ($N=56$) of French-immersion primary children in Canada, gauge, their perceptions of the program's effectiveness for "allophone" students who spoke languages other than English or French at home. The findings of these studies were mostly positive, particularly for majority-language students. Teachers at one school in Bower's (2020) study reported that learners' comprehension and communication skills in French had improved, and that their enhanced linguistic skills were being transferred to other subjects. Similarly, secondary teachers had a positive perception of students' use of vocabulary and syntax in the CLIL lesson they prepared for Hunt's (2011) study, and teachers in Ozfidan's (2014) study believed that programs with more instruction in Basque led to more complete bilingualism in Spanish and Basque. Also in the Basque Country, teachers in Gartziaarena et al.'s (2024) study reported that parents' linguistic attitudes influenced students' learning of Basque, and students' Basque proficiency in turn influenced their academic results. Finally, with regard to the impact of French-immersion on allophone students, educators in Davis et al.'s (2019) study were divided: they believed these students were learning French effectively, but many were less positive about their English acquisition, with some reporting that their lower levels of English proficiency put them at a disadvantage.

Turning to the studies reporting students' perspectives, three studies surveyed one group of students on their self-perceptions of their abilities in the language of instruction at one point in time (Bower, 2019; Coyle, 2013; Rehamo and Harrell, 2020), and three studies compared their perceptions either between programs, languages or before and after an intervention (Cross and Gearon, 2013; Macleod, 2014; Wright and Scullion, 2007). In terms of the studies surveying a single group of students at one point in time, Coyle (2013) conducted questionnaires and interviews with lower secondary students ($N=670$) enrolled in French, German and Spanish CLIL programs in England and Scotland. Bower (2019) used questionnaires and focus groups with lower secondary students ($N=55$) studying in French CLIL programs in England. Rehamo and Harrell (2020) used questionnaires with students at all educational levels ($N=3,500$) enrolled in Nuosu Yi bilingual education programs in China. The findings of these studies were mixed. Coyle (2013) showed that CLIL students in Scotland and England felt they had

improved their speaking skills, translation skills and vocabulary in both their first and second language, but they struggled in writing and perceived writing negatively in CLIL classes at most schools. Students in [Bower's \(2019\)](#) study reported more positive outcomes, perceiving rapid progress in listening, concentration and writing skills in French. In contrast, Nuosu-Chinese bilingual students in [Rehamo and Harrell's \(2020\)](#) study indicated very low confidence levels in their reading ability in Nuosu, which the authors attribute to the limited number of hours of instruction conducted in this language (2–3 per week).

In terms of the comparative studies, [Wright and Scullion \(2007\)](#) used questionnaires with primary 4 and 7 students enrolled in Irish-medium ($N = 218$) and English-medium ($N = 205$) programs in Scotland to compare their self-evaluations of their Irish abilities. [Macleod \(2014\)](#) used questionnaires and interviews with Gaelic-medium students ($N = 18$) in Scotland to compare their self-evaluations of their abilities in Gaelic and English. [Cross and Gearon \(2013\)](#) administered questionnaires to Australian primary and secondary students before and after a “CLIL trial” where they received 5 weeks of instruction in French, German or Spanish. The findings of these studies were also mixed: Irish-medium students had more positive self-evaluations of their abilities in Irish than their English-medium counterparts ([Wright and Scullion, 2007](#)), Gaelic-medium students felt nearly as confident in their Gaelic speaking skills as their English speaking skills but less confident in Gaelic for reading, writing and spelling ([Macleod, 2014](#)), and Australian students felt more confident in their speaking skills but less confident in their listening skills after the “CLIL trial” ([Cross and Gearon, 2013](#)).

3.2.3 Social and cultural outcomes

By “social and cultural outcomes,” we refer to students’ awareness and respect of other cultures and/or pride in their own heritage culture. Seven studies reported on social and cultural outcomes in CLIL in LOTE, with three focusing on primarily majority language students in programs teaching through a foreign language, three on mixed groups (heritage and majority) in programs teaching through a minoritized language, and one on speakers of other languages in a program teaching through a minoritized language. All studies used subjective measures, either hetero-reported (4) or self-reported (3). Five studies reported positive outcomes, one reported mixed outcomes and one reported negative outcomes.

The studies using hetero-reported measures either gauge, teachers’ and administrators’ perspectives ([Bower, 2020](#); [Ozfidan, 2014](#)) or parents’ perspectives ([Ní Dhiorbháin et al., 2023](#); [O’Hanlon et al., 2010](#)) on several issues, among which social and cultural outcomes were included. [Bower \(2020\)](#) conducted semi-structured interviews with leaders ($N = 12$) at French CLIL schools in England, [Ozfidan \(2014\)](#) questionnaires with teachers ($N = 26$) at schools in the Basque country, [O’Hanlon et al. \(2010\)](#) interviews with parents of Gaelic-immersion primary students ($N = 85$) and [Ní Dhiorbháin et al. \(2023\)](#) interviews with multilingual parents of Irish-immersion primary students ($N = 15$). All four studies reported positive social and cultural outcomes. School leaders in [Bower's \(2020\)](#) study reported that learners demonstrated greater intercultural awareness, such as empathy with French perspectives. Teachers in [Ozfidan's \(2014\)](#) study believed full Basque-immersion programs promoted biculturalism and self-esteem around Basque culture. Some interviewed parents in [Ní Dhiorbháin et al.'s \(2023\)](#) study believed that their children's

awareness of different ways of communicating would foster respect for linguistic and cultural diversity, and many parents in [O’Hanlon et al. \(2010\)](#) study chose Gaelic-medium education because they believed it would help preserve their heritage.

As for the studies using self-reported measures, they drew on questionnaire data: [Bower \(2019\)](#) with year 7/8 students ($N = 55$) in French CLIL programs in England, [Cross and Gearon \(2013\)](#) with primary and secondary students before and after they participated in the CLIL trial in Australia ($N = 93$), and [Macleod \(2014\)](#) with primary 5/6 students ($N = 14$) in a Gaelic-medium program in Scotland. Their findings were more mixed. Students in [Bower's \(2019\)](#) study expressed awareness of the link between language and culture, but those in [Cross and Gearon's \(2013\)](#) reduced their expectations about CLIL's ability to contribute their intercultural awareness after receiving CLIL instruction. Finally, the Gaelic-medium students in [Macleod's \(2014\)](#) study held positive attitudes toward bilingualism in general, but expressed uncertainty about the utility of Gaelic beyond education.

3.2.4 Behavioral and affective outcomes

“Behavioral outcomes” refer to observable occurrences in performance or behavior that come from learners’ engagement in different activities, and “affective outcomes” refer to perceptions, attitudes or beliefs in relation to one’s learning experiences such as satisfaction, motivation, etc. ([Wei et al., 2021](#)). Six studies reported behavioral and affective outcomes in CLIL in LOTE, with all six focusing on (primarily) majority-language students studying through a foreign language. All six studies used subjective measures, with three focusing on teachers’ and administrators’ perspectives (hetero-reported) and three focusing on students’ perspectives (self-reported). Three of the six studies reported positive outcomes and three reported mixed outcomes: positive and negative results on different subcategories. In terms of teachers’ and administrators’ perspectives ([Bower, 2020](#); [Méndez García and Pavón Vázquez, 2012](#); [Hunt, 2011](#)), [Méndez García and Pavón Vázquez \(2012\)](#) used semi-structured interviews with French CLIL primary and secondary school teachers in Andalusia, Spain ($N = 15$), to elicit their views on the effects of using two languages in the same classroom. [Bower \(2020\)](#) also used semi-structured interviews to examine school leaders’ ($N = 12$) perspectives on French CLIL in England. [Hunt \(2011\)](#) took a classroom-based action research approach, administering questionnaires to UK teachers ($N = 17$) who had prepared and carried out a CLIL lesson for her study. The findings showed positive behavioral outcomes. In the [Méndez García and Pavón Vázquez \(2012\)](#) study, teachers reported that CLIL learners were better able to cope with classroom routines and tended to work more effectively and efficiently. They believed that the development of these skills resulted from students’ efforts to master both the target language and the content, which in turn reinforced their understanding. In the same vein, teachers participating in [Hunt's \(2011\)](#) study perceived higher levels of student concentration, engagement and time management during the CLIL lesson, and they attributed these changes to the greater cognitive demand of activities carried out in an additional language. As for the affective outcomes, they were also mostly positive. Teachers in [Méndez García and Pavón Vázquez's \(2012\)](#) study reported that students were more motivated to take part in classroom activities, and that using the target language for a specific purpose made learning more enjoyable. Similarly, leaders at the three schools in [Bower's \(2020\)](#) study reported that students displayed greater enjoyment and more positive attitudes

toward learning a language, as well as greater external and internal motivation due to parental pressure and teachers' attitudes. Improvements in students' motivation and attitude were also reported by teachers in Hunt's (2011) study, irrespective of students' learning abilities. However, these gains were partly dependent on students' interests: the use of the target language increased motivation for students who otherwise disliked the subject, but decreased motivation for students who perceived it as an added difficulty.

With reference to students' perspectives, Coyle (2013) used questionnaires and interviews to investigate the perceptions of "successful learning" held by lower secondary CLIL students ($N=670$) at 11 schools across England and Scotland. Cross and Gearon (2013) used questionnaires to examine Australian students' ($N=93$) engagement, motivation and attitudes toward CLIL before and after receiving CLIL instruction. Bower (2019), by means of questionnaires and focus groups, explored the motivation of lower secondary CLIL students ($N=55$) in England. In terms of the findings, positive behavioral outcomes were reported in the Coyle (2013) study, where students reported frequent use of study skills in independent research projects. Students also mentioned that learning content with language promotes engagement, interaction, and opportunities for richer discussions and collaborative learning. The findings for affective outcomes were more mixed. Cross and Gearon's (2013) comparison of students' perspectives before and after receiving CLIL instruction showed an increase in their confidence to learn content through another language and an increase in their perceptions of the effectiveness of CLIL, but a decrease in their enjoyment of learning languages and a decrease in their belief that CLIL enhances creative thinking. The researchers found that academically advanced students with higher levels of self-esteem tended to have more positive attitudes toward CLIL. In Bower's (2019) study, students at two schools reported high levels of concentration, effort, enjoyment and progress, and perceived a high degree of cognitive challenge in CLIL. However, students at the school of higher socioeconomic status viewed this challenge as motivating, while those at the school of lower socioeconomic status viewed it as demotivating. This accords with Cross and Gearon's (2013) observation that a select group of students had more positive attitudes toward learning content through another language.

3.2.5 Cognitive and metacognitive outcomes

"Cognitive outcomes" refer to the acquisition of knowledge and intellectual skills (Bloom, 1956), and "metacognitive outcomes" refer to the acquisition of higher-order thinking skills relevant to the organization and monitoring of learning (Harris, 2000). Furthermore, cognitive strategies are the procedures used to accomplish specific goals, such as understanding a text, and metacognitive strategies are used to confirm that the goals have been accomplished, such as checking that the text is understood (Livingston, 2003). Six studies reported cognitive and metacognitive outcomes in CLIL in LOTE, with four focusing on majority-language students learning through a foreign language, and two on speakers of a third language learning through a minoritized language. Five of the studies used subjective measures, either hetero-reported measures which gathered the views of parents (3) or school administrators (1), or self-reported measures which gathered the views of students (1). One study used objective measures. In terms of findings, four studies reported positive outcomes, one negative and one neutral. In the studies eliciting

parents' perspectives (Cross and Gearon, 2013; Davis et al., 2019; Ní Dhiorbháin et al., 2023), Ní Dhiorbháin et al. (2023) used semi-structured interviews to examine the perceptions of Irish-medium primary education held by parents ($N=15$) who spoke a language at home other than Irish or English. Cross and Gearon (2013) used questionnaires to evaluate parents' ($N=51$) perceptions of the effects of CLIL before and after a trial in Australia. Davis et al. (2019) used questionnaires and interviews to gauge parents' ($N=23$) perspectives of the appropriateness of French immersion programs for "allophone" primary students who spoke a language other than French or English at home. As for school leaders, Bower (2020) used semi-structured interviews to explore their views ($N=12$) on several issues, including cognitive and metacognitive outcomes, in secondary schools teaching through French in England. In terms of findings, these stakeholders' views on cognitive and metacognitive outcomes in CLIL in LOTE were mixed. On the positive side, Ní Dhiorbháin et al. (2023) showed that one of the main reasons multilingual parents chose to register their children in Irish-medium primary schools was their awareness of the "cognitive and metacognitive skills that arise from the learning of additional languages" (p. 10), even if the language learnt would rarely be used outside of school for most learners. Similarly, in the study by Davis et al. (2019), some "allophone" parents believed French immersion programs offered an additional academic challenge which made the program suitable for their children, with one parent referencing the cognitive flexibility derived from language switching. In contrast, Cross and Gearon (2013) identified a decline in parents' perceptions of the benefits and relevance of CLIL and languages, including CLIL's impact on intellectual development, after their children had received CLIL instruction. Noteworthy is that some groups had more favorable views than others about CLIL's contributions to intellectual development and problem-solving skills, including secondary school parents over primary school parents and, perhaps surprisingly, monolingual parents over multilingual parents. Finally, Bower (2020) found that school leaders reported greater cognitive challenge across CLIL subjects, which could impact learners in different ways according to their academic profiles, as we noted above.

Regarding students' perspectives, Coyle (2013) used questionnaires and discussions with CLIL secondary students from England and Scotland ($N=670$) to examine several issues, including cognitive and metacognitive ones. She found that CLIL learners mentioned higher levels of cognitive challenge when learning through other languages and felt that they were rising to this challenge. As for students' performance on objective measures, Cape et al. (2021) used three tasks of attentional control with primary year 5 students in Scotland to compare the executive function of Gaelic-medium ($N=29$) and English-medium ($N=30$) students. They found that the Gaelic-medium students significantly outperformed their English-medium peers on one of the three tasks, but no significant differences were observed on the other two. The authors conclude that the context of the bilingual experience is important in shaping the cognitive effects of exposure to more than one language.

4 Discussion

This paper has presented a comprehensive review of 29 studies addressing student learning outcomes in CLIL in LOTE, according to

assessment tasks (objective measures) and stakeholder perspectives (subjective measures). Within these studies, 47 learning outcomes were identified, and a careful analysis led to their categorization into five types. As mentioned above, most studies (26 of 29) addressed academic and/or linguistic outcomes, with some studies additionally addressing social/cultural outcomes, behavioral/affective outcomes, and/or (meta) cognitive outcomes. Of the 47 learning outcomes reported, 25 (53%) were positive, five (11%) were negative, four (9%) were neutral, eight (17%) were mixed and four (9%) identified factors influencing outcomes. After excluding studies on factors influencing outcomes, we can see that the percentage of positive outcomes varies between categories, listed here from smallest to greatest percentage of positive outcomes: academic outcomes (44% positive), cognitive and metacognitive outcomes (50% positive), linguistic outcomes (60% positive), behavioral and affective outcomes (67% positive), social and cultural outcomes (71% positive). This section discusses the implications of the findings within each category in turn.

The findings on academic outcomes in CLIL in LOTE were highly context dependent, varying according to program type, geographical location and the typical profile of enrolled students. In the reviewed studies, majority language speakers seemed to fare better overall than heritage speakers of minoritized languages, but variation was observed within the latter group. In CLIL programs serving primarily majority language students, CLIL students either outperformed (Bower, 2020; Surmont et al., 2016) or performed on par with (O'Hanlon et al., 2010; Wright and Scullion, 2007) their non-CLIL peers, which suggests that these CLIL students may gain additional exposure to the target language with no harm to or at the benefit of their academic achievement. However, to our knowledge only one of these studies (O'Hanlon et al., 2010) matched students on potentially confounding variables, such as previous academic achievement, motivation and socioeconomic status, so it is unclear whether the perceived gains can be attributed to the benefits of CLIL or to the profile of student who tends to enroll in CLIL in these contexts (as mentioned in 2.3, this is the case for most of the reviewed studies across categories). In the immersion programs serving primarily heritage speakers of the target language, academic outcomes were mixed. In the USA, immersion programs worked to the benefit of both English-speaking and Spanish-speaking students, who outperformed their mainstream peers (Collier and Thomas, 2004; Marian et al., 2013), but in New Zealand, Maori-immersion students either performed on par with or underperformed their mainstream Maori peers (Murray, 2007; Stewart, 2011). We argue that further research should pinpoint which specific aspects of immersion programs contribute to or detract from minority-language students' success. As a starting point, our analysis suggests that mixed groups of minority- and majority-language speakers being taught through both languages (as in the USA) may produce better results than programs in which minority-language students are separated from majority-language students (as in New Zealand), but this is purely speculative. An additional consideration which is applicable across contexts is that the standardized testing carried out in the majority language may negatively impact the results of minority language speakers and the decisions these results inform (Murtagh and Seoighe, 2022).

The findings on linguistic outcomes suggest that CLIL in LOTE contributes positively to students' development of linguistic proficiency in the target language overall, and they also help us identify areas that may need more attention and factors potentially

influencing students' achievement. Studies using objective measures found that CLIL students attained higher levels of proficiency (Pérez et al., 2016) and wrote more syntactically and lexically complex texts (Bulon et al., 2017) in the target language than their non-CLIL peers, and that their patterns of linguistic development were in line with those of fully bilingual students (Hermanto et al., 2012) and in line with those in their other CLIL language: English (Baten et al., 2020). Similarly, teachers reported positive perceptions of students' overall linguistic development in CLIL in LOTE (Bower, 2020; Hunt, 2011; Ozfidan, 2014), though they raised concerns about its suitability for speakers of neither the mainstream nor the target language, who they believed may be at a disadvantage (Davis et al., 2019). For their part, students claimed to be more confident speaking the target language (Coyle, 2013; Cross and Gearon, 2013; Macleod, 2014) than reading or writing it (Coyle, 2013; Macleod, 2014; Rehamo and Harrell, 2020), which they viewed as more challenging. We contend that students' struggles in reading and writing may be related to the fact that interactive, oral activities are often foregrounded in CLIL, sometimes at the expense of written ones (Pérez Cañado, 2018a). Greater attention to disciplinary literacies in CLIL in LOTE may help students feel more confident reading and writing the academic language of the school disciplines and improve both linguistic and academic outcomes. Other factors which may positively influence students' linguistic outcomes, according to the reviewed studies, are parents' positive attitudes toward the target language (Gartziarena et al., 2024), students' investments in being perceived positively by parents and teachers (Potowski, 2004), and engagement with homework tasks (Birnle, 2022). In light of the key role played by parents in their childrens' linguistic outcomes, we advocate for regular communication between CLIL schools and parents to help promote positive attitudes about target language use in the home, as well as bilingual homework materials which empower parents to support their childrens' regular engagement with these tasks even when they do not speak the language of instruction.

In terms of social and cultural outcomes in CLIL in LOTE, stakeholders' perceptions were mostly positive, but those of parents, teachers and school leaders were more positive than those of the students themselves. Interviewed educators made strong claims about their program's positive effects on learners' intercultural awareness (Bower, 2020) and promotion of biculturalism and self-esteem around the heritage culture in bilingual communities (Ozfidan, 2014) which were echoed by parents (Ní Dhiorbháin et al., 2023; O'Hanlon et al., 2010). But students were more cautious: some acknowledged a link between language and culture (Bower, 2019), while others reduced their expectations about CLIL's contributions to intercultural awareness after receiving CLIL instruction (Cross and Gearon, 2013), and still others questioned the utility of their particular language of instruction beyond education (Macleod, 2014). In light of these findings, we believe that any CLIL program seeking to promote intercultural awareness or preserve heritage cultures should clearly delineate its aims in this regard and its plans for reaching them, beyond simply using the target language in question, to better ensure that the social and cultural value of CLIL in LOTE is perceived by its students.

In studies addressing behavioral and affective outcomes in CLIL in LOTE, stakeholders reported that there were observable changes in performance and behavior as well as in perceptions and attitudes. With reference to behavioral outcomes, teachers and administrators

in two studies (Méndez García and Pavón Vázquez, 2012; Hunt, 2011) reported that CLIL learners tended to work with greater effectiveness and efficiency, which they attribute to students' efforts to master both the target language and content, noting their concentration and focus when carrying out the ongoing tasks. As for students, they attribute the improvements to their teachers' use of group discussions and collaborative learning activities, which facilitate active engagement and interaction. As such, there may be internal and external determinants/factors related to student learning outcomes in CLIL in LOTE, such as students' readiness and interest to learn or teachers' classroom strategies. Further studies taking these factors into account will need to be carried out in order to confirm their association with behavioral outcomes in CLIL in LOTE during CLIL learners' involvement in various tasks or activities.

With reference to affective outcomes, both teachers and students reported positive changes in students' attitudes toward CLIL. A possible explanation for this positive attitude change might be the observed increases in learners' motivation (Bower, 2019; Méndez García and Pavón Vázquez, 2012), in their confidence in learning through the foreign language, and in their beliefs about the effectiveness of CLIL (Bower, 2020; Cross and Gearon, 2013; Hunt, 2011). However, these gains do not seem to be universal: students who were less interested in the content or the language or who struggled academically were less motivated and less receptive to CLIL (Bower, 2019; Cross and Gearon, 2013). In contrast, students who were academically advanced, had higher levels of self-esteem or attended schools of higher socioeconomic status showed more positive attitudes toward CLIL (Bower, 2019; Cross and Gearon, 2013). This is in line with what Keith Stanovich (1986, pp. 80–83) described as the Matthew Effect in Reading 40 years ago, whereby “the rich get richer, and the poor get poorer.” It can therefore be assumed that academically advanced students have greater confidence and this in turn contributes to higher levels of motivation, while students with learning difficulties might be at greater risk of developing negative attitudes toward CLIL and of experiencing challenges with CLIL instruction considering that they need extra support to process information during instruction. Taken together, additional research is needed to determine the possible factors related to students' positive vs. less positive attitudes toward CLIL.

In most studies addressing cognitive and metacognitive outcomes, teachers, parents and students reported positive perceptions regarding CLIL's effects on students' intellectual development. For example, many stakeholders referred to the cognitive benefits of learning through more than one language (Bower, 2020; Ní Dhíorbháin et al., 2023), which requires a certain degree of cognitive flexibility (Davis et al., 2019). This flexibility can be beneficial for many learners, but it also supposes an additional academic challenge that they may not be equally prepared to meet. For instance, students with learning difficulties have found it demotivating (Bower, 2020), and primary school parents were less positive about its benefits than secondary school parents (Cross and Gearon, 2013). It may be the case that, in order for at-risk learners and younger learners to reap the cognitive benefits of CLIL, additional accommodations must be made. For example, the target language may need to be introduced more gradually or only in less cognitively demanding activities. What we might infer from these findings is that the cognitive benefits gained through exposure to more than one language are more challenging for students with slower pace of learning or at-risk learners. Taken together, it seems that further research is required to identify potential

factors related to students' motivation toward the cognitive demands of CLIL.

Finally, this review makes practical educational contributions. The results of the reviewed studies reveal a number of stakeholder suggestions for improving the outcomes reported in the previous sections, which we believe are fundamental to progress in the field. One major challenge to CLIL in LOTE is insufficient teacher training. In Méndez García and Pavón Vázquez's (2012) study, French CLIL teachers in Andalusia, Spain reported training deficits which led content teachers to downplay language matters and their assistants to fail to integrate content and language. The authors maintain that CLIL teaching practice could be enhanced by providing clear methodological training and by better coordinating the roles and responsibilities of co-teachers. Teachers in the Basque Country of Spain also emphasized the importance of both pre-service and in-service teacher training programs specific to bilingual education (Ozfidan, 2014). Similarly, Rehamo and Harrell (2020) report a shortage of qualified Nuosu bilingual teachers in China. The second challenge is concerned with difficulties in accessing appropriate CLIL materials and resources. In the Spanish context, Ozfidan (2014) highlights the importance of government funding for teaching materials and resources that will be effective in captivating students' interest in CLIL, including audio-visual resources and software. In the Irish context, Murtagh and Seoighe (2022) point to a need for Irish-language teaching resources and materials to reduce the translation burden on Irish-medium teachers. They also note a need for standardized tests in Irish, in order to test Irish-medium students in their primary school language and to better inform the educational decisions taken on the basis of these tests. In the Canadian context, French-immersion educators report that students speaking languages other than French or English at home need more formal linguistic instruction in CLIL classrooms and in their English as an Additional Language classes in order to be able to participate effectively in this program (Davis et al., 2019).

4.1 Limitations and future research

The main limitation of the present review is that the state of the art relies heavily on case and cross-sectional designs, which account for 24 of the 29 studies included. While such designs offer avenues for describing the current state of affairs, generating hypotheses and exploring data to establish correlations between variables (Cohen et al., 2017), conclusions about the effects of CLIL in LOTE on students' learning outcomes cannot be reached without further longitudinal and (quasi-)experimental studies. Both designs are essential because neither support causal inferences without the contribution of the other (Bradley and Bryant, 1983; Pittas and Nunes, 2014): (quasi-)experimental studies which match students for potentially confounding variables can reveal the effects of CLIL exposure on relevant independent variables, while longitudinal studies can reveal how those effects develop over time. They are thus instrumental in confirming associations between variables and providing evidence for how early performance on educational tests is associated with later success on the same tests.

Although we recognize the difficulty of controlling variables in an educational environment, we maintain that causal links between

CLIL and students' learning outcomes cannot be established without longitudinal and experimental studies, especially Randomized Control Trials in which participants are randomly allocated into experimental and comparison groups (Robson, 2024). A synthesis of the findings of these types of studies would permit evidence-based responses to the question of "what works," which will advance both research in the field and the teaching practice which it intends to inform. At present, we have identified only four studies employing a longitudinal design to investigate students' learning outcomes in CLIL in LOTE, and only one study employing a quasi-experimental design. Therefore, further research is needed to test the causal connection between CLIL instruction and learning outcomes if we wish to paint an accurate picture of student learning outcomes in CLIL in LOTE.

Additionally, only a few studies reviewed here were designed to identify the factors linked with student learning outcomes in CLIL in LOTE. As such, further research is necessary to establish the possible factors related to students' learning outcomes in all categories, and to students' positive vs. less positive attitudes toward CLIL. It may also be feasible to explore whether the factors which stakeholders speculate about, such as students' enjoyment while learning through the foreign language, their efforts to master both the target language and content, and their increased efforts to focus and concentrate, in fact contribute to student learning outcomes in CLIL in LOTE. In addition to this, the findings of this review highlight that CLIL in LOTE supposes additional cognitive demands for students, and some students may need further support in meeting these demands, such as at-risk learners and younger learners. Further research must thus address both how to support these students and how to improve their attitudes toward CLIL in LOTE.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

References

- Admiraal, W., Westhoff, G., and De Bot, K. (2006). Evaluation of bilingual secondary education in the Netherlands: students' language proficiency in English. *Educ. Res. Eval.* 12, 75–93. doi: 10.1080/13803610500392160
- Anghel, B., Cabrales, A., and Carro, J. M. (2016). Evaluating a bilingual education program in Spain: the impact beyond foreign language learning. *Econ. Inq.* 54, 1202–1223. doi: 10.1111/eicn.12305
- Baten, K., Hiel, S. V., and Cuypere, L. D. (2020). Vocabulary development in a CLIL context: a comparison between French and English L2. *Stud. Second Lang. Learn. Teach.* 10, 307–336. doi: 10.14746/ssllt.2020.10.2.5
- Bermingham, N. (2021). Countering decapitalisation: examining teachers' discourses of migration in Galicia. *Lang. Cult. Curric.* 34, 337–351. doi: 10.1080/07908318.2021.1874965
- Birnie, I. (2022). Blended learning to support minority language Acquisition in Primary School Pupils: lessons from the 'taking Gaelic home study'. *Aust. Int. J. Rural. Educ.* 32, 126–141. doi: 10.47381/aijre.v32i2.329
- Bloom, B. S. (1956). *Taxonomy of educational objectives, handbook: The cognitive domain*. New York: David McKay.
- Bower, K. (2019). 'Speaking French alive': learner perspectives on their motivation in content and language integrated learning in England. *Innov. Lang. Learn. Teach.* 13, 45–60. doi: 10.1080/17501229.2017.1314483
- Bower, K. (2020). School leaders' perspectives on content and language integrated learning in England. *Lang. Cult. Curric.* 33, 351–367. doi: 10.1080/07908318.2019.1667367
- Bradley, L., and Bryant, P. (1983). Categorizing sounds and learning to read—a causal connection. *Nature* 301, 419–421. doi: 10.1038/301419a0
- Brevik, M., and Moe, E. (2012). "Effects of CLIL teaching on language outcomes" in *Collaboration in language testing and assessment*. eds. D. Tsagari and I. Csépes (Berlin, Germany: Peter Lang), 213–227.
- Broca, Á. (2016). CLIL and non-CLIL: differences from the outset. *ELT J.* 70, 320–331. doi: 10.1093/elt/ccw011
- Bruton, A. (2011). Is CLIL so beneficial, or just selective? Re-evaluating some of the research. *System* 39, 523–532. doi: 10.1016/j.system.2011.08.002
- Bulon, A., Hendriks, I., Meunier, F., and Van Goethem, K. (2017). Using global complexity measures to assess second language proficiency: comparing CLIL and non-CLIL learners of English and Dutch in French-speaking Belgium. *Papers of the Linguistic Society of Belgium* 11, 1. <http://hdl.handle.net/2078.1/184992>.
- Cape, R., Vega-Mendoza, M., Bak, T. H., and Sorace, A. (2021). Cognitive effects of Gaelic medium education on primary school children in Scotland. *Int. J. Biling. Educ. Biling.* 24, 1065–1084. doi: 10.1080/13670050.2018.1543648
- Collier, V. P., and Thomas, W. P. (2004). The astounding effectiveness of dual language education for all. *NABE J. Res. Pract.* 2, 15–20.
- Cohen, L., Manion, L., and Morrison, K. (2017). *Research methods in education*. 8th Edn. London: Routledge.
- Coyle, D. (2013). Listening to learners: an investigation into 'successful learning' across CLIL contexts. *Int. J. Biling. Educ. Biling.* 16, 244–266. doi: 10.1080/13670050.2013.777384

Author contributions

EP: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. LT: Data curation, Formal analysis, Investigation, Methodology, Resources, Validation, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/feduc.2024.1447270/full#supplementary-material>

- Cross, R., and Gearon, M. (2013). Research and evaluation of the content and language integrated learning (CLIL) approach to teaching and learning languages in Victorian schools. Available at: <https://minerva-access.unimelb.edu.au/handle/11343/55778> Accessed on February 14, 2024.
- Dallinger, S., Jonkmann, K., Hollm, J., and Fiege, C. (2016). The effect of content and language integrated learning on students' English and history competences – killing two birds with one stone? *Learn. Instr.* 41, 23–31. doi: 10.1016/j.learninstruc.2015.09.003
- Dalton-Puffer, C. (2011). Content-and-language integrated learning: from practice to principles? *Annu. Rev. Appl. Linguist.* 31, 182–204. doi: 10.1017/S0267190511000092
- Davis, S., Ballinger, S., and Sarkar, M. (2019). The suitability of French immersion for allophone students in Saskatchewan: exploring diverse perspectives on language learning and inclusion. *Can. J. App. Ling.* 22, 27–63. doi: 10.7202/1063773ar
- Fernández Sanjurjo, J., Arias Blanco, J. M., and Fernández-Costales, A. (2018). Assessing the influence of socio-economic status on students' performance in content and language integrated learning. *System* 73, 16–26. doi: 10.1016/j.system.2017.09.001
- Gartziarena, M., Villabona, N., and Olave, B. (2024). In-service teachers' multilingual language teaching and learning approaches: insights from the Basque Country. *Lang. Educ.* 38, 203–217. doi: 10.1080/09500782.2023.2176714
- Gorard, S. (2013). Research design: Creating robust approaches for the social sciences. Sage.
- Harris, V. (2000). Teaching learners how to learn: Strategy training in the ML classroom. London: CILT, Center for Information on Language Teaching and Research.
- Hermanto, N., Moreno, S., and Bialystok, E. (2012). Linguistic and metalinguistic outcomes of intense immersion education: how bilingual? *Int. J. Biling. Educ. Biling.* 15, 131–145. doi: 10.1080/13670050.2011.652591
- Hunt, M. (2011). UK teachers' and learners' experiences of CLIL resulting from the EU-funded project ECLILT. *Lat. Am. J. Cont. Lang. Int. Learn.* 4, 27–39. doi: 10.5294/lacilil.2011.4.1.3
- Hüttner, J., and Smit, U. (2023). "Policy, practice and agency: making CLIL work? Insights from Austrian upper secondary technical education" in Global CLIL: Critical, ethnographic and language policy perspectives. ed. E. Codó (New York, NY: Routledge), 125–149.
- 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. *Lang. Educ.* 19, 147–168. doi: 10.1080/09500780508668671
- Jexenflcker, S., and Dalton-Puffer, C. (2010). "The CLIL differential: comparing the writing of CLIL and non-CLIL students in higher colleges of technology" in Language use and language learning in CLIL classrooms. eds. C. Dalton-Puffer, T. Nikula and U. Smit (Amsterdam, the Netherlands: John Benjamins), 169–190.
- La Valle, C., Johnston, E., and Tager-Flusberg, H. (2022). A systematic review of the use of telehealth to facilitate a diagnosis for children with developmental concerns. *Res. Dev. Disabil.* 127:104269. doi: 10.1016/j.ridd.2022.104269
- Lasagabaster, D. (2008). Foreign language competence in content and language integrated courses. *Open App. Ling. J.* 1, 30–41. doi: 10.2174/1874913500801010030
- Livingston, J. (2003) Metacognition: An overview. Available at: <https://eric.ed.gov/?id=ED474273>. Accessed on April 12, 2024.
- Llinares, A., and Dafouz, E. (2010). "Content and language integrated programmes in the Madrid region: overview and research findings," in CLIL in Spain: Implementation, results and teacher training, eds. D. Lasagabaster and Zarobe Y. Ruiz de (Newcastle, UK: Cambridge Scholars).
- Lo, A. W. T. (2024). Unlocking CLIL success: exploring the interplay between students' self-regulation levels, linguistic challenges and learning outcomes in Hong Kong secondary education. *Lang. Educ.* 1–19, 1–19. doi: 10.1080/09500782.2024.2314135
- Lo, Y. Y., and Murphy, V. A. (2010). Vocabulary knowledge and growth in immersion and regular language-learning programmes in Hong Kong. *Lang. Educ.* 24, 215–238. doi: 10.1080/09500780903576125
- Lorenzo, F., Granados, A., and Rico, N. (2021). Equity in bilingual education: socioeconomic status and content and language integrated learning in monolingual southern Europe. *Appl. Linguist.* 42, 393–413. doi: 10.1093/applin/amaa037
- Macleod, M. (2014). Young speakers' use of Gaelic in the primary classroom: A multi-perspectival pilot study, Scotland: University of Aberdeen.
- Mady, C. (2015). Immigrants outperform Canadian-born groups in French immersion: examining factors that influence their achievement. *Int. J. Multiling.* 12, 298–311. doi: 10.1080/14790718.2014.967252
- Manterola, I., Almgren, M., and Idiazabal, I. (2013). Basque L2 development in immersion school settings. *Int. J. Biling.* 17, 375–391. doi: 10.1177/1367006912438996
- Marian, V., Shook, A., and Schroeder, S. R. (2013). Bilingual two-way immersion programs benefit academic achievement. *Biling. Res. J.* 36, 167–186. doi: 10.1080/15235882.2013.818075
- Mediavilla, M., Mancebón, M. J., Gómez-Sancho, J. M., and Pires, L. (2019). Bilingual education and school choice: a case study of public secondary schools in the Spanish region of Madrid. Available at: http://diposit.ub.edu/dspace/bitstream/2445/134081/1/IEB19-01_Mediavilla%20et.al.pdf Accessed on June 5, 2024.
- Méndez García, M. D. C., and Pavón Vázquez, V. (2012). Investigating the coexistence of the mother tongue and the foreign language through teacher collaboration in CLIL contexts: perceptions and practice of the teachers involved in the plurilingual programme in Andalusia. *Int. J. Biling. Educ. Biling.* 15, 573–592. doi: 10.1080/13670050.2012.670195
- Murray, S. (2007). Achievement at Maori immersion and bilingual schools. Update for 2005 results. Available at: https://www.educationcounts.govt.nz/publications/schooling/11240/maori_immersion_2005 Accessed on February 15, 2024.
- Murtagh, L., and Seoighe, A. (2022). Educational psychological provision in Irish-medium primary schools in indigenous Irish language speaking communities (Gaeltacht): views of teachers and educational psychologists. *Br. J. Educ. Psychol.* 92, 1278–1294. doi: 10.1111/bjep.12499
- National Heart, Lung, and Blood Institute (NHLBI) Quality Assessment tools. (2023). Study quality assessment tools. <https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools> Accessed on February 5, 2024.
- Ní Dhorbáin, A., Nic Aindriú, S., Connaughton-Crean, L., and Duibhir, P. Ó. (2023). It's more the invisible benefits – multilingual parents' experiences of immersion education and their reasons for choosing immersion. *Lang. Educ.* 1–20, 1–20. doi: 10.1080/09500782.2023.2238680
- O'Hanlon, F., McLeod, W., and Paterson, L. (2010). Gaelic-medium education in Scotland: choice and attainment at the primary and early secondary school stages Bòrd na Gàidhlig. Available at: <http://www.gaidhlig.org.uk/Downloads/Aithisg%20202010%20Report%20-%20GME%20Choice%20and%20Attainment.pdf> Accessed on February 15, 2024.
- Ozfidan, B. (2014). The Basque bilingual education system: a model for a Kurdish bilingual education system in Turkey. *J. Lang. Teach. Res.* 5, 382–390. doi: 10.4304/jltr.5.2.382-390
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Br. Med. J.* 372:n71. doi: 10.1136/bmj.n71
- Pavón Vázquez, V. (2018). Learning outcomes in CLIL programmes: a comparison of results between urban and rural environments. *Porta Linguarum* 29, 9–28. doi: 10.30827/Digibug.54020
- Pérez, A., Lorenzo, F., and Pavón, V. (2016). European bilingual models beyond lingua franca: key findings from CLIL French programs. *Lang. Policy* 15, 485–504. doi: 10.1007/s10993-015-9386-7
- Pérez Cañado, M. L. (2018a). CLIL and pedagogical innovation: fact or fiction? *Int. J. Appl. Linguist.* 28, 369–390. doi: 10.1111/ijal.12208
- Pérez Cañado, M. L. (2018b). The effects of CLIL on L1 and content learning: updated empirical evidence from monolingual contexts. *Learn. Instr.* 57, 18–33. doi: 10.1016/j.learninstruc.2017.12.002
- Pérez-Cañado, M. L. (2012). CLIL research in Europe: past, present, and future. *Int. J. Biling. Educ. Biling.* 15, 315–341. doi: 10.1080/13670050.2011.630064
- Pittas, E., and Nunes, T. (2014). The relation between morphological awareness and reading and spelling in Greek: a longitudinal study. *Read. Writ.* 27, 1507–1527. doi: 10.1007/s11145-014-9503-6
- Pittas, E., and Papanastasiou, E. (2023). Effects of COVID-19 on the educational performance of children with special educational needs and disabilities: a systematic review according to children's/youth's and caregivers' perspectives. *Res. Dev. Disabil.* 143:104635. doi: 10.1016/j.ridd.2023.104635
- Pladevall-Ballester, E., and Vallbona, A. (2016). CLIL in minimal input contexts: a longitudinal study of primary school learners' receptive skills. *System* 58, 37–48. doi: 10.1016/j.system.2016.02.009
- Potowski, K. (2004). Student Spanish use and Investment in a Dual Immersion Classroom: implications for second language acquisition and heritage language maintenance. *Mod. Lang. J.* 88, 75–101. doi: 10.1111/j.0026-7902.2004.00219.x
- Rehama, A., and Harrell, S. (2020). Theory and practice of bilingual education in China: lessons from Liangshan Yi autonomous prefecture. *Int. J. Biling. Educ. Biling.* 23, 1254–1269. doi: 10.1080/13670050.2018.1441259
- Riera Ventura, L. (2021). Effects of CLIL on the students' motivation and learning of English as a foreign language, Catalunya, Spain: Universitat de Vic.
- Robson, C. (2024). Real world research. United Kingdom: Blackwell.
- Ruiz de Zarobe, Y. (2008). CLIL and foreign language learning: a longitudinal study in the Basque country. *Int. CLIL Res. J.* 1, 60–73.
- Ruiz de Zarobe, Y. (2010). "Written production and CLIL: an empirical study" in Language use and language learning in CLIL classrooms. eds. C. Dalton-Puffer, T. Nikula and U. Smit (Amsterdam, the Netherlands: John Benjamins), 191–212.
- Somers, T., and Llinares, A. (2021). Students' motivation for content and language integrated learning and the role of programme intensity. *Int. J. Biling. Educ. Biling.* 24, 839–854. doi: 10.1080/13670050.2018.1517722
- Stanovich, K. E. (1986). Matthew effects in Reading: some consequences of individual differences in the acquisition of literacy. *Read. Res. Q.* 21, 360–407. doi: 10.1598/RRQ.21.4.1
- Stewart, G. (2011). Science in the Māori-medium curriculum: assessment of policy outcomes in Pūtaiao education. *Educ. Philos. Theory* 43, 724–741. doi: 10.1111/j.1469-5812.2009.00557.x

Surmont, J., Struys, E., Van Den Noort, M., and Van De Craen, P. (2016). The effects of CLIL on mathematical content learning: a longitudinal study. *Stud. Second Lang. Learn. Teach.* 6, 319–337. doi: 10.14746/ssllt.2016.6.2.7

Van Mensel, L., Hiligsmann, P., Mettwie, L., and Galand, B. (2020). CLIL, an elitist language learning approach? A background analysis of English and Dutch CLIL pupils in French-speaking Belgium. *Lang. Cult. Curric.* 33, 1–14. doi: 10.1080/07908318.2019.1571078

Wei, X., Saab, N., and Admiraal, W. (2021). Assessment of cognitive, behavioral, and affective learning outcomes in massive open online courses: a systematic literature review. *Comput. Educ.* 163:104097. doi: 10.1016/j.compedu.2020.104097

Wright, M., and Scullion, P. (2007). Quality of school life and attitudes to Irish in the Irish-medium and English-medium primary school. *Ir. Educ. Stud.* 26, 57–77. doi: 10.1080/03323310601125278