

The effect of CLIL combined with language instruction on language learning and the role of individual and institutional factors in students' perspectives: Empirical evidence from Kazakhstan

Malik Satayev^{1,2} , Elvira Barrios³ , Alberto Fernandez-Costales⁴ , Shakhriyat Agaidarova² ,
Rimma Izbassarova¹ , Nuri Balta^{2*} 

¹ Kazakh National Pedagogical University, Almaty, KAZAKHSTAN

² Suleyman Demirel University, Kaskelen, KAZAKHSTAN

³ University of Malaga, Malaga, SPAIN

⁴ University of Oviedo, Oviedo, SPAIN

Received 18 May 2022 • Accepted 23 August 2022

Abstract

Content and language integrated learning (CLIL) is still a growing area of research. In this study, we investigated the effect of a one-semester CLIL intervention combined with English language instruction on university students' general English skills the associations between institutional and individual factors and views on CLIL. A total of 105 students from two academic programs (biology-chemistry, and physic-informatics) at a private university in Kazakhstan participated in the study. Quantitative data were collected through two different instruments. MANOVA and correlation analyses were carried out to analyze the data. A large effect of CLIL combined with language instruction was observed on language gains. Additionally, the study identified two factors that are significantly associated with student recognition of higher language and disciplinary knowledge learning, more positive attitudes to and higher satisfaction with CLIL: being a non-first-year student and having CLIL experience in pre-university education. In light of the results obtained, our findings reinforce the case for integrating language instruction that includes general, academic, and subject-specific goals in English-medium programs and courses in higher education.

Keywords: CLIL, higher education, Kazakhstan, English learning, students' perspectives

INTRODUCTION

The effects of CLIL on the learning process have attracted scholarly attention—especially in primary and secondary education—, and there is a relevant corpus of research analyzing the impact of studying through English on students' language competence in the second language (L2) (Lasagabaster & Ruiz de Zarobe, 2010; Pérez-Cañado, 2018) and the satisfaction of students with CLIL programs (Lasagabaster & Doiz, 2016). Similarly, the implementation of English-taught programs in higher education—under English-Medium Instruction or EMI—has also been examined in the last decade, especially concerning the implications of teaching through English, the relevance of teacher

training, and the need for suitable instructional and methodological support to maximize student performance (Doiz & Lasagabaster, 2021; Fernández-Costales, 2022; Macaro et al., 2018).

Notwithstanding the progress in this line of enquiry, research has been primarily focused on the results of bilingual programs in the European scenario, and further examination on the implementation of English-taught programs in other settings is required, as it has only been recently approached. Moreover, the central issue in English-taught programs is still to evaluate the improvement in students' L2 competence and analyze if no negative effects are observed regarding content learning. As Barrios and López-Gutiérrez (2021, p. 2) argue, "the questions remain as to whether students

Contribution to the literature

- The article analyses the impact that studying through English has on students' competence in the L2 and on their CLIL-related perspectives.
- The current paper contributes to CLIL research by offering empirical evidence on the advantages of English-taught programs that integrate explicit language goals in tertiary education.
- The study contributes to the optimization of CLIL provisions in higher education in Kazakhstan (and other multilingual settings beyond the European Union).

make any significant progress in their language proficiency and to whether they acquire the purported benefit of the added value of EMI in terms of English language learning". In this regard, Macaro et al. (2018) underline the need for further research evaluating the impact of EMI on students' English proficiency. In particular, there is a dearth of studies using objective tests instead of learners' self-reports to measure language gains. In addition, as Macaro and Akincioglu (2018) claimed, individual and institutional factors deserve further attention, as relevant variables such as gender, student previous experience in bilingual programs (CLIL or EMI), and the academic program of participants has not been sufficiently examined in scientific literature.

In this sense, we believe that the current paper might contribute to the field of CLIL and EMI in higher education (HE) by providing empirical evidence on the impact that studying through English has on students' competence in the L2 and on their CLIL-related perspectives. To this aim, the objective of the present study is two-fold:

1. to investigate the effect of a one-semester CLIL intervention that included English language instruction on university students' general English skills, and
2. to assess how institutional and individual factors are associated with perspectives on CLIL.

More precisely, the following research questions (RQs) are addressed:

1. **RQ1:** Does a one-semester CLIL intervention in higher education combined with general and discipline-specific English language instruction have a significant effect on university students' grammar and listening?
2. **RQ2:** Are there statistically significant differences in student perspectives on the CLIL intervention associated with students' gender, academic program, year group, and having lessons with the CLIL approach previously?
3. **RQ3:** Are there statistically significant relationships between the students' proficiency level and their perspectives on CLIL?

Although Kazakhstan is a multilingual setting with more than 130 languages (Zharkynbekova, 2017), it has been overlooked in the international panorama, and

there is a paucity of studies reporting on empirical evidence on English-taught programs in this context. By investigating language gains and the perception of students learning content through English, we may contribute to optimizing the implementation of CLIL provisions in higher education in Kazakhstan (and other multilingual settings beyond the European Union). Furthermore, our study will offer empirical evidence to gauge the relevant advantages of English-taught programs that integrate explicit language goals in tertiary education.

LITERATURE REVIEW

Language Gains and Student Perceptions in English-Medium Education

When analyzing the implementation of English-taught programs, research has been devoted to measuring the implications of using English as a medium of instruction (Galloway & Rose, 2021), the promotion of multilingualism in higher education (Lasagabaster, 2021), teacher satisfaction with EMI/CLIL programs (Fernández-Costales & González-Riaño, 2015), teacher training (O'Dowd, 2015), or the meta discourse used by university lecturers (Doiz & Lasagabaster, 2022). Undoubtedly, the enhancement of language competence in the L2 and student perceptions of the learning process are two major concerns for universities and educators. Although research generally reports positive outcomes in terms of students' performance, the heterogeneity and plurality of English-taught programs do not allow to draw straightforward conclusions (Doiz et al., 2013). The results of available studies suggest that the variables, the specific conditions, and the context of the studies need to be carefully taken into consideration. Also, most studies rely on participants' self-perceived language level, and very few investigations have used standardized tests to examine students' competence in the L2. Likewise, individual variables such as gender have been largely overlooked in CLIL/EMI studies in higher education (Macaro & Akincioglu, 2018).

Starting with the group of studies relying on students' self-reported L2 competence, Muñoz (2001) investigated student perceptions towards the use of English in content subjects in a Spanish university. This study determined that participants' progress was more

noticeable in receptive than in productive skills, with a special focus on the gains in self-confidence. However, as it happens with most studies examining English-taught programs in higher education, the analyzed program did not include language-specific goals and was limited to using English as the language of tuition. Gender differences or other variables, such as students' prior CLIL or EMI experience, were not explored.

In the same vein, Toledo et al. (2011) also conclude that students learning content through English at the university see their receptive skills enhanced. This study examined the attitudes of 39 university students towards English-taught programs and the effect on their motivation and academic output. Data were collected through an ad hoc questionnaire (a Likert scale with 45 items) to survey student perceptions of their learning process. The study concludes that participants improved their written skills in English, especially regarding vocabulary acquisition and written comprehension. Findings also determine that students' attitudes and English competence may affect their perceptions of the bilingual program, although gender differences were not examined.

The study conducted by Ismail et al. (2011) in the Malaysian context aimed to investigate students' inclination toward English as the medium of instruction in the teaching of science and mathematics in higher education. Using a closed survey, the paper investigated the perceptions of 291 undergraduate students and found that participants' inclination towards English as the medium of instruction is positive, with individual differences according to the origin of participants. However, no statistically significant differences were found according to the gender of the students. The authors justify the favorable perception of students towards English by the fact that the participants had studied mathematics and science in the L2 in high school.

Aguilar and Rodríguez (2012) scrutinize the perception of students in a CLIL pilot experience in a Spanish university. The study used an open-ended questionnaire to analyze data of pupils enrolled in an engineering school, and the most salient finding is that most participants (59%) acknowledged they had not learned any English after studying through English for one semester. When asked about the areas where they perceived an improvement in the L2 competence, participants identified the acquisition of technical, specialized vocabulary as the most significant improvement.

Hengsadeeikul et al. (2014) investigated the motivation of students enrolled in English-taught programs in three institutions of higher education in Bangkok. Using a convenience sampling process, the paper screened 2,252 Thai undergraduate students from nine fields: business, education, engineering, English,

information technology, international business, law, nursing, and vocational education. Although the effect sizes were small, statistically significant differences were established according to the degree, with students enrolled in international business and English majors showing a more integrative inclination towards the L2. Results also underline that gender is a determining variable, as statistically significant differences were found, with females being more inclined towards learning the L2.

Fernández-Costales (2017) analyzed the perception of undergraduate students in a Spanish university towards an English-taught program. The study focused on the satisfaction of participants with EMI and its impact on their language competence in English, the promotion of their international dimension, and the improvement in their career prospects. The investigation used a questionnaire to collect data from a sample of 255 Spanish students enrolled in several degrees (accountancy, business and marketing, economy, engineering, and tourism). The most significant finding is that 80% of participants estimated their L2 competence improved since they studied through the medium of English. This tendency is more visible in students with higher levels, as participants with C1 and C2 reported having fewer difficulties following the lessons in a foreign language. Statistically significant differences were found according to the English level, as less proficient students (those with B1 and B2) report noticing less relevant language gains than learners with higher L2 competence.

The study by Hernández-Nanclares and Jiménez-Muñoz (2017) scrutinized written assignments and video recordings of tutorial groups of two first-year modules in a BA program in business administration in a Spanish university. Data were analyzed throughout one academic year, and the language performance of participants was contrasted with descriptors of the Common European Framework of Reference—the CEFR—for the Language (Council of Europe, 2001) to accomplish content-related assignments. The results of the study confirm that students' progress was less than half of a CEFR level in one year. Taking into account that the CEFR estimates that 200 tuition hours are required to progress from one level to the next one (e.g., from B1 to B2), the results of the study are rather positive, although they should be taken with caution, as other determinant factors—such as the time of exposure to the L2 outside the classroom, the fact that students attended private lessons or not, etc.—were not estimated.

In their large-scale study of university students in Turkey, Macaro and Akincioglu (2018) scrutinized the perception of 989 students from 18 universities. Following a quantitative research design, the study utilized a 54-item questionnaire to analyze the motivation of participants to study via EMI. The results of the survey reported significant differences in terms of

year of study and gender, with females being more positive than men regarding their improvement in English. In particular, female students perceived a larger improvement in their general and subject-specific English proficiency than their male counterparts.

As it has already been mentioned, few studies used standardized tests to measure students' competence in the L2. Among these investigations, Rogier (2012) analyzed the language improvement of female Emirati undergraduates over a four-year program. The most salient finding is that participants progressed .05 of a level of the CEFR on average. Results show statistically significant improvements in all areas tested by the IELTS exam, with speaking being the most noticeable gain, followed by reading comprehension, writing, and listening skills. Although this investigation did not include a control group, the study results were compared with the ones reported by studies of general English programs indicating that students need between 200 and 240 tuition hours to improve their CEFR level by .05 (Elder & O'Loughlin, 2003). According to Macaro et al. (2018), the gains shown by this study can hardly be linked with the isolated impact of the program on language learning.

Aguilar and Muñoz (2014) report on a study on engineering students that reveals less proficient pupils obtain higher gains in listening skills and grammar than those with a higher command of English. In contrast, higher proficiency learners did not get any language gains in any of the tests (and they even scored significantly worse in the grammar post-test than in the pre-test). This study employed an international standard language test—the Oxford Placement Test (OPT) 1 and 2 (Allan, 2005a, 2005b)—to measure the changes in participants' L2 competence after studying through English.

In the Chinese context, Lei and Hu (2014) used a general ability language test to determine student proficiency before and after taking part in an EMI program in business administration. Unlike other studies in the field, this research included a control group with students learning content in their L1 (Chinese). Standardized English proficiency tests, focus groups, and motivation surveys were used to collect data from 136 students from the first and second years. The most salient finding of the study is that students learning through English did not outperform the Chinese-medium group after one year. The paper also underlines that prior English competence is the stronger predictor of L2 proficiency after studying in EMI programs.

The study conducted by Heath et al. (2020) in Japan explored the links between L2 proficiency, motivation, and academic language skills in EMI. Analyzing a sample of 146 students of business administration in a Japanese university, Heath et al. (2020) used

questionnaires and standard tests, followed by semi-structured interviews. Participants' subject exam scores and English proficiency tests revealed that L2 competence was a key predictor of the success of the EMI program.

In a recent paper, Barrios and López-Gutiérrez (2021) researched the language development of university students over a four-year partially English-taught program and the learners' perceptions concerning language gains and their experience learning through English. The study used a mixed research design with questionnaires and focus groups. The language competence of 71 students enrolled in a Spanish university was measured through the OPT (1 and 2). The results of the analyses reveal that the least proficient students obtained the highest language gains in the four years, especially concerning listening skills. In contrast, most proficient pupils underperformed in the post-test, with worse results in grammar and listening than in the pre-test. As for the student perception of the EMI program, pupils show a positive attitude towards studying through English, although students with a higher command of the L2 were less enthusiastic about the program. Since the EMI program being analyzed did not have language-specific targets and there was no instructional design intended to support students in acquiring language skills, Barrios and López-Gutiérrez (2021) emphasizes the need for further research on the impact of EMI on language learning, the specific context that would promote an improvement in L2 competence, and—in general—the perception of participants on their learning process on English-taught programs.

The findings of the studies reviewed in this section allow us to identify some conclusions. First, taking into account the surge of attention on EMI and CLIL, the number of studies employing standard tests to assess language gains in students learning content through English at the tertiary level is still scarce. Moreover, the language gains identified in most studies are modest compared to the expectations for language learning in EMI programs. Having said that, the perception of students towards English-taught programs is generally positive, and the overall belief is that studying through English may have a favorable effect on their language competence.

English-Medium Education in Kazakhstan

The case of Kazakhstan is of interest regarding the implementation of English-taught programs, as it has been overlooked so far and there is a dearth of empirical research reporting on language gains and student perceptions. Although the Republic of Kazakhstan has made relevant efforts to promote multilingualism—with Kazakh, Russian and English—in education through the so-called "Trinity of Languages" program (Dearden, 2014; Neuendorf, 2019; Ozdemir, 2018), there is a need

for studies providing empirical results on the use of English as a medium of instruction.

In Kazakhstan, English-taught programs have proliferated in the last years, following the adaptation to the three-year degree model proposed by the Bologna process. Institutions of higher education have made relevant efforts and more than 40 universities offer special groups where English is used as the language of instruction in more than 30% of the courses being offered (Seitzhanova et al., 2015). As it happens with CLIL and EMI programs in other settings, one of the most fundamental characteristics of English-taught itineraries at university level is the lack of a homogeneous model with consistent characteristics, organizational similarities, and pedagogical approaches. Nevertheless, studies analyzing the perception of Kazakh teachers towards the implementation of CLIL (Huertas & Shashken, 2021) confirm there is a lack of support and specific training on this approach. This qualitative study explored the perception of five teachers working in primary and secondary schools in the cities of Pavlodar and Shymkent. The strengths, weaknesses, opportunities and threats (SWOT) analysis reveals the “urgent need of teachers for methodological support and the exchange of experiences on an ongoing basis to improve teaching practice and the development of the teacher’s linguistic skills” (Huertas & Shashken, 2021, p. 268). The study also identifies the scarcity of specific materials as one of the most noticeable challenges for teachers delivering content through an additional language in Kazakhstan.

Karabassova (2019) investigated the implementation of CLIL in 20 state funded elite Nazarbayev Intellectual Schools (NIS). Through semi-structured interviews with five teachers, Karabassova (2019) concludes that CLIL lessons do not significantly differ from traditional first language lessons which focus on teaching content without paying attention to language issues. The main explanation is that teachers do not receive sufficient CLIL principles and methodology training.

Satayev et al. (2022) recently studied if engaging university students in CLIL can increase their achievement in biology and English language. In a quasi-experimental study with a sample of 25 students from Almaty, they found that CLIL effectively taught content and language for science subjects.

Vitchenko (2017) took a mixed research design to analyze the perception of stakeholders on the implementation of CLIL. Using semi-structured interviews with administrators and questionnaires with 207 students and 15 teachers from the Karaganda State Technical University, this study aimed to assess the participants’ beliefs on the role of English as a language of instruction and the possible benefits of using CLIL in Kazakhstan. Although students and lecturers show a positive view on the multilingualism policies promoted

in Kazakhstan, participants are reluctant to implement CLIL. The main reason is that the current conditions in the country and the lack of appropriate support hamper the introduction of fundamental changes in foreign language education. Interestingly, students being surveyed in this project identified low language proficiency as the main challenge for CLIL implementation.

Most studies so far have been devoted to exploring the possibilities of implementing CLIL in Kazakhstan and analyzing teacher and student attitudes and beliefs about this approach. However, no studies have reported empirical results on the effect of CLIL programs on university students’ competence in the L2. The current paper is intended to fill this gap and contribute to the knowledge of the effects of EMI at the tertiary level concerning the students’ learning process (Dewaele, & Saito, 2022).

METHODOLOGY

Research Design

Considering the data collection tools and the data analysis, we employed a quantitative research design in this study. Quantitative research is a type of research that is describing events by gathering numerical data that are evaluated utilizing mathematics-based methods (Creswell, 1994). This study had no goals to make huge assertions by generalizing the outcomes. Thus, this research could be counted as an exploratory case study (Mills et al., 2009; Scott & Morrison, 2006).

Participants

This study was conducted at a medium-sized university in the southeast region of Kazakhstan. This university has approximately 7,300 students, four faculties and over 20-degree programs. The sample includes students from two academic programs (biology-chemistry and physic-informatics) who learned through CLIL for three months at different age groups. In the biology-chemistry or physic-informatics programs, students get two diplomas once they are graduated. instance, students who graduated from the biology-chemistry program can teach chemistry and biology.

Student ages ranged from 17 to 24 years. A total of 105 students participated in this study (77 male, 26 female, while two did not specify). All students were informed about the aims and process of the research and joined voluntarily. A total of 54 of students were from the biology-chemistry program, and 41 were from the physics-informatics program. Ten students did not specify their program. The total number of participants was divided into two groups: first-year students (n=59), and second and above-year students (n=36); ten students did not specify their year group. Finally, 37 students had

prior experience with CLIL in primary and/or secondary education, while 51 participants had not learned content through an additional language before. The students have to undertake general English and English for special purposes (ESP) courses during the first year of their academic program in order to be able to continue their education through EMI. In the first semester, students have general English lessons at the pre-intermediate level for six hours per week, focusing on explicit language acquisition. Course materials included a textbook named *Pathway* (Mitchel & Malkogianni, 2015), which helped develop four language skills and grammar, vocabulary, and pronunciation development. The number of hours of general English language instruction during the second semester of study decreased to three hours each week. The same textbook was used at the next higher level. Additionally, an ESP course that focused on topic terminology and content-specific language was also offered for three hours per week. For instance, the biology-chemistry students followed the ESP courses like introduction to chemistry and terminology and introduction to biology and terminology. Except for Kazakh/Russian languages and history courses, students take all subjects in English all along their education at the university as the university adheres to the EMI approach to be competitive in academia. According to the university curriculum, both academic programs have approximately six courses per semester that are taught using English medium instruction.

Instruments

We used two data collection tools; two forms of a standardized English proficiency test and one questionnaire to determine students' perspectives on CLIL. The tools provided data on students' English language attainment, their perceptions of language and content learning, and their attitudes to and satisfaction with CLIL program implementation.

Oxford placement test

We obtained a measure for English achievement with 41 first-year students as they were new to learning academic content through English. To determine students' language knowledge before and after the CLIL intervention, we used the OPT (Allan, 2005a, 2005b) as they are standardized, reliable and validated tests and contain two equivalent tests that can be used for pre- and post-testing. The OPT results are claimed to be calibrated against the CEFR levels (A1-C2), Cambridge ESOL exam levels, and IELTS band scores. These tests have been used in previous studies on language gains in EMI (Aguilar & Muñoz, 2014; Barrios & López-Gutiérrez, 2021). Comparability across investigations was facilitated by using the same language testing tool. Each test consists of two subtests (a listening section and a grammar section), each of them consisting of 100 items.

The correct choice and the distractors were embedded in the stem of the questions so that test-takers easily provided their responses. A typical item was as follows: Water be /freezing/ is freezing/ freezes/ at a temperature of 0°C.

Perspectives on CLIL questionnaire

A questionnaire was used to elicit students' perspectives on five dimensions: content learning in CLIL (five items), English learning in CLIL vs. EMI (seven items), English learning in CLIL (seven items), attitude towards content learning in CLIL (seven items), and satisfaction with being taught through CLIL (six items) (**Appendix A**). The Cronbach's alpha internal consistency coefficients for the subdimensions of the questionnaire were 0.72, 0.90, 0.92, 0.77, and 0.87, respectively. An item from PCQ is as follows: I think taking part in this CLIL intervention has improved my level of English language. Items were answered on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The questionnaire sections were validated through content validity procedures. Items were used or adapted from several confirmed instruments (Aguilar & Rodríguez, 2012; Barrios & López-Gutiérrez, 2021; CLIL Activities © Cambridge University Press, 2012; Lasagabaster & Doiz, 2016; McDougald, 2015; Roiha, 2014), and their contents were discussed and decided by the researchers through email communications. Furthermore, we created a shared, editable Google Doc, where we collect questions concerning the development of the questionnaire. All authors contributed with further questions and comments. The tool was validated by three university lecturers working in CLIL/EMI-related fields belonging to the areas of education, English studies, and research methods. The suggestions and remarks of the experts were considered, and the final version of the questionnaire was produced. The different dimensions in PCQ constituted summative scales, and a higher score on the scale represented a higher level in the corresponding construct. Negatively phrased items were reverse coded for analysis. The questionnaire was administered in English.

Data Collection

Data were collected in the autumn 2022 term using OPT and PLCTCQ. The instruments were administered during lesson time. Responses from different year groups, gender, degree program, and previous CLIL were gathered. The collected data were not anonymous, so we were able to match the participants' demographic information with the rest of the information obtained through the questionnaire.

We considered the ethical issues throughout the data collection process. Students voluntarily participated in this study. Participants were informed about

confidentiality that the information they supplied would only be used for this study. Students were also given information on the study objectives, as well as how their data would be processed and how their private information would be maintained.

Data Analysis

To measure any changes in students' grammar and listening skills before and after CLIL implementation, a one-way MANOVA was run. In the case of the significant effect of CLIL, univariate analysis (ANOVA) was conducted for group differences. To calculate the gain or effect size, we preferred omega squared (ω^2), which is an estimate of how much variance in the dependent variables (grammar and listening) are accounted for by the independent variable (CLIL approach with language tuition). ω^2 is recommended for small samples (Olejnik & Algina, 2003) and categorized as small effect=0.01, medium effect=0.06, and large effect=0.14.

PCQ had five dimensions. For the effect of each independent variable (gender, degree program, year group, and previous CLIL experience) on the five dimensions of the questionnaire (perceptions about content learning in CLIL, perceptions about English learning in CLIL vs EMI, perceptions about English learning in CLIL, attitude towards learning in CLIL, and satisfaction with being taught through CLIL) we conducted one-way MANOVA. All the independent variables of this study have two levels. We also looked at the correlations between the OPT and the five dimensions of the PCQ.

CLIL Intervention

One of the co-authors is an instructor and designer of a CLIL course with a background in multilingual education and ten years of teaching experience. The CLIL implementing teachers were science teachers of Physics and Biology. These subject teachers have teaching experiences ranging from two to 27 years. None of them was a native English speaker; however, throughout their teaching careers, they had exclusively taught courses in English.

1. The physics teacher-1 was a senior teacher with over 27 years of teaching experience. He had a master's degree in physics education. He taught the courses optics, mechanics, magnetism, and thermodynamics in English.
2. The biology teacher-1 was a qualified teacher in the process of obtaining his/her PhD degree in biology. This teacher had 20 years of experience teaching biology in English, completed a teacher development program in CLIL in her previous workplace, and had approximately a two-year CLIL teaching experience. This instructor taught

human anatomy, physiology, animal biology, and plant biology in English.

3. The biology teacher-2 had taken a CLIL course as part of his pre-service teacher education program. This was a novice teacher with two years of teaching experience.

The CLIL implementation process started with workshops for these science teachers delivered by the CLIL instructor, who also performed lesson observations via an observation protocol. The CLIL instructor ran two training workshops. The first introductory workshop covered key concepts of the CLIL approach such as types of CLIL, historical background, CLIL operating factors, challenges, and benefits, CLIL and content, cognition, culture, and communication, scaffolding, trans-languaging, BICS and CALP, teacher talking time and student talking time, and HOTS and LOTS. The second practical workshop focused on lesson planning, including language and content goals and lesson sequencing. It covered key elements of lesson planning such as objectives, timing, sequencing, differentiation, and assessment. The subject teachers demonstrated how to write clear CLIL lesson objectives. For instance, one of the lesson objectives of the biology course was "to establish the relationship between the structure and function of chloroplast". In order to integrate language and content, the objective was further developed, and language objectives were also included as follows: "Students will be able to explain the relationship between the structure and function of chloroplast using functional, content-obligatory and content-compatible languages by doing interpersonal activities". Another concept that subject teachers were briefly instructed on was that of differentiation. Differentiation is about tailoring instruction to meet individual needs. The subject teachers were advised to provide multiple modes of learning and input materials to give students options to choose from in order to facilitate their learning. A traditional face-to-face format was used in the first workshop, while a flipped format was used in the second. Before this second workshop, a recorded video explanation was sent to each subject teacher for them to watch. The workshop discussion focused on practical planning and implementation issues.

Following the training of the instructors, they used the CLIL approach in their classes throughout one semester. They were supported in preparing the CLIL lesson plans. The duration of the lessons were 50 minutes. We observed three lessons of each teacher for treatment fidelity through an observation protocol. After each observation, feedback was provided to the instructors. This included details related to the percentage of English language usage, types of interaction, activity types, scaffolding techniques, and communication difficulties. Observations also focused on how lesson goals were being met. The feedback was done in written form with oral one-to-one explanations

Table 1. Descriptive statistics for pre- and post-test scores

	Group	n	Mean	SD	SE
Listening	Post-test	41	76.8	17.0	2.65
	Pre-test	41	61.4	13.3	2.08
Grammar	Post-test	41	77.9	19.9	3.11
	Pre-test	41	54.8	19.6	3.06

Table 2. MANOVA results

	Value	F	df1	df2	p
Pillai's trace	0.301	17.0	2	79	<.001
Wilks' lambda	0.699	17.0	2	79	<.001
Hotelling's trace	0.431	17.0	2	79	<.001
Roy's largest root	0.431	17.0	2	79	<.001

and clarifications to avoid any misunderstandings. Emails and social media chats were used as means by which the subject teachers could reach out to the CLIL instructor at any time if they had questions or concerns about the CLIL planning or implementation process.

RESULTS

Effect of the Combination of a CLIL-Based Approach and Language Instruction

The descriptive statistics shown in **Table 1** include values of grammar and listening for students subjected to the CLIL approach. For both linguistic measures, the mean scores of the two groups (pre- and post-test) did not overlap, which indicates a difference between the groups. Both measures produced higher results in the post-tests than in the pre-tests.

A one-way MANOVA was run to compare the pre-test and post-test score means on the two linguistic measures: grammar and listening. The assumption of homogeneity of covariance matrices was met ($\chi^2=2.49$, $df=3$, $p=.478$). That is to say; the individual group covariance matrices are homogeneous across groups. Moreover, through Shapiro-Wilk multivariate normality test, we found a normal distribution of the scores ($W=0.978$, $p=0.168$).

Based on the results displayed in **Table 2**, the Pillai's trace values ($F[2, 79]=17.0$, $p<.001$), it can be concluded

Table 3. Univariate tests results

	DV	SS	df	MS	F	p	ω^2
Group	Listening	4871	1	4871	21.0	<.001	0.196
	Grammar	10960	1	10960	28.0	<.001	0.248
Residuals	Listening	18577	80	232			
	Grammar	31301	80	391			

Note. DV: Dependent variable; SS: Sum of squares; & MS: Mean square

that there were significant differences between the pre-test and post-test means of the students on the two dimensions of language learning. Thus, the CLIL approach together with formal language tuition, significantly increased students' language learning.

According to the univariate tests results presented in **Table 3**, both linguistic measures were significantly influenced by the CLIL approach coupled with formal language learning. There is a significant CLIL effect on listening ($F[1, 80]=21.0$, $\omega^2=0.196$, $p<.001$). It also showed a significant effect of CLIL on the grammar ($F[1, 80]=28.0$, $\omega^2=0.248$, $p<.001$). The CLIL approach used in this study has a large effect on listening and grammar ($\omega^2=0.196$, and $\omega^2=0.248$, respectively).

As seen in **Table 1**, post-test means are higher than pre-test means for both grammar and listening, and this difference was proved to be statistically significant through the MANOVA analysis presented in **Table 2** and **Table 3**.

Student Perceptions Concerning the CLIL-Based Intervention According to Individual and Institutional Factors

As mentioned above, the perspectives on CLIL questionnaire (PCQ) had five dimensions. For the effect of each independent variable (gender, program, year group, and previous CLIL experience) on the five dimensions of the questionnaire, we conducted one-way MANOVAs. The MANOVA results for the main effects are presented in text, and the univariate test results are shown in **Table 4**.

Table 4. Univariate tests

	Dependent variable	df	F	p	ES
Gender	Perceptions about content learning in CLIL	1	1.93	0.168	-0.328
	Perceptions about English learning in CLIL vs. EMI	1	0.23	0.635	-0.107
	Perceptions about English learning in CLIL	1	0.67	0.417	-0.188
	Attitude towards learning in CLIL	1	1.06	0.305	-0.232
	Satisfaction with being taught through CLIL	1	0.69	0.408	-0.196
Academic programme	Perceptions about content learning in CLIL	1	0.04	0.847	-0.047
	Perceptions about English learning in CLIL vs. EMI	1	5.00	0.028	-0.512
	Perceptions about English learning in CLIL	1	4.18	0.044	-0.452
	Attitude towards learning in CLIL	1	1.55	0.216	-0.264
	Satisfaction with being taught through CLIL	1	0.19	0.667	-0.09
Year group	Perceptions about content learning in CLIL	1	5.36	0.023	0.507
	Perceptions about English learning in CLIL vs. EMI	1	3.99	0.049	0.438
	Perceptions about English learning in CLIL	1	5.51	0.021	0.517
	Attitude towards learning in CLIL	1	4.81	0.031	0.488
	Satisfaction with being taught through CLIL	1	6.34	0.014	0.551

Table 4 (Continued). Univariate tests

	Dependent variable	df	F	p	ES
Previous CLIL	Perceptions about content learning in CLIL	1	0.99	0.323	-0.233
	Perceptions about English learning in CLIL vs. EMI	1	11.09	0.001	-0.752
	Perceptions about English learning in CLIL	1	13.86	<.001	-0.849
	Attitude towards learning in CLIL	1	5.30	0.024	-0.525
	Satisfaction with being taught through CLIL	1	3.32	0.072	-0.409

The results of the MANOVA showed a nonsignificant main effect for the ‘gender’, ‘program’, and ‘year group’ factors ($F=0.430$; $df=5-86$; $p=.826$), ($F=1.95$; $df=5-86$; $p=.095$), and ($F=1.82$; $df=5-86$; $p=.118$), respectively. The p-values associated with these analyses are .826, .095, and .118, respectively. The univariate test was not significant for the relatively high value (.826) of the “p” for gender. However, we found significant univariate test results for those (.095, and .118) close to .05 (see **Table 4**). On the other hand, the results of the MANOVA showed a significant main effect for the “previous CLIL experience” factor ($F=3.19$; $df=5-73$; $p=.012$).

The univariate results for four independent variables across five dimensions of the perception survey are presented in **Table 4**.

Univariate analyses of the ‘gender’ factor was not significant for any of the dimensions of the questionnaire. Univariate analyses of the ‘program’

factor were significant for the perceptions about English learning in CLIL vs EMI ($F=5.00$; $df=1$; $p=.028$) dimension and for the perceptions about English learning in CLIL ($F=4.18$; $df=1$; $p=.044$) dimension of the questionnaire. Univariate analyses of the ‘year group’ factor was significant for all questionnaire dimensions. Statistical details can be seen in **Table 4**. Except for the perceptions about content learning in CLIL, all other dimensions of the questionnaire are significant for the ‘previous CLIL experience’ factor. All significant “p values” in **Table 4** are bolded.

The descriptive statistics for all independent variables across the questionnaire dimensions are presented in **Table 5**. All the independent variables in this study have two levels: that is, gender (female and male), program (chemistry-biology and physics-informatics), year group (first year and higher years), and previous CLIL experience (yes and no). To determine the significant univariate test results in favor

Table 5. Descriptive statistics across variables

Variable	Dimension	Group	n	Mean	Median	SD
Gender	Perceptions about content learning in CLIL	Female	66	3.39	3.2	0.766
		Male	26	3.12	3.1	0.955
	Perceptions about English learning in CLIL vs. EMI	Female	67	3.74	3.86	0.79
		Male	26	3.65	3.64	0.951
	Perceptions about English learning in CLIL	Female	67	3.8	3.71	0.759
		Male	26	3.64	3.79	1.061
	Attitude towards learning in CLIL	Female	67	3.43	3.43	0.625
		Male	26	3.27	3.21	0.839
	Satisfaction with being taught through CLIL	Female	67	3.71	3.67	0.735
		Male	26	3.56	3.55	0.844
Academic program	Perceptions about content learning in CLIL	CB	50	3.33	3.3	0.789
		PI	33	3.29	3.2	0.94
	Perceptions about English learning in CLIL vs. EMI	CB	51	3.86	3.86	0.765
		PI	33	3.45	3.43	0.855
	Perceptions about English learning in CLIL	CB	51	3.88	4	0.789
		PI	33	3.50	3.6	0.914
	Attitude towards learning in CLIL	CB	51	3.42	3.43	0.671
		PI	33	3.24	3	0.701
	Satisfaction with being taught through CLIL	CB	51	3.66	3.67	0.737
		PI	33	3.59	3.5	0.844
Year group	Perceptions about content learning in CLIL	First	62	3.18	3.1	0.747
		Above	30	3.59	3.8	0.924
	Perceptions about English learning in CLIL vs. EMI	First	63	3.60	3.57	0.841
		Above	30	3.96	4	0.778
	Perceptions about English learning in CLIL	First	63	3.62	3.71	0.883
		Above	30	4.05	4.07	0.707
	Attitude towards learning in CLIL	First	63	3.28	3.14	0.697
		Above	30	3.61	3.71	0.631
	Satisfaction with being taught through CLIL	First	63	3.54	3.5	0.797
		Above	30	3.95	3.92	0.617

Note. CB: Chemistry-biology & PI: Physics-informatics

Table 5 (continued). Descriptive statistics across variables

Variable	Dimension	Group	n	Mean	Median	SD
Previous CLIL	Perceptions about content learning in CLIL	Yes	35	3.46	3.6	0.926
		No	44	3.26	3.2	0.801
	Perceptions about English learning in CLIL vs. EMI	Yes	35	4.07	4.14	0.87
		No	44	3.46	3.43	0.762
	Perceptions about English learning in CLIL	Yes	35	4.17	4.29	0.857
		No	44	3.48	3.57	0.776
	Attitude towards learning in CLIL	Yes	35	3.62	3.67	0.677
		No	44	3.25	3.29	0.727
	Satisfaction with being taught through CLIL	Yes	35	3.86	3.83	0.814
		No	44	3.54	3.42	0.756

Note. CB: Chemistry-biology & PI: Physics-informatics

Table 6. Correlation between English proficiency and students' perceptions, attitudes, and satisfaction

	Post-test		Pre-test	
	Pearson's r	p	Pearson's r	p
Perceptions about content learning in CLIL	0.455	0.033	0.28	0.207
Perceptions about English learning in CLIL vs. EMI	0.097	0.666	0.096	0.672
Perceptions about English learning in CLIL	0.096	0.671	0.163	0.47
Attitude towards learning in CLIL	-0.045	0.842	-0.193	0.39
Satisfaction with being taught through CLIL	-0.086	0.704	0.149	0.507

of groups, we need the means of the groups (see Table 5).

The results presented in Table 5 can be summarized according to the findings of univariate analyses. Namely, for the gender groups, no significant group difference between females and males was observed. Chemistry-biology students have more positive perceptions than physics-informatics students for the program groups in the second (English learning in CLIL versus EMI) and third (English learning in CLIL) dimensions of the questionnaire. For year groups, the higher-year students have more positive perceptions in all dimensions of the questionnaire than the first-year students. For the previous CLIL experience groups, except for the first dimension, students with previous CLIL experience have more positive perceptions in all dimensions than those who were new to CLIL.

The effect sizes were also calculated for practical significance from the differences between the means of the groups. Using Cohen's classification of effect size (small, $d=0.10-0.29$; medium, $d=0.30-0.49$; large, $d>0.50$), it can be said that the difference between the means of the groups is practically significant for many groups. For example, in the 'year group' variable, an effect size of 0.517 for the *perceptions about English learning in CLIL* shows that the average score of higher graders is 0.517 standard deviations above the average score of the first graders. Similarly, in the previous CLIL experience variable, an effect size of -0.752 for the perceptions about English learning in CLIL vs EMI indicates that the average score of those who did not have previous experience in CLIL is 0.752 standard deviations below the average score of those who had previous CLIL experience.

Students' Proficiency Level and Perspectives on CLIL

We further investigated the correlation between students' proficiency level and their perceptions, attitudes and satisfaction with learning content and language through the CLIL approach. We correlated the students' scores from the pre-test and post-test scores and the five perspective dimensions in the study. The results are presented in Table 6.

As seen in Table 6, except for one, all correlations are insignificant. In other words, there is only a significant positive relationship between students' post-test scores and *perceptions about content learning in CLIL*.

DISCUSSION

Research so far in HE English-medium education contexts has only attested to modest gains in language competence (Aguilar & Muñoz, 2014; Barrios & López-Gutiérrez, 2021; Rogier, 2012; Yang, 2015), which may fall short of the language learning expectations that English-medium education may create. However, according to our study results, implementing a CLIL approach in HE coupled with formal general and domain-specific language learning was highly effective for general English learning. Our results thus seem to provide evidence that language instruction is needed in CLIL/EMI to secure the language progress these approaches purport to ensure. Additionally, the study contributes to ameliorating the dearth of research objectively assessing the effect of English-medium education on English proficiency through language tests that has been highlighted in the literature (Airey, 2004; Macaro et al., 2018, p. 57). Furthermore, our research strengthens the case for including language learning objectives in English-medium education and setting up

the necessary conditions to ensure that this learning occurs (e.g., Pecorari et al., 2011), as it was the case with the CLIL initiative investigated here.

The large effect size of the difference in English language skills before and after the CLIL intervention coupled with language lessons is most probably associated with the students' English proficiency before the intervention. Both in Aguilar and Muñoz (2014) and Barrios and López-Gutiérrez (2021), the students who made the greatest gains in language proficiency were the least proficient—A2 to B1—students according to the CEFR (Council of Europe, 2001, 2020). In our study, the students scored a mean of 116.2 in the overall OPT1 pre-test. As OPT scores are calibrated against CEFR proficiency levels, this means that before the one-semester CLIL and language tuition intervention, the sample proficiency as a group was near the top of an A2 CEFR level (OPT scores 105-119 are claimed to correspond to an A2 level [Allan, 2005a]). After the intervention, the mean score for the OPT2 test was 154.7, which represents an average of more than a band score gain on CEFR (scores of 135-149 on the OPT are considered to equate to a B2 level). This indicates that the combination of the general and domain-specific (ESP) English tuition and the CLIL intervention, most probably in association with high academic motivation—as language proficiency is key to success in EMI—, can be very effective in enhancing English language proficiency. This finding lends support to the claim that English-medium education initiatives need to incorporate explicit language learning goals (Rogier, 2012), including academic and discipline-specific ones (Airey et al., 2017; Galloway & Ruegg, 2020; Kuteeva & Airey, 2014), and a corresponding provision of general, academic, and disciplinary language support (Kling, 2017; Kuteeva & Airey, 2014; Rose et al., 2020).

Our research also yields interesting results as to how the institutional and individual factors under study are associated with perspectives related to CLIL, namely perceptions about content and about English learning in CLIL, perceptions about English learning in CLIL as compared to English learning in their EMI program, attitudes towards learning in CLIL, and satisfaction with being taught through CLIL.

Female students' perspectives about CLIL are invariably more enthusiastic than those held by their male counterparts and hold the view that they learn more content and English, that they learn more English than in pure EMI with no content-related language component, and that their attitude towards and satisfaction with CLIL are also more positive than those expressed by male students. Differences, however, do not reach significance level. Contradictory results concerning gender differences in views on CLIL are found in the literature. As in our study, Ismail et al. (2011) found no significant gender differences between science and mathematics female and male students in

Malaysia in inclination towards EMI. In contrast, Hengsaddeekul et al. (2014) reported that females expressed a significantly higher preference for English-medium graduate studies than males in Thailand. The female students in the study conducted by Macaro and Akincioglu (2018) in Turkish universities reported a significantly larger improvement in their general and discipline-specific English proficiency than male students. Somewhere in the middle lay the results obtained by Lasagabaster (2016), who found that attitudes to CLIL were significantly more positive in female university students. However, the small effect size led him to conclude that “gender differences tend to disappear in EMI university contexts” (p. 326).

As to the academic program, biology-chemistry students recognize significantly higher English language gains in CLIL than the physics-informatics students. The teacher effect could partly explain the dissimilar results in the two groups. It could also be the case that, because of the features of the discipline-specific discourse, biology-chemistry students are more exposed to larger amounts of English input than physics-informatics students and/or they need to produce more output in English. This possible explanation must be confirmed or disproved by further studies. Hengsaddeekul et al. (2014) identified a significant difference across nine different fields of study in Thailand, although the difference was in terms of self-rated English language proficiency, not in objective language gains. As **Table 5** shows, the rest of the perceptions, attitudes towards and satisfaction with CLIL also depict a more favorable picture of this approach among the biology-chemistry students. Biology-chemistry students' average higher grade point average can be another explanation for their higher gains in English language. Several studies (e.g., Humphreys et al., 2012; Schoepp, 2018) have suggested a relationship between language proficiency and academic achievement.

CLIL experience in pre-university education seems to be an additional factor affecting students' perspectives on CLIL, as students with previous experience have more positive views than those who are new to CLIL in all dimensions being examined. The former consider that they learn more content and English in CLIL, show more positive attitudes, and are more satisfied with CLIL. In terms of practical significance, effect sizes range from medium to large. Whether or not students with CLIL experience also had a higher level of proficiency was not controlled by the present investigation, and future studies will need to determine if familiarity with the CLIL approach in pre-university education is a predictor of students' views of CLIL or EMI at tertiary education. To the best of the authors' knowledge, the only study that aimed to find the association between students' perceptions—more precisely, inclination towards choosing EMI—and previous experience in learning through English is that of Ismail et al. (2011). They did

not find a statistically significant difference in inclination toward choosing EMI and the language for learning at primary and secondary levels. However, their study results led them to conclude that participants who studied science and mathematics in English during their pre-university education are in favor of EMI.

Concerning the year group, all measures reveal that first-year students feel that they learn less content and English during the CLIL experience, show less favorable attitudes about CLIL, and are less satisfied with CLIL than the students in higher years of study. Significant differences were identified when comparing both groups of students in all the five CLIL perspective dimensions, with effect sizes ranging from medium to large, i.e., substantial enough to be considered practically significant. It needs to be noted that these university students undertake an English skills together with an ESP development course for an academic year and that the general English competence they reach after that year may be a necessary condition for them to learn both the content in English and the discipline-specific discourse. It has been argued (e.g., Johnson & Swain, 1997; Leaver & Stryker, 1997) that a certain threshold level in the language is necessary in order to acquire appropriate disciplinary knowledge. In the study conducted by Rose and McKinley (2018), students above a proficiency of IELTS 6 (i.e., a CEFR C1) experienced fewer challenges learning content through English than those below the threshold. Rose et al. (2020) also found that English knowledge was a statistically significant predictor of success in EMI. Curiously enough, among the correlations between English proficiency and the dimensions of perspectives on CLIL explored in our study, only that between proficiency and perceptions of content learning was significant before the CLIL intervention accompanied by language instruction. The students' mean English proficiency level was at basic user CEFR A2 at that time, which may be indicative that students with a low proficiency felt that they were finding content learning in CLIL challenging, while the reverse was the case with students with high English proficiency. At the end of the intervention, however, this correlation ceases to exist, and this coincides with a significantly higher English proficiency mean score. In sum, both our findings and previous ones suggest that students have distinct EMI experiences concerning content learning depending on their level of linguistic proficiency. However, contrary to findings in previous studies (Barrios & López-Gutiérrez, 2021; Barrios et al., 2022; Fernández-Costales, 2017; Muñoz, 2001), where the higher the students' English proficiency, the less satisfied they were with their language improvement from participation in EMI and with the EMI program, our findings do not confirm these associations. It is worth noting that, unlike other studies in which self-reports of proficiency were used, a standardized English proficiency measure was used in the present study.

CONCLUSIONS

The study behind this article aimed to explore the effect of a one-semester CLIL intervention combined with English language instruction on university students' general English skills and the associations between institutional and individual factors and views on CLIL. It has provided evidence that a combination of general and domain-specific language instruction within a framework of CLIL has a great potential for the language development that is claimed to be a necessary condition for success and content learning in English-medium education. Interesting results have also been found concerning how institutional and individual factors relate to students' views about CLIL. In terms of the practical significance of the findings, two factors seem to be of particular relevance for student recognition of higher levels of language and disciplinary knowledge learning, positive attitude towards and higher satisfaction with CLIL: being a non-first-year student and having CLIL experience in pre-university education. In light of the results, our findings reinforce the case for integrating language instruction that includes general, academic, and subject-specific goals in English-medium programs and courses in HE.

Despite these interesting findings concerning under-researched areas in HE CLIL, some limitations must be acknowledged. Firstly, our investigation only reports on quantitative data: qualitative research is required to have a better understanding of the impact of CLIL on the students' learning process. Secondly, we must bear in mind that this investigation was carried out in one institution of HE, so findings may be taken with caution as they cannot be directly generalized to other settings. And thirdly, it is also worth mentioning that we did not analyze the isolated impact of CLIL from the effect of general and discipline-specific English instruction. Although no claims are made in the study attributing language gains to CLIL alone, further research is needed that isolates both the impact of CLIL/EMI and also that of general and domain-specific instruction in language learning in HE English-medium education.

Notwithstanding these limitations, our study provides valuable results in an area that is in dire need of empirical evidence to inform institutional policies and practices in HE English-medium education concerning the effect and value of language development initiatives for content and language learning. Besides, the study explores the association between perceptions of, attitudes towards and satisfaction with CLIL, and institutional and individual factors that have rarely been studied by CLIL or EMI research. Finally, the study takes place in a context, that of HE Kazakhstan, that is currently under-explored in the literature on English-taught programs. In this sense, it contributes to expanding the research focus in this field to include areas of the world where the situation of HE English-

medium education is poorly represented in the scientific literature. Prospective lines of enquiry to be pursued may focus on the cumulative effect of CLIL programs in longitudinal studies. In addition, qualitative and ethnographic research aiming to better understand the impact of CLIL on both the learners' and the teachers' learning process is welcome.

Author contributions: All authors have sufficiently contributed to the study and agreed with the results and conclusions.

Funding: No funding source is reported for this study.

Declaration of interest: No conflict of interest is declared by authors.

Data sharing statement: Data supporting the findings and conclusions are available upon request from the corresponding author.

REFERENCES

- Aguilar, M., & Muñoz, C. (2014). The effect of proficiency on CLIL. Benefits in engineering students in Spain. *International Journal of Applied Linguistics*, 24(1), 1-18. <https://doi.org/10.1111/ijal.12006>
- Aguilar, M., & Rodríguez, R. (2012). Lecturer and student perceptions on CLIL at a Spanish university. *International Journal of Bilingual Education and Bilingualism*, 15(2), 183-197. <https://doi.org/10.1080/13670050.2011.615906>
- Airey, J. (2004). Can you teach it in English? Aspects of the language choice debate in Swedish higher education. In R. W. (Ed.), *Integrating content and language: Meeting the challenge of a multilingual higher education* (pp. 97-108). Maastricht Universitaire Pers.
- Airey, J., Lauridsen, K. M., Räsänen, A., Salö, L., & Schwach, V. (2017). The expansion of English-medium instruction in the Nordic countries: Can top-down university language policies encourage bottom-up disciplinary literacy goals? *Higher Education*, 73(4), 561-576. <https://doi.org/10.1007/s10734-015-9950-2>
- Allan, D. (2005a). *Oxford placement test 1*. Oxford University Press.
- Allan, D. (2005b). *Oxford placement test 2*. Oxford University Press.
- Barrios, E., & López-Gutiérrez, A. (2021). Language learning in a partially English-taught teacher education programme: Language gains and student perceptions. *International Journal of Multilingualism*. <https://doi.org/10.1080/14790718.2021.1888956>
- Barrios, E., López-Gutiérrez, A., & Lopez-Agudo, L.-A. (2022). Language-related perceptions: How do they predict student satisfaction with a partial English medium instruction in higher education? *Journal of English for Academic Purposes*, 57, 101121. <https://doi.org/10.1016/j.jeap.2022.101121>
- CLIL Activities © Cambridge University Press. (2012). *Checklist: How 'CLIL' are you?* https://ateneu.xtec.cat/wiki/form/wikiexport/_media/cmd/llc/cls/modul_6/checklist_how.pdf
- Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge University Press.
- Council of Europe. (2020). *Common European framework of reference for languages: Learning, teaching, assessment—Companion volume*. Council of Europe Publishing.
- Creswell, J. W. (1994). *Research design: Qualitative & quantitative approaches*. SAGE.
- Dearden, J. (2014). *English as a medium of instruction—a growing global phenomenon*. British Council.
- Dewaele, J. M., & Saito, K. (2022). Positive psychology can combat the pernicious native speaker ideology. *The European Educational Researcher*, 5(2), 225-234. <https://doi.org/10.31757/euer.526>
- Doiz, A., & Lasagabaster, D. (2022). Looking into English-medium instruction teachers' metadiscourse: An ELF perspective. *System*, 105, 102730. <https://doi.org/10.1016/j.system.2022.102730>
- Doiz, A., Lasagabaster, D., & Sierra, J. M. (Eds.) (2013). *English-medium instruction at universities: Global challenges*. Multilingual Matters. <https://doi.org/10.21832/9781847698162>
- Elder, C., & O'Loughlin, K. (2003). Investigating the relationship between intensive English language study and band score gain on IELTS. *IELTS Research Reports*, 4. <https://www.ielts.org/for-researchers/research-reports/volume-04-report-6>
- Fernández-Costales, A. (2017). Evaluación de las percepciones del alumnado sobre la utilización del inglés como lengua de instrucción en la Universidad [Evaluation of students' perceptions about the use of English as the language of instruction at the university]. *Didáctica. Lengua y Literatura [Didactics. Language and Literature]*, 29, 43-63. <https://doi.org/10.5209/dida.57129>
- Fernández-Costales, A. (2022). Cognitive development in content and language integrated learning. In D. Banegas, & S. Zappa-Hollman, S. (Eds.), *The Routledge handbook of CLIL*. Routledge.
- Fernández-Costales, A., & González-Riaño, X. A. (2015). Teacher satisfaction concerning the implementation of bilingual programmes in a Spanish university. *Porta Linguarum*, 23, 93-108. <https://doi.org/1697-7467>
- Galloway, N., & Ruegg, R. (2020). The provision of student support on English medium instruction programmes in Japan and China. *Journal of English for Academic Purposes*, 45, 100846. <https://doi.org/10.1016/j.jeap.2020.100846>

- Galloway, N., & Rose, H. (2021). English medium instruction and the English language practitioner. *ELT Journal*, 75(1), 33-41. <https://doi.org/10.1093/elt/ccaa063>
- Hengsadeekul, C., Koul, R., & Kaewkuekool, S. (2014). Motivational orientation and preference for English-medium programs in Thailand. *International Journal of Educational Research*, 66, 35-44. <https://doi.org/10.1016/j.ijer.2014.02.001>
- Huertas Abril, C., & Shashken, A. (2021). Exploring the potential of CLIL in Kazakhstan: A qualitative study. *Revista Complutense de Educación [Complutense Journal of Education]*, 32(2), 261-271. <https://doi.org/10.5209/rced.68345>
- Humphreys, P., Haugh, M., Fenton-Smith, M., Lobo, A., Michael, R., & Walkinshaw, I. (2012). Tracking international students' English proficiency over the first semester of undergraduate study. In J. Osborne, & G. Lim (Eds.), *IELTS research report series* (pp. 1-41). IDP: IELTS Australia.
- Ismail, W. R., Mustafa, Z., Muda, N., Abidin, N. Z., Isa, Z., Zakaria, A. M., Suradi, N. R. M., Mamat, N. J. Z., Nazar, R. M., Ali, Z. M., Rafee, N. M., Majid, N., Jaaman, S. H., Darus, M., Ahmad, R. R., Shahabuddin, F. A., Rambely, A. S., Din, U. K. S., Hashim, I., ... Azlan, M. I. (2011). Students' inclination towards English language as medium of instruction in the teaching of science and mathematics. *Procedia-Social and Behavioral Sciences*, 18, 353-360. <https://doi.org/10.1016/j.sbspro.2011.05.050>
- Johnson, R. K., & Swain, M. (1997). *Immersion education: International perspectives*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139524667>
- Karabassova, L. (2019). CLIL or "just good teaching" in Kazakhstan. *Revista Nebrija de Lingüística Aplicada a la Enseñanza de Lenguas [Nebrija Journal of Linguistics Applied to Language Teaching]*, 13(27), 56-71. <https://doi.org.10.26378/rmlael1327341>
- Kling, J. (2017). English medium instruction and the international classroom. In A. M. Snow, & D. M. Brinton (Eds.), *The content-based classroom: perspectives on integrating language* (pp. 216-227). University of Michigan Press.
- Kuteeva, M., & Airey, J. (2014). Disciplinary differences in the use of English in higher education: Reflections on recent policy developments. *Higher Education*, 67(5), 533-549. <https://doi.org/10.1007/s10734-013-9660-6>
- Lasagabaster, D. (2021). EMI in Spain: Striving to maintain a multilingual balance. In R. Wilkinson, & R. Gabriëls (Eds.), *The Englishization of higher education in Europe* (pp. 77-96). Amsterdam University Press. <https://doi.org/10.2307/j.ctv21ptzkn.7>
- Lasagabaster, D., & Doiz, A. (2016). CLIL students' perceptions of their language learning process: Delving into self-perceived improvement and instructional preferences. *Language Awareness*, 25(1-2), 110-126. <https://doi.org/10.1080/09658416.2015.1122019>
- Lasagabaster, D., & Doiz, A. (Eds.) (2021). *Language use in English-medium instruction at university: International perspectives on teacher practice*. Routledge. <https://doi.org/10.4324/9781003134534>
- Lasagabaster, D., & Ruiz de Zarobe, Y. (2010). *CLIL in Spain: Implementation, results and teacher training*. Newcastle-upon-Tyne: Cambridge Scholars.
- Leaver, B. L., & Stryker, S. B. (1997). *Content-based Instruction in foreign language education: Models and methods*. Georgetown University Press.
- Lei, J., & Hu, G. W. (2014). Is English-medium instruction effective in improving Chinese undergraduate students' English competence? *International Review of Applied Linguistics in Language Teaching*, 52(2), 99-126. <https://doi.org/10.1515/iral-2014-0005>
- Macaro, E. (2018). *English medium instruction: Content and language in policy and practice*. Oxford University Press.
- Macaro, E., & Akincioglu, M. (2018). Turkish university students' perceptions about English medium instruction: Exploring year group, gender and university type as variables. *Journal of Multilingual and Multicultural Development*, 39(3), 256-270. <https://doi.org/10.1080/01434632.2017.1367398>
- Macaro, E., Curle, S., Pun, J., An, J., & Dearden, J. (2018). A systematic review of English medium instruction in higher education. *Language Teaching*, 51(1), 36-76. <https://doi.org/10.1017/S0261444817000350>
- McDougald, J. (2015). Teachers' attitudes, perceptions and experiences in CLIL: A look at content and language. *Colombian Applied Linguistics Journal*, 17(1), 25-41. <https://doi.org/10.14483/udistrital.jour.calj.2015.1.a02>
- Mills, A. J., Durepos, G., & Wiebe, E. (Eds.). (2009). *Encyclopedia of case study research*. SAGE. <https://doi.org/10.4135/9781412957397>
- Mitchel, H. Q., & Malkogianni, M. (2015). *Pioneer elementary student's book*. MM Publication.
- Muñoz, C. (2001). The use of the target language as the medium of instruction. University students' perceptions. *Anuari de Filologia [Yearbook of Philology]*, 23(10), 71-82.
- Neuendorf, J. (2019). Kazakhstan's trinity of languages program—A model of success or failure? In

- Proceedings of the Central Eurasian Studies Society Conference*. The Central Eurasian Studies Society.
- O'Dowd, R. (2015). The training and accreditation of teachers for English medium instruction: An overview of practice in European universities. *International Journal of Bilingual Education and Bilingualism*, 28(5), 553-563. <https://doi.org/10.1080/13670050.2018.1491945>
- Olejnik, S., & Algina, J. (2003). Generalized eta and omega squared statistics: Measures of effect size for some common research designs. *Psychological Methods*, 8(4), 434-447. <https://doi.org/10.1037/1082-989X.8.4.434>
- Ozdemir, C. (2018). Causes of problems students face while learning Turkish as a foreign language: The case of Kazakhstan. *The European Educational Researcher*, 1(1), 17-33. <https://doi.org/10.31757/euer.112>
- Pecorari, D., Shaw, P., Irvine, A., & Malmström, H. (2011). English for academic purposes at Swedish universities: Teachers' objectives and practices. *Ibérica [Iberian]*, 22, 55-77.
- 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>
- Rogier, D. (2012). *The effects of English-medium instruction on language proficiency of students enrolled in higher education in the UAE* [PhD thesis, University of Exeter].
- Roiha, A. S. (2014). Teachers' views on differentiation in content and language integrated learning (CLIL): Perceptions, practices and challenges. *Language and Education*, 28(1), 1-18. <https://doi.org/10.1080/09500782.2012.748061>
- Rose, H., & McKinley, J. (2018). Japan's English-medium instruction initiatives and the globalization of higher education. *Higher Education*, 73(1), 111-29. <https://doi.org/10.1007/s10734-017-0125-1>
- Rose, H., Curle, S., Aizawa, I., & Thompson, G. (2020). What drives success in English medium taught courses? The interplay between language proficiency, academic skills and motivation. *Studies in Higher Education*, 45(11), 2149-2161. <https://doi.org/10.1080/03075079.2019.1590690>
- Satayev, M., Balta, N., Shaymerdenovna, I. R., Fernández-Cézar, R., & Alcaraz-Mármol, G. (2022). Content and language integrated learning implementation through team teaching in Biology lessons: A quasi-experimental design with university students. *Frontiers in Education*, 7, 867447. <https://doi.org/10.3389/feduc.2022.867447>
- Schoepp, K. (2018). Predictive validity of the IELTS in an English as a medium of instruction environment. *Higher Education Quarterly*, 72(4), 271-285. <https://doi.org/10.1111/hequ.12163>
- Scott, D., & M. Morrison. (2006). *Key ideas in educational research*. Continuum.
- Seitzhanova, A., Plokhikh, R., Baiburiev, R., & Tsaregorodtseva, A. (2015). English as the medium of instruction: Modern tendency of education in Kazakhstan. *Perspectives of Innovations, Economics and Business*, 15(3), 113-117. <https://doi.org/10.15208/pieb.2015.11>
- Toledo, I., Rubio, F. D., & Hermosín, M. (2011). Creencias, rendimiento académico y actitudes de alumnos universitarios principiantes en un programa plurilingüe [Beliefs, academic performance and attitudes of beginning university students in a multilingual program]. *Porta Linguarum [The Gate of Languages]*, 18, 213-229. <https://doi.org/10.30827/Digibug.15836>
- Yang, W. (2015). Content and language integrated learning next in Asia: Evidence of learners' achievement in CLIL education from a Taiwan tertiary degree programme. *International Journal of Bilingual Education and Bilingualism*, 18(4), 361-382. <https://doi.org/10.1080/13670050.2014.904840>
- Zharkynbekova, S., Akynova, D., & Aimoldina, A. (2017). The linguistic landscape in Kazakhstan: Globalization, triglossia and other challenges. *Cross-Cultural Studies: Education and Science*, 2(1), 89-99.

APPENDIX A

Table A. Perspectives on CLIL questionnaire

Dimension	Item	Score
Perceptions about content learning in CLIL	I would have learned more academic/disciplinary content if lessons had been delivered in non-CLIL environment.	
	I would rather learn academic/disciplinary content in non-CLIL environment.	
	I find the activities in these CLIL lessons more interesting than in non-CLIL environment.	
	I feel I do not learn academic/disciplinary content well because it is delivered in CLIL method.	
	Dealing with language content in CLIL has facilitated me to learn the subject content better than in non-CLIL environment.	
Perceptions about English learning in CLIL vs. EMI	I have improved my listening in English more than in the regular English class (understanding of lecturers' explanations and presentations and similar activities).	
	I have improved my speaking in English more than in the regular English class (delivering oral presentations and similar activities).	
	I have improved my oral interaction in English more than in the regular English class (participating in dialogues, debates, and similar activities).	
	I have improved my reading in English more than in the regular English class (understanding academic documents, online information, etc. and similar activities).	
	I have improved my writing in English more than in the regular English class (writing essays, exam questions, portfolios, and similar activities).	
	I have improved my vocabulary in English more than in the regular English class.	
	I have improved my grammar in English more than in the regular English class.	
Perceptions about English learning in CLIL	I think taking part in this CLIL intervention has improved my level of English language.	
	Taking part in the CLIL intervention has improved my listening skills in English (understanding of lecturers' explanations and presentations and similar activities).	
	Taking part in the CLIL intervention has improved my speaking skills in English (delivering oral presentations and similar activities).	
	Taking part in the CLIL intervention has improved my oral interaction skills in English (participating in dialogues, debates, and similar activities).	
	Taking part in the CLIL intervention has improved my reading skills in English (understanding academic documents, online information, etc. and similar activities).	
	Taking part in the CLIL intervention has improved my writing skills in English (writing essays, exam questions, portfolios, and similar activities).	
	My English proficiency is improving thanks to these CLIL lessons.	
Attitude towards content learning in CLIL	The CLIL lessons are more difficult than non-CLIL lessons	
	Subject lessons in non-CLIL environment are easier than in CLIL lessons	
	Subject lessons in non-CLIL environment are more understandable than in CLIL lessons	
	I prefer learning subject content through CLIL than through non-CLIL environment	
	Learning through CLIL has been a positive experience.	
	I am more motivated in these CLIL lessons (than in the regular English-delivered lessons)	
Satisfaction with being taught through CLIL	I would recommend other students to have lessons with CLIL.	
	In general, the quality of lessons in this CLIL experience is satisfactory.	
	I find the activities in CLIL lessons more interesting than those in non-CLIL environment.	
	CLIL lessons are easier to follow than non-CLIL lessons.	
	CLIL lessons are more understandable than non-CLIL lessons.	
	I am generally satisfied with the methodology used this CLIL experience	
I am satisfied with the academic/disciplinary content I have learned through CLIL.		

<https://www.ejmste.com>