

# Influence of teachers' linguistic competence on the implementation of CLIL in Physical Education

FRANCISCO JOSÉ MARTÍNEZ-HITA

*Universidad de Zaragoza*

MANUEL GÓMEZ-LÓPEZ

*Universidad de Murcia*

DAVID MANZANO SÁNCHEZ

*Universidad de Extremadura*

Received: 16/02/2022 / Accepted: 20/06/2023

DOI: <https://doi.org/10.30827/portalin.viVII.29170>

ISSN paper edition: 1697-7467, ISSN digital edition: 2695-8244

**ABSTRACT:** The aim of this paper was to know the influence of the qualification level in the second foreign language in order to see the differences in teacher perception in Bilingual Physical Education classes (AICLE). For this, 146 active teachers in Physical Education participated in the present study, to whom the Physical Education Session Evaluation questionnaire (CESEFA) was administered through. Effects based on the sex of the participants and age category were also checked. The analysis of results showed that the participants who had a C1 level of English had a more positive perception of all categories analyzed. In turn, no differences were found based on the sex of the participants, but were found based on the age category, finding that those under 36 with a B2 level had a better perception of their classes in the items: take advantage of breaks, adaptation of content and feedback. It is concluded that the degree of qualification influences the perception of Physical Education classes, and that in the case of the youngest subjects, there are elements related to the individual efficacy variable, where no differences were found according to the degree of qualification or even these are positive in favor of level B2.

**Key words:** evaluation, bilingualism, English, quality, PE.

## **Influencia de la competencia lingüística del profesorado en la implantación AICLE en Educación Física**

**RESUMEN:** El objetivo de este trabajo fue conocer la influencia del nivel de calificación en la segunda lengua extranjera para ver las diferencias en la percepción del profesorado en las clases de Educación Física Bilingüe (AICLE). Para ello, en el presente estudio participaron 146 profesores activos de Educación Física, a quienes se les aplicó el cuestionario de Evaluación de Sesiones de Educación Física (CESEFA). También se comprobaron los efectos según el sexo de los participantes y la categoría de edad. El análisis de resultados mostró que los participantes que tenían un nivel C1 de inglés tenían una percepción más positiva de todas las categorías analizadas. A su vez, no se encontraron diferencias según el sexo de los participantes, pero sí según la categoría de edad, encontrando que los menores de 36 años con un nivel B2 tuvieron una mejor percepción de sus clases en los ítems: aprovechar descansos, adaptación de contenido y comentarios. Se concluye que el grado de calificación influye en la percepción de las clases de Educación Física, y que en el caso de los sujetos más jóvenes existen elementos relacionados con la variable eficacia individual, donde no se encontraron diferencias según el grado de calificación o incluso estos. son positivos a favor del nivel B2.

**Palabras clave:** evaluación, bilingüismo, inglés, calidad, EF.

## 1. INTRODUCTION

Bilingual Spanish-English teaching using the Content and Language Integrated Learning (CLIL) methodology has become the dominant approach in the Spanish education system (Gil-López et al., 2021; Salvador-García et al., 2018).

The CLIL approach is understood from a dual perspective (Coyle et al., 2010) in which communication and content are integrated. In other words, it aims to teach non-linguistic content or a specific subject, such as physical education (PE) through a vehicular language, in most cases English. However, examples can also be found in other languages such as French and Portuguese, to a greater extent in the Autonomous Communities bordering these two countries, and in German in a smaller number of schools. The growing interest in this bilingual model stems from the need to improve students' language performance (EF EPI, 2020).

In this way, subjects such as PE assume an important role in this process of “transforming monolingual education systems into bilingual and multilingual ones” (Doiz & Lasagabaster, 2017, p.1), thus becoming one of the subjects chosen to be taught in a foreign language (Gil-López et al., 2021; Sánchez-Vegas et al., 2022). The choice of this subject is determined by its suitability, as it is one with an important experiential character which makes it a privileged and extraordinary setting for learning other languages (González et al., 2021).

As is evident, in order to become an ideal space for learning other languages (L2), correct didactic planning is necessary to produce CLIL sessions that allow for the balanced learning of subject content through another language (Coral & Lleixà, 2013; Martínez-Hita & García-Cantó, 2017; Martínez-Hita et al., 2023). However, there is currently a lack of pedagogical guidelines for the development of this bilingual approach in general, and in PE in particular (García-Calvo & Nieto, 2022), as well as a lack of research on the results obtained from the implementation of these educational approaches (Chiva-Bartoll et al., 2015; Martínez-Hita et al., 2022). Along the same lines, the lack of appropriate teacher training is another of the needs raised by the scientific community (Barrios & Milla-Lara, 2020; Codó, 2020; Custodio & García, 2019; Pena & Pladevall-Ballester, 2020; Pérez-Cañado, 2015), such training being necessary to achieve the expected results of quality bilingual education. Thus, Nieto (2019) highlights that despite the high degree of student satisfaction with CLIL bilingual programmes, there are practical shortcomings in their implementation.

In order to improve the goals mentioned above, more demanding requirements have been progressively introduced in order to be able to teach this type of education, such as requiring teachers to have a language level of C1 instead of B2, or requesting additional specific training in CLIL methodology, as is the case in the Autonomous Community of the Canary Islands (European Commission Eurydice, 2017). Recently Martínez-Hita (2022) found substantial improvements in the development of CLIL approaches and motor engagement time in PE classes following specific teacher training in CLIL for the organisation of PE sessions.

Despite the need for teacher training in language skills, there is still no consensus among teachers on the challenges that bilingual programmes may pose (Lozano-Martínez, 2017). Continuous teacher training must be carried out in conjunction with training in methodological questions and this requires collaboration, time, commitment, and motivation on the part of teachers and the public administrations involved (Doiz et al., 2019; González-Caballero, Cascales & Gomariz, 2022; Pavón & Méndez, 2017). Problems arise when the implemen-

tation of these programmes is rushed and does not allow time for the slow structures of training to adapt to the continuous changes that occur (Jover et al., 2016). For Escobar (2013), achieving quality bilingual education for teachers becomes one of the fundamental pillars of their success.

This research aims to discover the perception of teachers who teach PE CLIL in bilingual schools in terms of their recognised language level. At the same time, it seeks to differentiate the results obtained according to the age and gender of the participants.

It is hypothesized that teachers who have a C1 level of English will have a greater facility for conducting classes. In turn, the age of the participants may influence, with the youngest being the ones who find it easier to teach, without differences based on gender.

## 2. METHOD

### 2.1. Design

This study is a cross-sectional, quantitative study (Montero & León, 2002; Tashakkori & Teddie, 2010). To analyse the variables, a multiple-choice questionnaire was administered to the teachers. The study was conducted following the guidelines of the Declaration of Helsinki (World Medical Association, 2013) and approved by the Ethics Committee of the University of Murcia (4447/2023).

#### 2.1.1. Participants

The sample consisted of 146 PE teachers who taught English in different schools in the Region of Murcia, Aragon, the Canary Islands and Castile-La Mancha, teaching specifically at the Primary Education, Secondary Education or University Higher Education levels, with 41.8% being 36 years old or younger and 58.2% being 37 years old or older. The sample consisted of a total of 99 participants with a C1 level (67.8%) and 47 (32.2%) with a B2 level of English. In terms of gender distribution, 87 were male (59.6%) and 59 were female (40.4%). These data are reflected in Table 1:

**Table 1.** *Sociodemographic characteristics of the participants*

		<i>N</i>	<i>%</i>
Sex	Men	87	59.6%
	Women	59	40.4%
Age	36 or less	61	41.8%
	37 or more	85	58.2%
Education level	Primary education	53	36.3%
	Secondary education	68	46.6%
	University higher education	25	17.1%
English level	C1	99	67.8%
	B2	47	32.2%

## 2.2. Instruments

We used a multiple-choice questionnaire composed of a series of sociodemographic and professional variables as shown in Table 1 and the CESEFA evaluation instrument (<https://bit.ly/3qjIkCN>), developed and validated by Martínez-Hita et al. (2022). This questionnaire consisted of 15 items with a 4-point Likert scale, where 1 meant totally unfulfilled and 4 meant excellently fulfilled based on the opinion about the suitability of the specified criteria when teaching the classes, finding a total of 15 items that refer to the linguistic competence and the development of the classes of the PE teacher (Table 2), Cronbach's alpha values were.  $\alpha = .872$  for all 15 items. In the study of Martínez-Hita et al. (2022), items were linked in three categories: "Teacher-student interaction: attitude towards linguistic adaptation" (item 2, 3, 4, 5, 6, 7); "Organization and planning: Attitude towards linguistic knowledge in lesson planning" (item 1, 10, 11, 12, 14); "Individual efficacy, Attitude to teach PE through the L2" (item 8, 9, 13, 15). In order to describe each of the items more specifically, in the present study it was decided to analyze all the items separately and previously reliability values of three scales were checked and they were higher than .70.

**Table 2.** *CESEFA Questionnaire items.*

Item 1	Adequate_ Organisation	Use organisational scaffolding: routines that provide safety. Warming up or calming down. Taking the opportunity to present or evaluate what has been learnt in the L2 session.
Item 2	Adequate_ Climate	There is a fluid and open atmosphere in which learners are encouraged and stimulated by the teacher to speak in the L2.
Item 3	Chunk_ Use	Pupils make use of chunks or simple phrases for routine or day to day matters.
Item 4	Language_ Use	The teacher primarily uses the foreign language to convey information to students but may use L1 when necessary.
Item 5	Adapt_ message	Adapt the message, speed and intonation of speech to a slower pace with clear pronunciation, emphasising key words.
Item 6	Speech_ Redundancy	It incorporates redundancy in its design (rephrasing), use of synonyms, anglicisms and cognates, providing a correct and lengthy model.
Item 7	Demonstration	Use demonstration, body language or visual examples to facilitate or emphasize understanding of information.
Item 8	Feedback	The teacher provides concurrent, positive and continuous feedback to check that students have understood the proposal.
Item 9	Content_ Appropriateness	The methodological approaches of the area of PE are followed without the introduction of L2 detracting from them.

Item 10	Balance_Tasks_variable4Cs	Tasks with a balanced profile and/or are contextualized within the framework of the 4Cs are set out.
Item 11	Interdisciplinarity	There is interdisciplinarity and continuous communication with the L2 teacher.
Item 12	L2 and PE Content	There is planning in addition to the content PE subject in terms of grammar, vocabulary, use of scaffolding.
Item 13	Taking advantage of pauses	In the case of task and/or games with a high degree of motor demand, the pauses are used for linguistic development which is linked to the development of the content.
Item 14	Reading	The teacher provides material to facilitate the transition from reading to writing in L2.
Item 15	Cooperation	Cooperative learning is encouraged. The methodology emphasizes the use of activities that promote linguistic competence and communication among learners.

### 2.3. Procedure

Firstly, contact was made with the different centers from Primary Education, Secondary Education and University Higher Education, indicating the purpose of the study. Next, once the management team was informed, a Google Forms questionnaire was administered to different Bilingual Physical Education teachers. In this questionnaire, we included a brief video explanation with the different items in order to avoid misunderstandings in the different answers.

The questionnaire consisted of a total of four initial questions (email, sex, date of birth and level of English), together with the questionnaire called CESEFA that contained the 15 items that were analyzed (table 2). In this questionnaire, the guidelines for its completion were explained, indicating that it would only take about five minutes to complete, explaining the objectives of the research, and allowing them to opt out if they did not agree with the aims of the research. Since the questionnaire was carried out online, the researchers committed to the center, to resolve any questions that may exist prior to being answered.

Once the questionnaires were passed, the total number of responses was verified and the different participating centers were informed if they wanted to know the results of the study and the conclusions once drawn up.

### 2.4. Data analysis

First, a reliability analysis of the scale under assessment was performed and the Mahalanobis distance was used in order to detect and eliminate outliers or subjects who did not follow a logical pattern in the set of variables. In addition, the skewness and kurtosis values ( $>3$  or  $>10$  respectively) were analysed to check the normality of these data, and then the use of non-parametric procedures was tested using the one-sample Kolmogorov-Smirnov (KS) test for each of the items analysed.

Subsequently, a bivariate correlation analysis was carried out for the different items under study in order to check the relationship between them. After this procedure, the Mann-Whitney U test was used to check for differences according to the teachers' level of proficiency (level C1 or B2) in the variables, and the effect size was calculated using the d-Cohen formula (Cohen, 1988). In order to test for differences according to the gender and age range of the participants in contrast to the level of proficiency, a multivariate analysis (MANOVA) was performed by analysing the various univariate tests (ANOVAs). The statistical analyses were performed using the IBM SPSS 23.0 package.

### 3. RESULTS

#### 3.1. Descriptive and correlation analysis

Table 1 shows the descriptive analysis and correlations of the different variables under study, and the values of skewness and kurtosis show the descriptive analysis and correlations of the different variables under study. It should be noted that the values of skewness and kurtosis showed adequate values (<3 skewness and <10 kurtosis). In their turn, the correlations were significant at  $p < 0.01$  or  $p < 0.05$  in most of the items, especially the item adequacy\_organisation, which correlated significantly with all the variables except with 2L and PE content (item 12) and with cooperation (item 15).

**Table 3.** Correlation and descriptive analysis

	M	SD	S	K	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1 Adequate_organisation	3.425	0.778	-1.262	0.998	.289**	.226**	0.151	.173*	.264**	0.161	.405**	.363**	.479**	.381**	.460**	.375**	.330**	.307**
2 Adequate_climate	3.466	0.832	-1.494	1.375	1	.663**	.719**	.724**	.781**	.659**	0.148	0.136	.337**	.352**	.323**	0.103	.216**	0.093
3 Chunk_use	3.130	0.977	-0.849	-0.37		1	.562**	.760**	.650**	.662**	.174*	.168*	.206*	.196*	.200*	0.111	.163*	0.161
4 Language_use	3.432	0.838	-1.385	1.067			1	.612**	.758**	.684**	.196*	0.132	.206*	.264**	.260**	-0.012	.234**	.178*
5 Adapt_message	3.404	0.922	-1.429	0.932				1	.694**	.749**	0.158	0.091	.251**	.275**	.243**	0.051	.191*	0.154
6 Speech_redundancy	3.411	0.892	-1.442	1.112					1	.717**	.250**	.224**	.256**	.219**	.257**	0.13	.176*	.216**
7 Demonstration	3.438	0.871	-1.492	1.299						1	.246**	0.078	0.154	0.156	.178*	0.025	0.09	0.144
8 Feedback	3.493	0.754	-1.394	1.24							1	.382**	0.141	0.118	.168*	.345**	0.113	.610**
9 Content_appropriateness	3.541	0.744	-1.671	2.345								1	0.158	0.098	.217**	.664**	.172*	.470**
10 Balanced_Tasks_4Cs	3.390	0.773	-1.077	0.423									1	.535**	.672**	.242**	.469**	0.123
11 Interdisciplinarity	3.404	0.766	-1.117	0.56										1	.592**	0.134	.476**	0.136
12 L2 and PE content	3.418	0.812	-1.3	0.955											1	.263**	.552**	.185*
13 TAOP	3.541	0.744	-1.671	2.345												1	.172*	.331**
14 Reading	3.397	0.843	-1.214	0.5													1	0.085
15 Cooperation	3.541	0.735	-1.572	1.872														1

Legend: TAOP; Taking advantages of pauses; M = mean; SD = standard deviation; S = skewness; K = kurtosis; \*  $p < 0.05$ ; \*\*  $p < 0.01$

### 3.2. Differences according to the level of empowerment of the participants and the items under study.

Analysis using the Mann-Whitney U-test indicated statistically significant differences in all variables in favour of the group with a C1 level. Therefore, it was decided to use the d-Cohen effect size for a more detailed analysis of the results. Thus, an effect size was considered to be moderate for the variables: Language\_use1 (0.76). Feedback (0.43). Content\_appropriateness (0.44). Taking advantaged of pauses (0.45), Reading (0.59) and Cooperation (0.52) of these, Feedback ( $Z = -2.488$ ) and Content\_Appropriateness ( $Z = 2.305$ ) being significant at  $p$ -value  $< 0.05$  and the rest at  $p$ -level  $< 0.001$ .

**Table 4.** Analysis of results based on the qualification degree.

	C1 Level		B2 Level		Z	p	d
	M	SD	M	SD			
Adequate_organisation	3.69	0.51	2.87	0.95	-5.473	.000**	1.01
Adequate_climate	3.72	0.62	2.94	0.96	-5.503	.000**	0.92
Chunk_use	3.41	0.82	2.53	1.02	-5.064	.000**	0.93
Language_use	3.65	0.66	2.98	0.99	-4.416	.000**	0.76
Adapt_message	3.68	0.7	2.83	1.07	-5.359	.000**	0.90
Speech_redundancy	3.67	0.67	2.87	1.06	-4.970	.000**	0.86
Demonstration	3.7	0.63	2.89	1.05	-5.159	.000**	0.88
Feedback	3.61	0.65	3.26	0.9	-2.488	.013*	0.43
Content_appropriateness	3.66	0.61	3.3	0.93	-2.305	.021*	0.44
Balanced_Tasks_4Cs	3.62	0.6	2.91	0.88	-4.965	.000**	0.90
Interdisciplinarity	3.63	0.6	2.94	0.87	-4.988	.000**	0.88
L2 and PE content	3.67	0.59	2.89	0.96	-5.171	.000**	0.93
Taking advantaged of pauses	3.66	0.66	3.3	0.86	-2.861	.005**	0.45
Reading	3.57	0.7	3.04	1	-3.272	.001**	0.59
Cooperation	3.68	0.59	3.26	0.92	-2.940	.003**	0.52

Legend: M = mean; SD = standard deviation; \*\*  $p < 0.01$ ; \*  $p < 0.05$ ; d = d-Cohen effect size

### 3.3. Differential analysis taking into account the language level and gender of participants.

A multivariate analysis (MANOVA) was carried out in order to find out the differences according to the gender and language level of the participants. In this regard, the results showed statistically significant differences at  $p < 0.01$  in the Balanced Tasks variable 4CSs ( $F = 12.274$ ), 2L and PE content ( $F = 7.13$ ) and at  $p < 0.05$  in the reading variable ( $F =$

0.032). Specifically, the balanced tasks variable obtained a higher value in females in the case of level C1 (M = 3.78 vs. M = 3.5) and the opposite for level B2 (M = 2.56 vs. 3.14). For the variable 2L and PE content, the same occurred with higher values for women at the C1 level (M = 3.8 vs M = 3.57) and lower values for women at the B2 level (M = 2.61 vs M = 3.07). Finally, for the variable Reading, the values were M = 3.68 for women at level C1 and M = 2.78 at level B2, while for men the values were M = 3.48 and M = 3.21 respectively.

**Table 5.** *Analysis of results based on the degree of qualification, differentiating according to gender.*

	C1 Level				B2 Level				F	p	Eta <sup>2</sup>
	Men		Women		Men		Women				
	M	SD	M	SD	M	SD	M	SD			
Adequate_organisation	3.67	0.51	3.71	0.51	2.97	0.94	2.72	0.96	1.268	0.262	0.009
Adequate_climate	3.71	0.56	3.73	0.71	2.9	1.01	3.00	0.91	0.083	0.774	0.001
Chunk_use	3.45	0.75	3.37	0.92	2.45	0.99	2.67	1.08	0.863	0.354	0.006
Language_use	3.66	0.66	3.63	0.66	2.83	1.07	3.22	0.81	2.17	0.143	0.015
Adapt_message	3.69	0.63	3.66	0.79	2.79	1.05	2.89	1.13	0.173	0.678	0.001
Speech_redundancy	3.62	0.64	3.73	0.71	2.76	1.06	3.06	1.06	0.397	0.53	0.003
Demonstration	3.78	0.46	3.59	0.81	2.86	1.06	2.94	1.06	0.908	0.342	0.006
Feedback	3.6	0.62	3.61	0.7	3.24	0.91	3.28	0.89	0.012	0.911	0.000
Content_appropriateness	3.59	0.65	3.76	0.54	3.31	0.85	3.28	1.07	0.586	0.445	0.004
Balanced_Tasks_4Cs	3.5	0.66	3.78	0.47	3.14	0.69	2.56	1.04	12.274	0.001**	0.08
Interdisciplinarity	3.62	0.59	3.63	0.62	3.07	0.75	2.72	1.02	2.038	0.156	0.014
L2 and PE content	3.57	0.68	3.8	0.4	3.07	0.8	2.61	1.14	7.13	0.008**	0.048
Taking advantaged of pauses	3.53	0.75	3.83	0.44	3.38	0.68	3.17	1.1	3.775	0.054	0.026
Reading	3.48	0.75	3.68	0.61	3.21	0.9	2.78	1.11	4.681	0.032*	0.032
Cooperation	3.66	0.58	3.71	0.6	3.24	0.99	3.28	0.83	0.004	0.952	0.000

Box M = 718.883, F = 1.503, Sig = 0.000; Wilk's Lambda = F= 1.236; p = .253;

Legend: M = mean; SD = standard deviation; \*\* p < 0.01; \* p < 0.05; eta2 = partial eta squared

### 3.4. Differential analysis taking into account the language level and age range of participants.

Finally, a multivariate analysis (MANOVA) was carried out in order to find out the differences according to the participants' age and skills. In this regard, the results showed statistically significant differences at  $p < 0.05$  in the variables Feedback ( $F = 5.61$ ), Content\_appropriateness ( $F = 5.977$ ), 2L and PE content ( $F = 3.988$ ), Taking advantage of pauses ( $F = 6.358$ ), Reading ( $F = 4.467$ ) and Cooperation ( $F = 4.17$ ). More specifically, we contrast how those under 36 years of age show statistically significant results with higher values at the B2 level in the Feedback variable ( $M = 3.67$  vs  $M = 3.06$ ), Content\_appropriateness ( $M = 3.73$  vs  $M = 3.09$ ), 2L and PE content ( $M = 3.27$  vs  $M = 2.72$ ), Taking advantage of pauses ( $M = 3.67$  vs  $M = 3.13$ ), Reading ( $M = 3.33$  vs  $2.91$ ) and Cooperation ( $M = 3.53$  VS  $M = 3.12$ ).

**Table 6.** Analysis of results based on the degree of qualification, differentiating according to the age range of the participants

	C1 Level				B2 Level				F	p	Eta <sup>2</sup>
	36 or less		37 or more		36 or less		37 or more				
	M	SD	M	SD	M	SD	M	SD			
Adequate_organisation	3.65	0.48	3.72	0.53	3	0.93	2.81	0.97	0.988	0.322	0.007
Adequate_climate	3.76	0.67	3.68	0.58	3.13	0.83	2.84	1.02	0.555	0.457	0.004
Chunk_use	3.43	0.78	3.4	0.86	2.53	1.13	2.53	0.98	0.012	0.913	0
Language_use1	3.63	0.77	3.66	0.55	3.13	0.83	2.91	1.06	0.778	0.379	0.005
Adapt_message	3.67	0.73	3.68	0.67	2.8	1.08	2.84	1.08	0.015	0.902	0
Speech_redundancy	3.67	0.79	3.66	0.55	3.07	0.88	2.78	1.13	0.803	0.372	0.006
Demonstration	3.7	0.66	3.7	0.61	2.73	1.03	2.97	1.06	0.626	0.43	0.004
Feedback	3.59	0.62	3.62	0.69	3.67	0.62	3.06	0.95	5.61	0.019*	0.038
Content_appropriateness	3.65	0.57	3.66	0.65	3.73	0.59	3.09	1	5.977	0.016*	0.04
Balanced_Tasks_4Cs	3.61	0.58	3.62	0.63	3.2	0.86	2.78	0.87	2.77	0.098	0.019
Interdisciplinarity	3.61	0.61	3.64	0.59	3.2	0.94	2.81	0.82	2.647	0.106	0.018
L2 y contenidos EF	3.67	0.56	3.66	0.62	3.27	0.96	2.72	0.92	3.988	0.048*	0.027
Taking advantaged of pauses	3.59	0.72	3.72	0.6	3.67	0.62	3.13	0.91	6.358	0.013*	0.043
Reading	3.46	0.78	3.66	0.62	3.33	0.98	2.91	1	4.467	0.036*	0.03
Cooperation	3.61	0.61	3.74	0.56	3.53	0.74	3.12	0.98	4.17	0.043*	0.029

M de box = 444.387, F = 1.523, Sig = 0.000; Lamda de Wilks = F= 1.223; p = .263

Legend: M = mean; SD = standard deviation; \*\* p < 0.01; \* p < 0.05; eta<sup>2</sup> = partial eta squared

## 4. DISCUSSION

The aim of the present study was to explore the perception of teachers who teach CLIL in bilingual schools in Spain based on their recognised language level. In this respect, it should be noted that around twenty years have passed since the first bilingual programmes were introduced, but despite this, research into them has not been as rapid as their expansion in the Spanish education system. Another added problem is the number of different qualifications that exist in the Common European Framework of Reference for Languages (CEFR) that accredit the level of English, which further disperses the knowledge of the language. Thus, Pavón (2018) mentions that there is still a lack of a solid knowledge base on the effects of this approach in the different subjects where a foreign language is introduced as a vehicle for teaching, given that, in many cases, they have been implemented without really knowing what the consequences of these could be (Martínez-Hita & García-Cantó, 2017).

In our case, taking into account the level of language accredited, we have been able to verify how practically unanimously, having a higher level of knowledge of the language (C1 vs B2), allowed for more positive results in all the items of the CESEFA questionnaire (Martínez-Hita et al. 2022). This is something that could be suggested by authors such as Mosquera (2017) when indicating that teachers who teach CLIL should have a high level of English and adequate training in embedded methodologies in order to be able to deliver bilingual teaching programmes with quality. This second suggestion could be taken into account for future studies in which bilingual PE teaching is applied, using the active methodologies that are so common in this field.

Looking at the differences according to the items, the study by Martínez-Hita et al. (2022) classifies this section into three categories, which we will describe according to the results obtained. The items referring to teacher-student interaction all had effect sizes considered to be large ( $>0.80$ ), which indicated that the greatest differences were in this variable. This means that the aspects where most differences existed were those where communication with learners (climate, use of chunks, use of language, intonation and redundancy of discourse) were items that were to a large extent, difficult for B2 participants to focus on probably because these are items where the dependence is exclusive on having a wide and varied second language level, probably because these items allow the complexity of the teaching-learning process to be assessed during interaction and improvisation. It is necessary to remember that it is not about doing the same thing in another language, in this dual approach language becomes a means to an end and not an end in itself, in other words, it is a tool and not an end goal (Deller & Price, 2007).

Secondly, regarding the items referring to organisation and planning, these, generally had a medium effect size, except for item 1, which had the largest effect size (adequate organisation). This shows how the organisational aspect had an intermediate difficulty level, which may be due to the possibility of prior planning, something that did not exist on occasions where interaction with students was immediate (without the possibility of anticipating the response) as in the previous category. Adequate organisation, however, was where there were most differences, perhaps because it is a more global item which encompasses all the methodological/organisational procedures of CLIL EF teaching.

Finally, the third category, 'individual effectiveness' in teaching EF in a second language, was the only one in which all items had a small or medium effect size. Here it may

be considered less important to have a higher level of English, as it includes items such as feedback or the use of pauses (where 'fixed phrases' can be found) or the appropriateness of content and the use of cooperative learning (which can be pre-planned).

It is not surprising that authors such as Pérez et al. (2016) point to the existing deficit in CLIL research, especially because of the difficulty of establishing common guidelines for planning, implementing, and evaluating programmes at the national level. In our study, we did not carry out an evaluation as such, but we did find that, in both men and women, having a lower level of English meant greater difficulties when it came to teaching. This is even more important if we take into account that teachers are increasingly aware of the importance of really integrating the second language with the curricular content of the subject (Durán-Martínez, et al. 2016).

As for the differences found according to the age of the participants, it should be noted that those who were under 36 years of age had significant and higher values in the category of content appropriateness (level B2). These participants with level B2 had higher values even than those with level C1 in both ages in this variable, something similar to what happened with feedback and the use of pauses. In addition, the results showed higher values for cooperation, reading and the use of the second language with the EF content.

In terms of gender, significant differences were only found at B2 level with the variable of balanced tasks 4Cs and 2L and PE content and reading, where men had higher values than women. On the other hand, in these same variables at level C1, it was women who had statistically higher values than men, although they were always higher in the case of level C1 than B2. In summary, there were some differences according to gender, but in no case did participants with level B2 have a higher perception than in the case of level C1.

As the main limitations of the study, it should be noted that the experience that teachers have in Physical Education could have been considered as a possible variant. In addition, the results where in some cases the B2 level was better than the C1 level, does not have an objective explanation. Finally, it could have been considered to analyze other variables such as satisfaction with the classes or to have given questionnaires to the students to find out their perception.

Finally, as future lines of work, it is recommended to expand the study sample, in addition to attending to the possible limitations specified. On the other hand to carry out these investigations in other subjects as well, it would be interesting to compare the results and see if they are replicated.

## 5. CONCLUSIONS

The level of language proficiency of the PE teachers who teach the subject in English has a specific influence on their ability to carry out their teaching. This is particularly the case when teachers have to work in the classroom without being able to plan in advance for aspects of bilingual teaching. At the same time, it is important to take into consideration the age of the teachers, since those under 36 years of age with a B2 level had similar or even higher values than those with a C1 level in certain categories related to the variable called "individual effectiveness", such as the appropriateness of the content of the classes, the use

of feedback or the use of breaks. In terms of gender, the differences were minimal, with both men and women at C1 level having higher values than those at B2 level.

**Note:** The results of this publication have been obtained from the treatment of the data of the Doctoral Thesis of Francisco José Martínez Hita entitled “Evaluation of bilingual teaching in CLIL Physical Education classes and its impact on motor engagement time”: [https://digitum.um.es/digitum/bitstream/10201/122883/1/Franciscojose.martinez.hita\\_sin.articulos.pdf](https://digitum.um.es/digitum/bitstream/10201/122883/1/Franciscojose.martinez.hita_sin.articulos.pdf)

## 6. REFERENCES

- Barrios, E. & Milla-Lara, M. D. (2020). CLIL methodology, materials and resources, and assessment in a monolingual context: an analysis of stakeholders’ perceptions in Andalusia, *The Language Learning Journal*, 48(1), 60-80. <https://doi.org/10.1080/09571736.2018.1544269>
- Chiva-Bartoll, O., Isidori, E., & Fazio, A. (2015). Bilingual physical education and critical pedagogy: an application based on Judo. *Retos*, 28, 110-115. <https://doi.org/10.47197/retos.v0i28.34812>
- Codó, E. (2020). The dilemmas of experimental CLIL in Catalonia. *Journal of Multilingual and Multicultural Development*, 43(4), 341-357. <https://doi.org/10.1080/01434632.2020.1725525>
- Cohen, J. (1988). *Statistical power analysis for the behavioural sciences*. Lawrence Erlbaum Associates Publishers.
- Coral, J. & Lleixà, T. (2013). Physical Education-in-CLIL tasks. Determining tasks characteristics through the analysis of the diaries. *Retos*, 24, 79-84. <https://doi.org/10.47197/retos.v0i24.34532>
- Coyle D., Hood P., & Marsh D. (2010). *Content and Language Integrated Learning*. Cambridge University Press
- Custodio, M. & García, J. M. (2019). Measuring teachers’ competence to plan CLIL and diagnosis teachers’ training needs. *Bordón*, 72(1), 31-48. <https://doi.org/10.13042/Bordon.2019.72250>
- Deller, S., & Price, C. (2007). *Teaching other subjects through english (CLIL)*. Oxford university press.
- Doiz, A., & Lasagabaster, D. (2017). Management teams and teaching staff: do they share the same beliefs about obligatory CLIL programmes and the use of the L1? *Language and Education*, 31(2), 93-109. <https://doi.org/10.1080/09500782.2017.1290102>
- Doiz, A., Lasagabaster, D., & Pavón, V. (2019). The integration of language and content in English-medium instruction courses: lecturers’ beliefs and practices. *Iberica: Revista de la Asociación Europea de Lenguas para fines Específicos (AELFE)*, 38, 151-176.
- Durán-Martínez, R., Beltrán-Llavador, F., & Martínez-Abad, F. (2016). A contrastive análisis between novice and expert teachers’ perceptions of school bilingual programmes. *Cultura y Educación*, 28(4), 738-770. <https://doi.org/10.1080/11356405.2016.1237339>
- EF EPI (2020). EF Epi, índice EF de nivel de inglés [EF English proficiency index]. [https://www.ef.com.es/assetscdn/WIBIwq6RdJvcD9bc8RMd/legacy/\\_/~/media/centralefcom/epi/downloads/full-reports/v10/ef-epi-2020-spanish-euro.pdf](https://www.ef.com.es/assetscdn/WIBIwq6RdJvcD9bc8RMd/legacy/_/~/media/centralefcom/epi/downloads/full-reports/v10/ef-epi-2020-spanish-euro.pdf)
- Escobar, C. (2013). Learning to become a CLIL teacher: teaching, reflection and professional development. *International Journal of Bilingual education and Bilingualism*, 16(3), 334-353. <https://doi.org/10.1080/13670050.2013.777389>

- European Commission Eurydice (2017). *Cifras clave de la enseñanza de lenguas en los centros escolares de Europa, edición 2017* [Key data on language teaching in schools in Europe, 2017 edition]. Oficina de Publicaciones de la Unión Europea. <https://sede.educacion.gob.es/publiventa/cifras-clave-de-la-ensenanza-de-lenguas-en-los-centros-escolares-de-europa-edicion-2017-informe-eurydice/ensenanza-lenguas-europa/22137>
- García-Calvo Rojo, S., & Nieto Moreno de Diezmas, E. L. (2022). Bilingual Physical Education-CLIL through teaching games for understanding: comparative analysis of intrinsic satisfaction by age and gender. *Retos*, 46, 458–466. <https://doi.org/10.47197/retos.v46.91685>
- Gil-López, V., González Villora, S., & Hortigüela-Alcalá, D. (2021). Learning foreign languages through content and language integrated learning in physical education: A systematic review. *Porta Linguarum*, 35, 165-182. <http://dx.doi.org/10.30827/portalin.v0i35.15785>
- González, H., Salazar, P. A., & Villota, J. A. (2021). Learning English as a foreign language through sports practices. *Revista Boletín Redipe*, 10 (4), 347-375. <https://doi.org/10.36260/rbr.v10i4.1276>
- González-Caballero, F., Cascales, A., & Gomariz, M.A. (2022). Las estrategias neurodidácticas en el área de lengua ingles en Educación Infantil y Primaria. *Espiral. Cuadernos del Profesorado*. 15(31), 43-56. <https://doi.org/10.25115/ecp.v15i31.7217>
- Jover, G., Fleta, T., & González, R. (2016). Pre-service education of primary school teachers in the context of foreign language bilingual teaching. *Bordón. Revista de Pedagogía*, 68(2), 120-135. <https://doi.org/10.13042/Bordon.2016.68208>
- Lozano-Martínez, L. (2017). Teachers in bilingual programs in Cantabria. *Elia, Estudios de Lingüística Inglesa Aplicada*, 17, 93-123. <https://doi.org/10.12795/elia.2017.i17.05>
- Martínez-Hita, F. J. (2022). Specific training in CLIL in Physical Education to improve motor engagement time. *Sportis. Scientific Journal of School Sport, Physical Education and Psychomotricity*, 8(2), 227-238. <https://doi.org/10.17979/sportis.2022.8.2.8712>
- Martínez-Hita, F. J., & García-Cantó, E. (2017). Influence of bilingualism on engagement motor time in Physical Education. *Retos*, 32, 178–182. <https://doi.org/10.47197/retos.v0i32.51805>
- Martínez-Hita, F. J., García-Cantó, E., Gómez-López, M. & Granero-Gallegos, A. (2023). Influence of clil bilingual physical education on the engagement motor time. *E-balonmano.com: Journal Sports Science*, 19(1), 55-66. <https://doi.org/10.17398/1885-7019.19.5>
- Martínez-Hita, F. J., Granero-Gallegos, A., & Gómez-López, M. (2022). Design and validation of a tool to evaluate CLIL in Physical Education sessions. *Porta Linguarum*, 37, 193-210. <https://doi.org/10.30827/portalin.vi37.17795>
- Montero, I. & León, O. (2002). Classification and description of research methodologies in Psychology. *International Journal of Clinical and Health Psychology*, 2, 503-508.
- Mosquera, I. (2017). Enseñanza bilingüe y formación del profesorado: dos caras de la misma moneda [Bilingual education and teacher training: two sides of the same coin]. *Unir Revista Online*. <https://www.unir.net/educacion/revista/ensenanza-bilingue-y-formacion-del-profesorado-dos-caras-de-la-misma-moneda/>
- Nieto, E. (2019). Students, teachers and management teams in bilingual programmes. *Journal of English Studies*, 17, 277-297. <https://doi.org/10.18172/jes.3564>
- Pavón, V. (2018). Learning Outcomes in CLIL Programmes: a Comparison of results between urban and rural environments. *Porta Linguarum*, 29, 9–28. <http://dx.doi.org/10.30827/Digibug.54020>
- Pavón, V. & Méndez, M. C. (2017). Analysing teachers' roles regarding cross-curricular coordination in Content and Language Integrated Learning (CLIL). *Journal of English Studies*, 15, 235-260. <http://doi.org/10.18172/jes.3227>

- Pena, C. & Pladevall-Ballester, E. (2020). Effects of focus on form on primary CLIL students' foreign language performance in task-based oral interaction. *Journal of Immersion and Content-Based Language Education*, 8, 53-79. <https://doi.org/10.1075/jicb.17028.alf>
- Pérez-Cañado, M. L. (2015). Evaluating CLIL programmes: instrument design and validation. *Revista de Educación Pulso*, 39, 79-112. <http://hdl.handle.net/10017/28338>
- Pérez, A., Lorenzo, F. & Pavón, V. (2016). European bilingual models beyond lingua franca: key findings from CLIL French programs. *Language Policy*, 15, 485–504. <https://doi.org/10.1007/s10993-015-9386-7>
- Salvador-García, C., Chiva-Bartoll, O., & Vergáz Gallego, J. J. (2018). Students' perception of the use of CLIL method in Physical Education: A Case Study. *Retos*, 33, 138-142. <https://doi.org/10.47197/retos.v0i33.53665>
- Sánchez-Vegas, V., Coterón, J., González-Peño, A., & Franco, E. (2022). Physical Education in Bilingual Schools: Effect of the language use on students' perceptions of teaching communication and motivation. *Retos*, 43, 352-360. <https://doi.org/10.47197/retos.v43i0.89269>
- Tashakkori, A., & Teddlie, C. (2010). *Handbook of mixed methods in social and behavioral research*. Sage: Thousand Oaks. <https://doi.org/10.4135/9781506335193>
- World Medical Association. (2013). World Medical Association Declaration of Helsinki: Ethical principles for medical research involving human subjects. *Jama*, 310(20), 2191–2194. <https://doi.org/10.1001/jama.2013.281053>