Systematic Review

English-Medium Instruction as a Pedagogical Strategy for the Sustainable Development of EFL Learners in the Chinese Context: A Meta-Analysis of Its Effectiveness

Jian-E Peng ¹ and Xiaowen (Serina) Xie ^{2,*}

¹ College of Liberal Arts, Shantou University, Shantou City 515063, China; pengjiane@stu.edu.cn

² Moray House School of Education and Sport, University of Edinburgh, Edinburgh EH8 8AQ, UK

* Correspondence: Serina.Xie@ed.ac.uk

Abstract: With English-medium instruction (EMI) as a pedagogical strategy being practiced world-wide in higher education (HE), extensive research has explored stakeholders' attitudes toward, and perceived benefits and challenges of EMI based on self-report data. However, the actual effectiveness of EMI on students' subject content and English language learning achievements tested with objective measures has accrued little evidence. This meta-analysis synthesized 44 independent samples (32 in medical disciplines) from 36 studies. The results show that EMI students performed significantly better in both subject content and English learning than students in Chinese-medium courses, but it should be noted that the difference in content learning was found only with students from medical disciplines. Discipline was the only significant factor moderating content learning, while disciplines, research design, and instruction time in English significantly moderated English learning. The findings provide implications for implementing EMI in similar contexts and highlight the importance of rigorous future research to examine the benefits of EMI.

Keywords: English-medium instruction; meta-analysis; effectiveness; subject content learning; English learning

Citation: Peng, J.-E; Xie, X. English-Medium Instruction as a Pedagogical Strategy for the Sustainable Development of EFL Learners in the Chinese Context: A Meta-Analysis of Its Effectiveness. *Sustainability* **2021**, *13*, 5637. <https://doi.org/10.3390/su13105637>

Academic Editors: Lawrence Zhang and Vincent T. Greenier

Received: 19 April 2021

Accepted: 11 May 2021

Published: 18 May 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Internationalization has been widely embraced in higher education institutions (HEIs), one collateral impact of which is “Englishization” in HEIs manifested in adopting English-medium instruction (EMI) as a pedagogical strategy [1]. Generally, EMI refers to the use of English to teach non-language academic subjects in “countries or jurisdictions where the first language of the majority of the population is not English” [2] (p. 37). It has often been likened to “content and language integrated learning” (CLIL) or “integrating content and language in higher education”, notions used in Europe. However, while both CLIL and EMI share the pedagogical objective of improving “students’ L2 proficiency by teaching subject matter through L2” [3] (p. 663), the “language component” is usually not the focus of teaching and assessment in EMI programs, and EMI instructors are usually not L2 specialists [4]. Despite the ostensibly straightforward definition of EMI, its actual practices vary greatly at geographical, institutional, and even classroom levels [2]. For instance, in the context of mainland China (henceforth referred to as China), the official term used to describe the adoption of English in non-language courses, partially or fully, is *shuangyu jiaoxue* (bilingual teaching) or *shuangyu jiaoyu* (bilingual education) [5]. However, bilingual teaching or education practices in China differ hugely from bilingual education models in international contexts [6]. While bilingual education in Western contexts offers lengthy exposure to two languages for various reasons, such as maintaining heritage language or promoting additive bilingualism, bilingual education in China mainly aims to promote content and language learning simultaneously through offering semester-based English-taught courses (for an in-depth discussion of the differences, see [6]). To be more

precise, EMI has been adopted as a pedagogical strategy for enhancing students' English proficiency without taxing their time and energy, and at a macro level, as a curricular strategy "intended to internationalize tertiary institutions and improve the quality of higher education" [7] (p. 1). EMI practice would range from a minimum use of spoken English to full English provision in class, which depends on instructors' strategic choices or institutional policies. Given the equivocal meanings attached to these terminologies, in this study, we followed Zhao and Dixon [5] and chose EMI as the cover term to refer to the pedagogical strategy of using English, to various extents, as the medium in teaching non-English academic subjects.

While EMI has been practiced and researched worldwide, there is still scarce evidence regarding the effectiveness of such a pedagogical strategy in promoting content learning, and, as expected by educators in contexts such as China, in enhancing English language learning. This may be partly due to the diversity of EMI programs and different research designs, which revealed different or even contradictory findings. Therefore, a meta-analysis of empirical findings regarding the effectiveness of EMI is of utter importance to inform future directions of EMI implementation. Since EMI practices diversify hugely across educational contexts, we here report a meta-analysis of the effectiveness of EMI in enhancing subject content and English learning achievements in China. More specifically, this meta-analysis focused on the context of mainland China while excluding Hong Kong, Macau, and Taiwan since in these regions, educational policies and practices have been historically different [8]. The promotion of EMI in China in the past two decades has rendered this context an epitome of worldwide efforts toward internationalization in higher education. Synthesizing findings on the effectiveness of EMI in this context can help educators and policy makers in many other similar contexts to critically evaluate this pedagogical strategy and the ways to improve it for sustaining EMI students' academic development, which should be the priority of the agenda of educational institutions that aspire to implement EMI.

2. Literature Review

2.1. Perceptions and Measured Effectiveness of EMI

Previous research on EMI has generated rich descriptions of learners' and teachers' perceptions of EMI, including their attitudes and perceived challenges and obstacles [9–14]. Generally, students' positive attitudes toward EMI have been reported. For instance, in their questionnaire survey and interviews conducted among 579 students at 12 HEIs and 28 staff members at 8 universities in Japan and China, Galloway et al. [10] reported that students mostly had a positive image about their EMI courses or programs, with nearly 90% of the students agreeing that EMI is appropriate at the university level. Notably, EMI was perceived to do a better job at improving their overall English language proficiency than their learning of subject knowledge. A similar positive result was reported in Kong and Wei's [13] survey of 282 Chinese students. In this study, the item "EMI helps improve my English language proficiency" received an endorsement from 70.3% of the respondents, with a mean score of 3.81 on a 5-point scale. In a recent systematic review on attitudes toward EMI in East Asia and the Gulf, Graham and Eslami [15] found that of the 20 studies included, 4 studies reported positive attitudes and 7 slightly positive attitudes, which accounted for 55% of the studies being reviewed. Teachers' positive attitudes toward EMI have also been documented. In Galloway et al.'s [10] research, faculty members believed that EMI was beneficial for students' overall language knowledge and subject knowledge. Twenty-five university teachers from Austria, Italy, and Poland in Dearden and Macaro's [9] interview study expressed positive attitudes toward EMI and believed that EMI could improve students' English.

However, despite the optimistic picture about EMI that has surfaced in the literature, many challenges and concerns have been raised, particularly in English-as-a-foreign-language (EFL) contexts. Ample criticism has been directed at the gaps between national-level (macro) and/or university-level (meso) EMI policies and classroom-level EMI prac-

tices (micro) [7,16,17]. Rose and colleagues' [16,17] analyses of government and institutional documents, as well as interviews with stakeholders (i.e., teachers and students), have shown that due to a lack of detailed stipulations or guidelines, policy enactment at the classroom level varied, which caused differing degrees of English provision in actual practice. Consequently, this could confuse faculty members and confound any intended efforts to examine the effectiveness of EMI as a pedagogical strategy.

Another repeatedly reported factor that affects the effectiveness of EMI is the English proficiency of both teachers and students [18–22]. Although instructors of content courses may have met the qualifications for EMI given their overseas academic training or other experiences, many of them are not sufficiently competent to use the English language at will. In Doiz et al.'s [19] focus group interviews with 13 EMI lecturers in Spain, the lecturers all mentioned that teaching in a foreign language was one of the biggest challenges, which resulted in a reduction of vocabulary at their disposal and reduced details in their explanations regarding the content subject. Due to inadequate English proficiency, EMI instructors may have to stay close to the textbooks and be unable to teach in "a spontaneous, interactive, freewheeling manner" [23] (p. 35). This may then reduce the transmission of lesson content and trigger cognitively lower-order student responses [24] and limit the variety of pedagogical activities otherwise available [25,26]. To worsen this situation, many EMI contexts lack sufficient and effective in-service training for EMI instructors to improve their pedagogical and language skills in managing their teaching [11,27].

Similarly, students' low English proficiency is a widely recognized obstacle that hinders them from comprehending lesson content [11,28]. For instance, Kim [29] found that 60% of the students understood less than 70% of class content in beginner classes in a Korean EMI course. This can result in consequences such as EMI teachers falling back on using students' first language (L1) as a learning resource [28,29], repeatedly explaining specific or technical vocabulary, or simplifying disciplinary content [11,19]. Consequently, EMI's initially assumed benefits for improving both English and content learning may not be materialized.

Based on the mixed perceptions of EMI presented above, a critical issue worth investigating is to what extent EMI effectively brings about both subject content and language learning achievements. Hitherto, many studies have focused on the effectiveness of EMI perceived by students. There is a dearth of studies concerning the objective measure of such effectiveness (see [21,22,30–34]), and inconclusive findings have been reported. For instance, while significant gains in English vocabulary, morphological awareness, and reading comprehension were identified in Li's [35] study of 53 students majoring in Early Childhood Education in one-semester EMI, when this group was compared to a "historical" group, the EMI students scored slightly lower than the "historical" group. In Lei and Hu's [21] study, no significant difference was found between 64 students in an EM program and 72 students in a Chinese-Medium (CM) program in their scores of the national standardized College English Test Band 6 (CET 6). Targeting students majoring in medical sciences, Yang et al. [30] reported no statistical difference between EMI students and CMI students on their scores on the Chinese Medical Practitioner Examination (CMPE) taken in both 2015 and 2016. Joe and Lee's [32] study among 64 Korean medical students, based on the comparison of the students' comprehension of a 50 min lecture in English and in Korean, respectively, showed that the medium of instruction had no effect on students' lecture comprehension, although more than half of the students perceived EMI as less satisfying than the Korean-medium instruction. These findings suggest that no consensus can be reached on the impact of EMI on students' subject content and English learning achievements.

Stronger evidence has been obtained from quasi-experimental research. For instance, Guo, Tong, Wang, Min, and Tang [36] conducted a quasi-experimental study among 18 EMI students and 25 CMI students in China and showed that while EMI students scored numerically higher on post-test of academic achievement and English language, the differences were not significant after adjusting for their performance in the pre-test. Similarly, in Han's [37] quasi-experimental study, 137 EMI students performed equally

well on an end-of-year mathematics test compared to their 137 non-EMI counterparts. These results seemed to have prompted researchers to dismiss the common concern documented in the literature as to whether EMI carries a detrimental effect on students' content area learning [32,36].

While mixed findings on the impact or effectiveness of EMI on language and subject content learning have arisen from individual studies, systematic reviews or meta-analyses have been conducted to provide a fuller understanding of the issues in question [2,15,38,39]. Macaro et al. [2] systematically reviewed 83 studies on EMI in higher education in different geographical regions and concluded that "the research evidence to date is insufficient to assert that EMI benefits language learning nor that it is clearly detrimental to content learning" (p. 36). In addition to Graham and Eslami's [15] systematic review on attitudes toward EMI among students from East Asia and the Gulf, as mentioned before, Tong et al. [39] systematically reviewed research on bilingual education in Chinese higher education published in English and Chinese. Tong et al.'s [39] study has attempted to explore what aspects of bilingual education in the Chinese context were explored and in what ways, as well as the publication trend. Although these reviews have contributed essential insights into EMI in different regions, they hardly employed quantitative techniques to precisely measure the effectiveness of EMI on promoting content and/or language learning. To this purpose, meta-analysis as a more powerful approach is highly needed. Meta-analysis is "a quantitative method of synthesizing empirical research results in the form of effect sizes" [40] (p. 7). In other words, the effect size (ES) computed in a meta-analysis can provide substantial evidence for assessing the existence of effects. Lo and Lo [38] conducted a meta-analysis in which they synthesized the results of 24 studies on EMI education in secondary schools in Hong Kong between 1970 and 2010. To the best of our knowledge, Lo and Lo's [38] study represented the first attempt to conduct a meta-analysis on EMI. However, the effectiveness of EMI programs at the tertiary level remains mostly unknown, which requires more endeavors to perform such a meta-analysis to shed light on this field.

2.2. EMI Practices in Tertiary Education in Mainland China

EMI practices in China have gained momentum since the turn of this century when the Ministry of Education of China [41] issued a directive stipulating that universities should strive to offer 5–10% of undergraduate courses in English or other foreign languages within three years. This top-down policy has been driven by multiple benefits assumed to be brought about by EMI. It is believed that EMI enables students to keep abreast with cutting-edge scientific and technologic achievements [7]. Additionally, the adoption of EMI is expected to raise the quality of higher education and thus increase the competitiveness of HEIs. At a fundamental level, EMI is generally believed to fulfil two goals with one shot, that is, enhancing both subject content and language learning [5].

Despite what the term EMI may suggest, actual practices of EMI in China are complicated. Hu [6] summarized four bilingual models: foreign language teaching in mainstream education, maintenance bilingual instruction, transitional bilingual instruction, and immersion bilingual instruction. The four models can be viewed as situated on a continuum between exclusive use of Chinese (i.e., foreign language teaching in mainstream education) and the predominant use of English (i.e., immersion bilingual instruction) in teaching curriculum content. In the case of maintenance bilingual instruction, Chinese is still used as the main medium of instruction, while English is used more frequently to provide explanations, descriptions, and illustrations; in contrast, the transitional bilingual instruction model is a reversal where English is more regularly used than Chinese [6]. A relatively simplified classification was proposed by Xu [42], who categorized three EMI models based on the amount of English and language used by instructors: Chinese as the main medium of instruction (CMMI), English and Chinese balanced instruction (ECBI), and English as the main medium (EMMI). In these three models, the proportion of English used is less than 30%, about 40–60%, and higher than 70%, respectively. Based on his review of the literature on EMI in China and his observations of EMI classes on the tertiary level, Xu [42]

noted that the textbooks, handouts, or PowerPoint slides used in EMI courses are often in English, and different models of EMI mainly differ in the oral language used.

Conforming to the international research landscape in EMI presented above, similar positive attitudes, concerns, and challenges perceived by students and instructors have been reported in many empirical studies of EMI in this local context (e.g., [7,11–14,42]). However, only a handful of studies have concentrated on examining the actual effectiveness of EMI (e.g., [36,37]), leaving this issue unresolved. Hence, this meta-analysis aims to address this issue by synthesizing studies that have statistically assessed students' subject content and/or English learning achievements in EMI courses compared to their counterparts in corresponding CMI courses. Additionally, we considered that the difference in learning achievements between EMI and CMI students may be moderated by academic disciplines, language provision in class, research designs, and even the degree of details disseminated in different types of work, such as unpublished theses and published articles in English (mostly in international journals) or Chinese (in local journals). Hence, we also examined the possible moderating effect of five variables: (a) discipline (medical disciplines, other hard disciplines, and soft disciplines); (b) research design (non-experimental, randomized controlled trial, and quasi-experimental); (c) instruction time in English (50% or above, below 50%, and unreported); (d) publication type (published articles and theses); and e) journal type (in English, in Chinese). Further information about the categories of these variables is presented in the next section.

3. The Study

3.1. Research Questions and Context

This study was guided by the following research questions:

1. Are there any significant differences in subject content and English learning achievements between students taking EMI and CMI courses?
2. To what extent do the variables of discipline, research design, instruction time in English, publication type, and journal type moderate the differences in subject content and English learning achievements between students taking EMI and CMI courses?

We originally classified discipline into a broad contrast between “hard” disciplines (e.g., medical science, physics) and “soft” disciplines (e.g., Mass Media, Education) in Becher and Trowler's [43] parlance. However, this variable was later categorized into medical disciplines, other hard disciplines, and soft disciplines, because 32 out of 44 independent samples from the collected studies were in the field of medical science, 8 in other hard disciplines (i.e., Physics, Mathematics, Computer Science, Chemistry-related), and 4 in soft disciplines (i.e., Advertising, Mass Media, Early Childhood Education, Business). Separating medical disciplines from other hard disciplines allowed us to perform fine-grained analyses on other hard disciplines (e.g., physics, mathematics). Referencing Tong et al. [39], we categorized research design into the non-experimental design, randomized controlled trial (i.e., experimental design), and quasi-experimental design. Although Xu's [42] above-mentioned framework for classifying English provision in teaching is useful, we found that many of the collected studies did not report the actual or approximate percentage of English used in teaching. Hence, instruction time in English was divided into three types: 50% or above of the instruction time, below 50% of the instruction time, and unreported. Publication type had two categories: published articles and unpublished theses. Journal type was classified into English and Chinese in terms of the language used.

3.2. Method

3.2.1. Literature Search

The literature search was performed using the keywords “medium of instruction”, “English medium instruction” or “bilingual program” and “China” or “Chinese”, targeting studies published or completed between January 2001 and July 2020. The databases searched included Web of Science, ERIC, SCOPUS, and ProQuest Dissertations and Theses. We also searched the China National Knowledge Infrastructure Net (CNKI), which is the

largest academic database integrating journal articles and unpublished masters' and doctoral theses in China. The thematic words used were *quanying jiaoxue* (full English teaching), *shuangyu jiaoyu* (bilingual education), or *shuangyu jiaoxue* (bilingual instruction/teaching). Since an extremely large number of studies was returned, the majority of which were not empirical but simply involved theoretical discussion, we decided to add the controlled term *xiaoguo* (effect/effectiveness) to the Abstract. Initial searches yielded a total number of 4610 studies, of which 79 are journal articles in English, 3917 journal articles in Chinese, and 614 unpublished theses in Chinese. Moreover, we performed forward and backward searches by consulting relevant published review articles, including Macaro et al. [2] and Tong et al. [39]. An addition of 35 studies was then procured, of which 32 were journal articles in Chinese, 2 book chapters in English, and 1 book in Chinese.

In further screening the literature [44], the following inclusion criteria were set up: (a) The study was conducted in tertiary education institutions in mainland China; (b) the study was empirical and involved two groups of students studying the same course in EMI and CMI, respectively; (c) the academic achievement was measured by final exams, tests, or other forms of standardized measures rather than gauged by students' self-reports. Meanwhile, the exclusion criteria employed in this analysis are as follows: (a) The study targeted students majoring in English; (b) the study was conducted in Hong Kong, Macao, Taiwan, and other regions outside mainland China; (c) the study examined bilingual programs involving Mandarin Chinese and an ethnic minority language; (d) the study compared the effectiveness of some pedagogical methods such as flipped classrooms or problem-based learning on students' academic achievement within bilingual programs. Authors of studies with insufficient data information were contacted, which enabled us to include Li's [35] study. Later analysis revealed a detected outlier [45], which was removed (presented later). Altogether, 36 studies involving 44 independent samples fit our criteria, of which 32 samples were notably from medical disciplines. The total sample was 7582 students, of whom 3117 took EMI and 4465 CMI courses.

3.2.2. Coding Procedure

All studies were coded by the two authors independently. We first established the protocol that specified several features to be coded under the consideration of the current research purposes and previous systematic reviews [2,39]. The features include the basic information (e.g., authors, year of publication, title of the work, title of journal or edited book) and the moderators to be examined (e.g., discipline, research design, instruction time in English, publication type, and journal type in terms of English or Chinese). To ensure the reliability of the coding, the first author conducted three rounds of coding, after which the second author independently coded all the studies. While high agreement was achieved (98.33%), the disagreement was in the coding of the variable of instruction time in English. This variable was initially coded into four categories in terms of the percentage of instruction time in which English was used (50% or above, between 20% and 50%, below 20%, and unreported). Ostensibly, this seemed a low-inference variable, but it was found that many of the studies did not report actual or estimated percentage of time when English was used as the MI. After negotiation, we decided to reduce the coding into three categories (50% or above, below 50%, and unreported), and intercoder reliability was 100%.

3.2.3. Computing Effect Size

To examine the effectiveness of EMI on students' academic achievements, the standardized mean difference (Cohen's *d*) in academic achievements between students receiving EMI and CMI was used as the ES. The magnitude of effect sizes was gauged as small, medium, and large, with the values of Cohen's *d* being 0.20, 0.50, and 0.80, respectively [46].

Random-effects models were chosen for this meta-analysis. Random-effects models are based on the assumption that the true effect may vary across the studies included in the analysis, which contrasts with their counterparts' fixed-effects models that assume homogeneity of the true effect in all studies [47]. Random-effects models yield more conservative

estimates and are generally more preferred over fixed-effects models [47]. Additionally, random-effects models are more appropriate for making generalized conclusions or unconditional inferences about the magnitude of the association between variables based on the meta-analytic findings [40].

Since the sample size was small, outliers were examined in that their presence would have a significant impact on the results. Following Li's [48] practice, outliers were detected by transforming the effect sizes from independent samples into *z* scores. Any resultant absolute values of the *z* scores larger than 2.0 indicated the presence of outliers. This procedure was performed in the analyses of both subject content and English learning achievements. Consequently, Zhang's [45] study was identified as an outlier in both analyses and removed from further analysis. Table 1 shows the coded characteristics of the studies included in this meta-analysis.

Table 1. Coded characteristics of the collected studies.

Study	Sample Size		Achievement	Effect Size (<i>d</i>)	SE	Discipline	Design	Instruction Time in English	Publication Type	Journal Type
	EMI	CMI								
Yu [49]	95	101	Subject content	−0.241	0.143	OHD (Physics)	NE	Unreported	Thesis	NA
Sun and Xiao [50]	52	39	Subject content	−0.156	0.212	OHD (Mathematics)	RCT	≥50%	Article	Chinese
Zhou [51]	29	31	Subject content English	−0.430 0.581	0.261 0.264	OHD (Physics)	QED	Unreported	Thesis	NA
Han [37]	137	137	Subject content English	0.217 0.706	0.121 0.128	OHD (Mathematics)	QED	≥50%	Article	English
Li et al. [52]	56	19	Subject content	0.178	0.330	MD	NE	Unreported	Article	Chinese
Peng [53]	36	36	English	0.734	0.254	OHD (Mathematics)	QED	≥50%	Thesis	NA
Wang [54]	59	59	Subject content	−0.265	0.185	OHD (Computer Science)	NE	Unreported	Article	Chinese
Zhao et al. [55]	53	54	Subject content	0.415	0.195	MD	NE	Unreported	Article	Chinese
Deng et al. [56]	147	133	Subject content English	0.380 0.896	0.121 0.126	MD	NE	Unreported	Article	Chinese
Lu et al. [57]	65	65	Subject content	−0.173	0.176	MD	NE	<50%	Article	Chinese
Wang and Zheng [58]	104	101	Subject content English	0.305 9.146	0.141 0.473	MD	RCT	Unreported	Article	Chinese
Yu [59] S1	119	115	Subject content	1.357	0.145	MD	RCT	<50%	Article	Chinese
Yu [59] S2	126	141	Subject content	1.557	0.140	MD	RCT	<50%	Article	Chinese
Yu and Han [60]	137	137	Subject content English	0.031 0.477	0.121 0.123	OHD (Chemistry-related)	NE	≥50%	Article	Chinese
Li [61] S1	111	119	Subject content	1.422	0.147	MD	NE	Unreported	Article	Chinese
Li [61] S2	128	136	Subject content	1.456	0.138	MD	NE	Unreported	Article	Chinese
Liu et al. [62]	64	67	Subject content	−0.124	0.175	OHD (Physics)	QED	≥50%	Article	Chinese

Table 1. Cont.

Study	Sample Size		Achievement	Effect Size (d)	SE	Discipline	Design	Instruction Time in English	Publication Type	Journal Type
	EMI	CMI								
Xiang et al. [63]	53	43	Subject content	0.198	0.198	MD	NE	Unreported	Article	Chinese
Chen [64]	36	33	Subject content	−0.165	0.241	MD	NE	≥50%	Article	Chinese
Chen [65]	60	146	Subject content	0.386	0.155	MD	NE	Unreported	Article	Chinese
Pu [66]	40	40	Subject content English	−0.049 0.535	0.224 0.228	SD (Advertising)	QED	≥50%	Thesis	NA
Qin et al. [67]	62	65	Subject content English	0.055 1.661	0.178 0.207	MD	NE	≥50%	Article	Chinese
Sha et al. [68]	245	245	Subject content	2.442	0.119	MD	RCT	Unreported	Article	Chinese
Li et al. [69]	80	87	Subject content	0.619	0.158	MD	NE	Unreported	Article	Chinese
Lu et al. [70]	50	50	Subject content	0.132	0.200	MD	NE	Unreported	Article	Chinese
He et al. [71]	10	10	Subject content English	1.005 1.081	0.475 0.479	MD	RCT	Unreported	Article	Chinese
Han [72]	80	78	Subject content	2.659	0.218	MD	NE	Unreported	Article	Chinese
Jiang et al. [73]	56	56	Subject content	2.871	0.269	MD	NE	Unreported	Article	Chinese
Zhuang et al. [74] S1	36	140	Subject content	0.866	0.192	MD	NE	≥50%	Article	Chinese
Zhuang et al. [74] S2	34	157	Subject content	1.399	0.202	MD	NE	≥50%	Article	Chinese
Zhuang et al. [74] S3	32	162	Subject content	0.844	0.198	MD	NE	≥50%	Article	Chinese
Zhuang et al. [74] S4	26	144	Subject content	1.324	0.225	MD	NE	≥50%	Article	Chinese
Zhuang et al. [74] S5	23	157	Subject content	1.837	0.243	MD	NE	≥50%	Article	Chinese
Zhuang et al. [74] S6	25	151	Subject content	1.408	0.229	MD	NE	≥50%	Article	Chinese
Guo et al. [36]	18	25	Subject content English	0.390 0.198	0.312 0.310	SD (Mass Media)	QED	≥50%	Article	English
Li [35]	53	100	Subject content	−0.150	0.170	SD (Early Childhood Education)	NE	≥50%	Article	English
Mi [75]	50	50	Subject content	2.395	0.262	MD	RCT	≥50%	Article	Chinese
Du and Lu [76]	105	87	Subject content	0.767	0.150	MD	NE	Unreported	Article	Chinese
He et al. [77]	40	40	Subject content English	2.456 1.955	0.296 0.272	MD	NE	Unreported	Article	Chinese
Lin and He [33]	187	311	Subject content	−0.061	0.093	SD (Business)	NE	Unreported	Article	English

Table 1. Cont.

Study	Sample Size		Achievement	Effect Size (d)	SE	Discipline	Design	Instruction Time in English	Publication Type	Journal Type
	EMI	CMI								
Liu [78]	37	36	Subject content English	0.063 1.819	0.234 0.278	MD	RCT	Unreported	Article	Chinese
Long et al. [79]	106	54	Subject content	−1.183	0.212	MD	RCT	<50%	Article	Chinese
Yang et al. [30] S1	30	226	Subject content	0.370	0.195	MD	NE	≥50%	Article	English
Yang et al. [30] S2	25	222	Subject content	0.410	0.212	MD	NE	≥50%	Article	English

MD = medical disciplines; OHD = other hard disciplines; SD = soft disciplines; RCT = randomized controlled trial; QED = quasi-experimental design; NE = non-experimental design; S1, S2, . . . , S6 = Sample 1, Sample 2, . . . , Sample 6; NA = not applicable.

3.2.4. Assessing Publication Bias

Publication bias refers to the possibility that studies reporting unexpected or non-significant results are less likely to be published, resulting in the phenomenon that the published literature may not represent the whole picture on a research topic [40]. To address this issue, we included unpublished works such as dissertations or theses in the literature search [47]. In addition, we also assessed the potential impact of publication bias on the current meta-analysis in two ways. The first was to produce two funnel plots (one for subject content achievement and the other for English achievement), shown in Figures 1 and 2. The funnel plots show the distribution of effect sizes in relation to their standard error. As shown in Figures 1 and 2, the studies included were generally distributed symmetrically, which does not suggest publication bias [40]. Second, non-significant results were obtained from the Begg and Mazumdar rank correlation test (subject content learning achievement: $T = 0.132$, $Z = 1.245$, $p > 0.05$; English learning achievement: $T = 0.256$, $Z = 1.220$, $p > 0.05$) and Egger's regression test (subject content learning achievement: b intercept = 1.737, 95% CI (3.491, 6.965), $t(41) = 1.163$, $p > 0.05$; English learning achievement: b intercept = 4.193, 95% CI (−5.280, 13.667), $t(11) = 0.974$, $p > 0.05$). The results suggested no indication of publication bias.

Figure 1. Funnel plot of effect sizes for subject content learning achievement.

Figure 2. Funnel plot of effect sizes for English learning achievement.

4. Results

4.1. Differences in Subject Content and English Learning Achievements

The first research question concerned the differences in subject content and English achievements between students taking EMI and CMI courses. The results of the meta-analysis are shown in Table 2. It can be seen that in 43 independent samples (since Peng's [53] study did not examine the participants' subject content learning achievement, the comparison of the difference in content learning was conducted with 43 independent samples), EMI students performed better in subject content learning, with a medium and significant ES, $d = 0.673$, $SE = 0.134$; $Z = 5.032$, $p < 0.001$ (see Appendix A for the forest plot showing effect sizes for content subject achievements). The homogeneity test showed that these effect sizes were heterogeneous, $Q(42) = 47.817$, $p < 0.001$, $T = 0.852$, $I^2 = 12.166\%$. This result justified the use of random-effects models and the need to explore the possible presence of moderators.

Table 2. Results of the meta-analysis based on the random-effects model.

Achievement	<i>k</i>	Mean <i>d</i>	<i>SE</i>	95% CI	<i>Z</i>	<i>p</i>
Subject content	43	0.673	0.134	(0.411, 0.936)	5.032	0.000
English	12	1.583	0.344	(0.908, 2.258)	4.598	0.000

It is also seen in Table 2 that in 12 independent samples whose English learning achievement was tested, EMI students had better achievement, with a large and significant ES, $d = 1.583$, $SE = 0.344$; $Z = 4.598$, $p < 0.001$ (see Appendix B for the forest plot showing effect sizes for English achievements). A heterogeneity of effect sizes was also detected, $Q(11) = 41.911$, $p < 0.001$, $T = 1.345$, $I^2 = 73.754\%$.

At this juncture, it should be noted that the significant and medium summary effect size for subject content learning was unexpected and somewhat inconsistent with the general view on the effectiveness of EMI found in the literature. This unexpected result may be attributed to the fact that out of the 43 independent samples, 32 (74.42%) were from medical disciplines. The overrepresentation of medical disciplines may bias the findings from other disciplines. Hence, we performed another round of analyses on studies in disciplines other than medical disciplines. The results showed that these EMI students had lower performance than CMI students in subject content learning, but the difference was not significant, $d = -0.064$, 95% CI $(-0.175, 0.046)$, $Z = -1.143$, $p > 0.05$. The EMI students had higher English achievement than their CMI counterparts, with a significant and medium ES, $d = 0.572$, 95% CI $(0.429, 0.715)$, $Z = 7.832$, $p < 0.000$. The homogeneous test showed non-significant results for both achievement in academic subjects, $Q(11) = 13.322$, $p > 0.05$, $T = 0.091$, $I^2 = 24.934\%$, and English learning, $Q(5) = 3.596$, $p > 0.05$, $T = 0.000$,

$I^2 = 0.000\%$. These results implied that studies from disciplines other than medical science did not seem to vary in their findings of the effectiveness of EMI. In other words, the studies in non-medical disciplines collected in this analysis may form a homogeneous group in terms of academic and English learning achievements.

4.2. Moderator Variables

To address the second research question, we assessed the effect sizes, 95% CI and between-group heterogeneity for the five moderators (discipline, research design, instruction time in English, publication type, and journal type) using the random-effects model [47]. The results for subject content learning achievement are shown in Table 3.

Table 3. Moderator analysis for subject content learning achievement.

Moderator	<i>k</i>	Mean <i>d</i>	SE	95% CI	<i>Q</i>	<i>p</i>
Discipline					14.190	0.000
MD	32	0.933	0.137	(0.665, 1.202)		
OHD	7	−0.134	0.290	(−0.701, 0.434)		
SD	4	0.023	0.3857	(−0.736, 0.782)		
Design					3.706	0.157
RCT	9	0.868	0.285	(0.310, 1.426)		
QED	5	−0.000	0.381	(−0.748, 0.747)		
NE	29	0.728	0.157	(0.420, 1.036)		
Instruction time in English						
≥50%	18	0.599	0.211	(0.185, 1.013)	0.851	0.653
<50%	4	0.404	0.444	(−0.466, 1.274)		
Unreported	21	0.790	0.196	(0.407, 1.174)		
Publication Type						
Published	40	0.741	0.136	(0.474, 1.008)	3.591	0.058
Thesis	3	−0.238	0.499	(−1.216, 0.739)		
Journal Type						
English	6	0.191	0.344	(−0.484, 0.865)	3.014	0.083
Chinese	34	0.839	0.145	(0.554, 1.124)		

MD = medical disciplines; OHD = other hard disciplines; SD = soft disciplines; RCT = randomized controlled trial; QED = quasi-experimental design; NE = non-experimental design.

As seen in Table 3, only one of the five variables (i.e., discipline) was a significant moderator, $Q(2) = 14.190$, $p < 0.001$. Specifically, the mean ES of studies in medical disciplines ($d = 0.933$) was significantly larger than those from other disciplines. Follow-up pairwise comparisons indicated that the mean ES in medical disciplines was considerably larger than that in other hard disciplines, $Q(1) = 10.286$, $p < 0.001$, and soft disciplines, $Q(1) = 4.078$, $p < 0.05$. No significant difference was found between studies in other hard disciplines and soft disciplines, $Q(1) = 0.126$, $p > 0.05$. The other four categorical variables did not significantly moderate the difference in subject content learning achievement between EMI and CMI students.

Table 4 shows the results of moderator analysis for English achievement. Three moderators were found to be significant: discipline, $Q(2) = 9.087$, $p < 0.05$; design, $Q(2) = 17.388$, $p < 0.001$; and instruction time in English, $Q(1) = 6.163$, $p < 0.05$.

Table 4. Moderator analysis for English learning achievement.

Moderator	<i>k</i>	Mean <i>d</i>	SE	95% CI	<i>Q</i>	<i>p</i>
Discipline					9.087	0.011
MD	6	2.672	0.506	(1.680, 3.664)		
OHD	4	0.624	0.607	(−0.565, 1.813)		
SD	2	0.369	0.867	(−1.331, 2.069)		
Design					17.388	0.000
RCT	3	3.935	0.663	(2.636, 5.235)		
QED	5	0.554	0.491	(−0.408, 1.516)		
NE	4	1.236	0.554	(0.171, 2.302)		
Instruction time in English					6.163	0.013
Above 50%	6	0.721	0.499	(−0.257, 1.699)		
Unreported	6	2.491	0.509	(1.492, 3.489)		
Pub Type					2.516	0.113
Published	9	1.916	0.412	(1.109, 2.723)		
Thesis	3	0.616	0.708	(−0.771, 2.004)		
Journal Type					2.579	0.108
English	2	0.457	1.049	(−1.599, 2.512)		
Chinese	7	2.370	0.565	(1.262, 3.478)		

MD = medical disciplines; OHD = other hard disciplines; SD = soft disciplines; RCT = randomized controlled trial; QED = quasi-experimental design; NE = non-experimental design.

Additionally, follow-up pairwise comparisons on discipline showed that the mean ES in medical disciplines was significantly larger than that in other hard disciplines, $Q(1) = 6.307$, $p < 0.05$, but not significantly larger than that in soft disciplines, $Q(1) = 2.687$, $p > 0.05$. No significant difference in mean ES was found between studies in other hard disciplines and soft disciplines, $Q(1) = 0.847$, $p > 0.05$. Post hoc pairwise comparisons on research design indicated that studies employing a randomized controlled trial yielded a significantly larger ES than non-experimental studies, $Q(1) = 6.270$, $p < 0.05$, and quasi-experimental studies, $Q(1) = 9.783$, $p < 0.01$. No significant difference was found between non-experimental studies and quasi-experimental studies, $Q(1) = 3.763$, $p > 0.05$. While instruction time in English was a significant moderator, half of the 12 studies did not report actual or estimated percentage of instruction time during which English was used. The percentage of instruction time in English was 50% or above in all the remaining six studies. Therefore, no further comparisons on this factor could be made between the six studies.

5. Discussion

5.1. Academic Achievements

This study aimed to systematically analyze the effectiveness of EMI in the mainland Chinese context through synthesizing empirical evidence from 36 studies with 44 independent samples. It was found that EMI students performed significantly better in subject content learning than their CMI counterparts, which suggested that EMI was beneficial for subject content learning. This result was somewhat unexpected, given the many concerns raised about students' inability to comprehend EMI lectures on curriculum content [29]. However, this result echoed Joe and Lee's [32] finding that students' content knowledge increased over the EMI course despite their dissatisfaction with the course, suggesting a mismatch between students' perceived and actual performance of lecture comprehension. Further analysis showed that the significant effect of EMI on subject content learning was only identified among students in medical disciplines. Likewise, Tong et al.'s [39] systematic review also found that in studies showing a significant difference in learning outcomes favoring bilingual programs, the majority were from medical science. This may be partly explained by medical education in the current context. Many medical colleges in China offer seven- or eight-year programs that combine undergraduate and postgraduate degrees [80], which requires students to invest much more effort in their disciplinary studies. Additionally, it has been found that medical students' English matriculation scores signifi-

cantly predicted their academic performance in college, which indicated the important role of English proficiency for medical students in this context [81]. The special prominence given to EMI in this field may be because this field has become a global enterprise entailing growing collaboration between medical schools in different countries, and thus, EMI is an essential means to the globalization of medical education [30,32]. That said, cautions are warranted when interpreting the current findings since the high number of medical-related studies included in this meta-analysis may not reflect the genuine effects of EMI on content learning in general domains.

The significant difference in subject content learning between EMI and CMI identified only with medical disciplines, a somewhat bewildering result, may lie in variations in the amount of English used across different disciplines. As can be found in Table 1, in the research articles from medical disciplines, the instruction time in English for 17 out of the 32 independent samples ($17/32 = 53.13\%$) was not reported. This means that the EMI medical students' more significant gains in content knowledge learning may capitalize on 53.13% of the samples for whom English provision in class was unknown. It was likely that English was not the major MI in these studies, and thus, much fewer linguistic barriers were imposed on these students. In contrast, of the 11 independent samples from other disciplines, 8 samples ($8/11 = 72.72\%$) received 50% English provision or above in class (see Table 1). The higher percentage of English provision in these studies in non-medical disciplines may demand higher levels of English proficiency from students and thus cause non-significant difference in content learning between EMI and CMI. Therefore, while the current finding of the beneficial effect of EMI on subject content learning may be consistent with students' and faculty members' favorable perceptions of EMI learning [10,13,15], it can only be interpreted within the medical disciplines in the current context and cannot be generalized to other disciplines. A case in comparison is Lo and Lo's [38] meta-analysis, which showed that EMI secondary students in Hong Kong performed significantly worse than the CMI group in science, history, and geography, while the difference in mathematics learning between the two groups was not significant. Hence, the non-significant difference in subject content learning favoring CMI students found in non-medical disciplines may still call into question the taken-for-granted positive effect of EMI on content knowledge learning. Nevertheless, the combination of the current findings and the literature [32,36] appears to suggest that EMI does not necessarily hamper students' content knowledge learning.

Compared to the mixed findings on the effectiveness of EMI on content learning, this meta-analysis revealed that EMI students obtained significantly higher achievement in English than their CMI peers. These findings were expected and consistent with the widely upheld beliefs that EMI can enhance English learning [9,10,13] and the results of Lo and Lo's [38] meta-analysis targeting secondary students in Hong Kong. It is reasonable that frequent exposure to English language input can facilitate students' English learning. Based on these findings, it may be tentatively concluded that EMI programs are beneficial for enhancing students' English learning in the current EFL context.

5.2. Moderators

This study also examined five factors that possibly moderated the differences in academic achievements between EMI and CMI students. It was found that discipline was the only significant moderator for subject content learning achievement. The studies in medical disciplines showed that EMI students performed significantly better than CMI students. In contrast, no significant difference was found between the two respective groups of students from other hard and soft disciplines. While the saliency of EMI in medical education has been addressed above, this result may also be accounted for from the perspective of future career development. Students in medical disciplines tend to engage in the medical profession and thus may expend much effort in acquiring their curriculum knowledge. In contrast, the four studies in other hard disciplines included in this meta-analysis (see Table 4) tested content learning achievement in EMI courses such as

Physics [51], Data Structure and Algorithm or Photonics and Lasers [54], Mathematics [37], and Chemistry [61], most of which are foundation instead of professional courses in other hard disciplines. These students may, therefore, not expend as much effort as medical students. Regarding soft disciplines, it has been strongly argued that knowledge in this field is structured “through new interpretations, or negotiations of phenomena, between contemporaries”, and academic literacy in these disciplines benefits more from using the local language rather than a foreign language as the medium of instruction [27] (p. 114). From this perspective, the significantly smaller ES found in soft disciplines than in medical disciplines seems plausible.

Unlike the case of content learning achievement, the difference in English learning achievement between EMI and CMI students was significantly moderated by discipline, research design, and instruction time in English. As indicated by the current results, EMI students in medical disciplines and soft disciplines appeared to benefit more in terms of English learning than students in other hard disciplines. As previously addressed, due to the pivotal consequences of their professional courses, medical EMI students may spend more time and effort on tackling English language problems to ensure comprehension, whereas EMI students in other disciplines may not expend as much effort on solving language difficulties in learning foundation courses that may not play decisive roles in their future professions. Therefore, discipline was a significant moderator for English learning achievement.

Research design and instruction time in English were also significant moderators. The most robust type, the randomized controlled trial, yielded the largest ES in the difference between EMI and CMI students. This result suggested the importance of adopting sound methodology when examining the effectiveness of EMI, which was also stressed by Tong et al. [39]. However, although instruction time in English was a significant factor, further inference about this factor’s moderation was not possible because half of the included studies did not report the instruction time using English (see Table 4).

What should be noted is that many of the studies included in this meta-analysis displayed several methodological issues, such as a lack of transparency regarding the procedures of implementing EMI and the percentage of instruction time in English; a lack of information about the subjects (e.g., age, gender) and test instruments; and insufficient evidence regarding the reliability and validity of research instruments. Many of these issues identified in this meta-analysis mirrored the problems identified in Tong et al.’s [39] systematic review of EMI studies, such as “nonrandom assignment, group incomparability” or “a lack of information on its implementation” (p. 16). These issues jointly compromised the validity of the findings of those studies. Thus, these methodological issues largely prevented an unequivocal conclusion on the effectiveness of EMI. It would be fair to say that given the evidence synthesized in this analysis, more empirical evidence harnessed with methodologically sound approaches is needed before unequivocal conclusions could be made.

With the recognition of the abovementioned methodological pitfalls observed in this inquiry, implications may be proposed. First, overall, it seemed that EMI did not undermine subject content learning based on the current findings, which should send a reassuring message to EMI teachers and program designers. However, cautious procedures such as allocation of English instruction and English materials are still in order, particularly in non-medical disciplines where small mean differences between EMI and CMI students were found. In the cases where both teachers’ and students’ English proficiency seemed insufficient, the practice of translanguaging may be an alternative to ensure that the transmission of subject content is not compromised [28]. The use of students’ L1 in class, as expected by EMI students [32], can be a source to assist students’ English language learning. EMI teachers may also expose students to authentic English more frequently, for instance, through using videotaped lectures by proficient English speakers that are widely available today. Further, to secure the effectiveness of EMI programs, English language enhancement courses or workshops for in-service teachers are in need to prepare for potential EMI implementation [11,27]. Rubio-Alcalá and Mallorquín [82] presented a detailed framework

for designing training activities for teachers in plurilingual programs. They accentuated the training of teachers' three dimensions of competencies—language, pedagogical, and emotional—and recommended activities taking such forms as workshops, team teaching, and reflective teaching. In addition, supplementing EMI courses with English-language support classes open for EMI students may also be an alternative worth exploring since teachers' expertise from the two fields could be complementary to each other in a way that brings out optimal learning opportunities for the students [10,16]. Meanwhile, although students' English score in the college entrance examination is often taken as the benchmark for registering EMI courses, scholars have challenged this practice and argued that continued language support should still be offered for students to maximize their gains in EMI courses [18,20–22]. This can find compelling evidence in Sánchez-Pérez's recent study [83] that found the frequencies of students' use of certain disciplinary–literacy variables such as moves, cohesive devices, technical words, and passive voice were significant indicators of their content proficiency in laboratory reports written in English. Therefore, as stressed by Sánchez-Pérez [83], providing subject-specific language and literacy support should be highlighted in EMI programs. Overall, as Aizawa and Rose [16] pointed out, a better understanding of how teachers comprehend and act on macro- and meso-level policies and what contextual obstacles face stakeholders is required to assist the implementation of EMI in various disciplines.

To measure the effectiveness of EMI and pinpoint when and how EMI is facilitative to students' subject content learning and English learning, more robust research designs and longitudinal studies are needed. EMI research so far appears to prevail with studies relying on students' self-reports, which could not guarantee true effects in real life. As Macaro et al. [2] rightly stated, the absence of research using objective tests indicates that “any cost-benefit evaluation of EMI is inconclusive at best and impossible at worst” (p. 64). Further, since it takes time for changes induced by EMI to occur, longitudinal studies are invaluable to discern the impact of EMI on students' academic achievements. In particular, to reveal the associations between teachers' EMI practice and students' academic gains, contextualized data about EMI practice enacted in classrooms are highly needed and can be obtained through methods such as classroom observations. Future research is also expected to examine individual and contextual factors that may influence students' academic success in EMI programs, such as students' learning motivation, affects, and engagement in inside and out-of-class learning activities. In other words, the effectiveness of EMI may also rest on factors beyond the classroom setting.

6. Conclusions

This meta-analysis examined 36 studies involving 44 independent samples that employed statistical methods comparing differences in subject content and English learning achievements between EMI and CMI students in the mainland Chinese context. It was found that EMI students generally showed significant higher academic performance than their CMI peers. However, when the samples from medical disciplines were set aside, there was no significant difference in subject content learning between the two groups, although the EMI students still had significantly higher achievement in English learning, which is probably because of the higher initial English proficiency needed for EMI students to meet registration requirements. Given the methodological insufficiency identified in the collected studies, caution is advised with regard to interpreting the validity of these findings, and more empirical evidence is needed.

EMI has become a trend and been implemented as a pedagogical strategy in many regions and countries worldwide in the quest for the internationalization and globalization of higher education. In EFL contexts, top-down policies have been issued to promote EMI, especially in HEIs. Although a large body of research has explored stakeholders' such as students' and teachers' perceptions of EMI and its perceived effectiveness, studies employing objective measures are still scarce. There are even fewer meta-analyses that assess the effectiveness of EMI drawing on effect sizes synthesized from empirical evidence.

This meta-analysis, albeit collecting studies conducted in the Chinese context, is directly pertinent to other contexts where EMI has continued to be valued for promoting both content and English language learning. The obstacles and challenges faced by EMI students reported in other contexts [16,19,27,32] are likewise found to be present in the Chinese context [7,11,28]. Therefore, the collective empirical evidence stemming from this meta-analysis can offer a sound frame of reference for policymakers and educators to weigh various factors at the macro, meso, and micro levels [16,17] when planning to adopt EMI as a pedagogical strategy. Findings of this study have also accentuated the importance of assessing the effectiveness of EMI using objective and rigorous measurements.

Author Contributions: Conceptualization, J.-EP.; methodology, J.-EP.; formal analysis, J.-EP. and X.X.; writing—original draft preparation, J.-EP. and X.X.; writing—review and editing, J.-EP. and X.X. Both authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: The authors would like to sincerely thank Miao Li for her generous sharing of data values which were necessary and significant to this study.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Figure A1. Forest plot showing effect sizes for content-subject achievement.

Appendix B

Figure A2. Forest plot showing effect sizes for English achievement.

References

- Galloway, N.; Ruegg, R. The provision of student support on English Medium Instruction programmes in Japan and China. *J. Engl. Acad. Purp.* **2020**, *45*, 100846. [\[CrossRef\]](#)
- Macaro, E.; Curle, S.; Pun, J.; An, J.S.; Dearden, J. A systematic review of English medium instruction in higher education. *Lang. Teach.* **2018**, *51*, 36–76. [\[CrossRef\]](#)
- Wannagat, U. Learning through L2—Content and language integrated learning (CLIL) and English as medium of instruction (EMI). *Int. J. Biling. Educ. Biling.* **2007**, *10*, 663–682. [\[CrossRef\]](#)
- Di Sabato, B.; Macaro, E. Language and content intersections in CLIL and EMI. *Rass. Ital. Linguist. Appl.* **2018**, *50*, 21–32.
- Zhao, J.; Dixon, L.Q. Introduction. In *English-Medium Instruction in Chinese Universities: Perspectives, Discourse and Evaluation*; Zhao, J., Dixon, L.Q., Eds.; Routledge: New York, NY, USA, 2017; pp. 1–19.
- Hu, G. The misleading academic discourse on Chinese-English bilingual education in China. *Rev. Educ. Res.* **2008**, *78*, 195–231. [\[CrossRef\]](#)
- Hu, G. English-medium instruction in higher education: Lessons from China. *J. Asia TEFL* **2019**, *16*, 1–11. [\[CrossRef\]](#)
- Hou, A.Y.C.; Morse, R.; Chiang, C.-L.; Chen, H.-J. Challenges to quality of English medium instruction degree programs in Taiwanese universities and the role of local accreditors: A perspective of non-English-speaking Asian country. *Asia Pac. Educ. Rev.* **2013**, *14*, 359–370. [\[CrossRef\]](#)
- Dearden, J.; Macaro, E. Higher education teachers' attitudes towards English medium instruction: A three-country comparison. *Stud. Second Lang. Learn. Teach.* **2016**, *6*, 455–486. [\[CrossRef\]](#)
- Galloway, N.; Kriukow, J.; Numajiri, T. Internationalisation, higher education and the growing demand for English: An investigation into the English medium of instruction (EMI) movement in China and Japan. *Br. Counc.* **2017**. Available online: https://englishagenda.britishcouncil.org/sites/default/files/attachments/h035_eltra_internationalisation_he_and_the_growing_demand_for_english_a4_final_web.pdf (accessed on 16 October 2019).
- Hu, G.W.; Lei, J. English-medium instruction in Chinese higher education: A case study. *High. Educ.* **2014**, *67*, 551–567. [\[CrossRef\]](#)
- Hu, J.; Wu, P. Understanding English language learning in tertiary English-medium instruction contexts in China. *System* **2020**, *93*, 102305. [\[CrossRef\]](#)
- Kong, M.; Wei, R. EFL learner's attitudes toward English-medium instruction in China: The influence of sociobiographical variables. *Linguist. Educ.* **2019**, *52*, 44–51. [\[CrossRef\]](#)
- Tong, F.; Shi, Q. Chinese–English bilingual education in China: A case study of college science majors. *Int. J. Biling. Educ. Biling.* **2012**, *15*, 165–182. [\[CrossRef\]](#)
- Graham, K.M.; Eslami, Z.R. Attitudes toward EMI in East Asia and the Gulf: A systematic review. *Lang. Probl. Lang. Plan.* **2019**, *43*, 8–31. [\[CrossRef\]](#)
- Aizawa, I.; Rose, H. An analysis of Japan's English as medium of instruction initiatives within higher education: The gap between meso-level policy and micro-level practice. *High. Educ.* **2019**, *77*, 1125–1142. [\[CrossRef\]](#)
- Rose, H.; McKinley, J. Japan's English-medium instruction initiatives and the globalization of higher education. *High. Educ.* **2018**, *75*, 111–129. [\[CrossRef\]](#)

18. Chou, M.H. Speaking anxiety and strategy use for learning English as a foreign language in full and partial English-medium instruction contexts. *TESOL Q.* **2018**, *52*, 611–633. [\[CrossRef\]](#)
19. Doiz, A.; Lasagabaster, D.; Pavon, V. The integration of language and content in English-medium instruction courses: Lecturers' beliefs and practices. *Iberica* **2019**, *38*, 151–175.
20. Hua, T.L. Understanding the learning challenges of English-medium instruction learners and ways to facilitate their learning: A case study of Taiwan Psychology students' perspectives. *Lat. Am. J. Content Lang. Integr. Laclil* **2019**, *12*, 321–340. [\[CrossRef\]](#)
21. Lei, J.; Hu, G. Is English-medium instruction effective in improving Chinese undergraduate students' English competence? *Int. Rev. Appl. Linguist. Lang. Teach.* **2014**, *52*, 99–126. [\[CrossRef\]](#)
22. Li, M. The effectiveness of a bilingual education program at a Chinese university: A case study of social science majors. *Int. J. Biling. Educ. Biling.* **2018**, *21*, 897–912. [\[CrossRef\]](#)
23. Hu, G.; Li, L.; Lei, J. English-medium instruction at a Chinese University: Rhetoric and reality. *Lang. Policy* **2014**, *13*, 21–40. [\[CrossRef\]](#)
24. Hu, G.; Li, X. Asking and answering questions in English-medium instruction classrooms: What is the cognitive and syntactic complexity level? In *English-Medium Instruction in Chinese Universities: Perspectives, Discourse and Evaluation*; Zhao, J., Dixon, L.Q., Eds.; Routledge: New York, NY, USA, 2017; pp. 184–203.
25. Vinke, A.A. English as the Medium of Instruction in Dutch Engineering Education. Ph.D. Thesis, Delft University of Technology, Delft, The Netherlands, 1995. Available online: <http://repository.tudelft.nl/islandora/object/uuid:491b55f9-fbf9-4650-a44d-acb9af8412a8?collection=research> (accessed on 22 March 2020).
26. Sánchez-Pérez, M.d.M. *Teacher Training for English-Medium Instruction in Higher Education*; IGI Global: Hershey, PA, USA, 2020.
27. Kim, J.; Kim, E.G.; Kweon, S.O. Challenges in implementing English-medium instruction: Perspectives of Humanities and Social Sciences professors teaching engineering students. *Engl. Specif. Purp.* **2018**, *51*, 111–123. [\[CrossRef\]](#)
28. Tong, F.; Tang, S. English-Medium Instruction in a Chinese University Math Classroom. In *English-Medium Instruction in Chinese Universities: Perspectives, Discourse and Evaluation*; Zhao, J., Dixon, L.Q., Eds.; Routledge: New York, NY, USA, 2017; pp. 128–144.
29. Kim, K.-R. Korean professor and student perceptions of the efficacy of English-medium instruction. *Linguist. Res.* **2011**, *28*, 711–741.
30. Yang, M.; O'Sullivan, P.S.; Irby, D.M.; Chen, Z.; Lin, C.; Lin, C. Challenges and adaptations in implementing an English-medium medical program: A case study in China. *BMC Med. Educ.* **2019**, *19*. [\[CrossRef\]](#)
31. Aguilar, M.; Muñoz, C. The effect of proficiency on CLIL benefits in Engineering students in Spain. *Int. J. Appl. Linguist.* **2014**, *24*, 1–18. [\[CrossRef\]](#)
32. Joe, Y.; Lee, H.-K. Does English-medium instruction benefit students in EFL contexts? A case study of medical students in Korea. *Asia Pac. Educ. Res.* **2013**, *22*, 201–207. [\[CrossRef\]](#)
33. Lin, T.; He, Y.Y. Does Bilingual Instruction Impact Students' Academic Performance in Content-Based Learning? Evidence from Business School Students Attending Bilingual and L1 Courses. *Sustainability* **2019**, *11*, 263. [\[CrossRef\]](#)
34. Yang, W. Content and language integrated learning next in Asia: Evidence of learners' achievement in CLIL education from a Taiwan tertiary degree programme. *Int. J. Biling. Educ. Biling.* **2015**, *18*, 361–382. [\[CrossRef\]](#)
35. Li, M. Evaluation of Learning Outcomes in an Education Course: Does It Work? In *English-Medium Instruction in Chinese Universities: Perspectives, Discourse and Evaluation*; Zhao, J., Dixon, L.Q., Eds.; Routledge: New York, NY, USA, 2017; pp. 147–164.
36. Guo, H.; Tong, F.; Wang, Z.; Min, Y.; Tang, S. English- vs. Chinese-medium instruction in Chinese higher education: A quasi-experimental comparison. *Sustainability* **2018**, *10*, 4230. [\[CrossRef\]](#)
37. Han, J. An empirical study on students' academic achievements in jointly sponsored bilingual programs. *Teach. Engl. China* **2009**, *32*, 3–12.
38. Lo, Y.Y.; Lo, E.S.C. A meta-analysis of the effectiveness of English-medium education in Hong Kong. *Rev. Educ. Res.* **2014**, *84*, 47–73. [\[CrossRef\]](#)
39. Tong, F.; Wang, Z.; Min, Y.; Tang, S. A systematic literature synthesis of 19 years of bilingual education in Chinese higher education: Where does the academic discourse stand? *Sage Open* **2020**, *10*. [\[CrossRef\]](#)
40. Card, N.A. *Applied Meta-Analysis for Social Science Research*; Guilford Publications: New York, NY, USA, 2012.
41. Minister of Education. Guanyu Jiaqiang Gaodeng Xuexiao Benke Jiaoxue Gongzuo Tigao Jiaoxue Zhiliang De Ruogan Yijian [Guidelines for Improving the Quality of Undergraduate Teaching]. 2011. Available online: http://www.moe.gov.cn/jyb_sjzl/moe_364/moe_302/moe_309/tnull_4682.html (accessed on 6 May 2018).
42. Xu, H. College Students' Attitudes toward English-Medium Instruction and the English Language. In *English-Medium Instruction in Chinese Universities: Perspectives, Discourse and Evaluation*; Zhao, J., Dixon, L.Q., Eds.; Routledge: New York, NY, USA, 2017; pp. 59–75.
43. Becher, T.; Trowler, P.R. *Academic Tribes and Territories: Intellectual Enquiry and the Culture of Disciplines*, 2nd ed.; SRHE and Open University Press: Philadelphia, PA, USA, 2001.
44. Moher, D.; Liberati, A.; Tetzlaff, J.; Altman, D.G. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Int. J. Surgery* **2010**, *8*, 336–341. [\[CrossRef\]](#) [\[PubMed\]](#)
45. Zhang, G. Jichu hulixue shuangyu jiaoxue chutan [A preliminary study of bilingual teaching of Basic Nursing course]. *Keji Chuangxin Daobao* **2012**, *20*, 185–187.
46. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed.; Lawrence Erlbaum Associates: Hillsdale, NJ, USA, 1988.

47. Borenstein, M.; Hedges, L.V.; Higgins, J.P.T.; Rothstein, H.R. *Introduction to Meta-Analysis*; John Wiley & Sons: Chichester, UK, 2009.
48. Li, S. The effectiveness of corrective feedback in SLA: A meta-analysis. *Lang. Learn.* **2010**, *60*, 309–365. [[CrossRef](#)]
49. Yu, F. Daxue Wuli Lixue Kecheng Shuangyu Jiaoxue De Shijian Yanjiu [A Study of Bilingual Teaching Practice on Mechanics Course of College Physics]. Master's Thesis, Capital Normal University, Beijing, China, 2004.
50. Sun, L.; Xiao, J. Gaoshi benke yingyong shuangyu jiaoxue de chubu yanjiu yu tansuo [The preliminary research and exploration of bilingual teaching in normal university]. *Guangdong Jiaoyu Xueyuan Xuebao* **2006**, *26*, 69–72.
51. Zhou, J. A Study of the Effects of Bilingual Teaching on College Students' English Proficiency and Subject Content Knowledge. Master's Thesis, Hunan University, Changsha, China, 2008.
52. Li, X.; Tian, W.; Bao, C.; Li, X.; Fan, L.; Gao, X.; Zhao, Y. 5nianzhi yufang yixue benkesheng liuxing bingxue shuangyu jiaoxue moshi de tansuo yu shijian [Exploration and practice on bilingual teaching mode of Epidemiology for five-year program preventive medical undergraduates]. *Zhongguo Gaodeng Yixue Jiaoyu* **2009**, *10*, 10–12.
53. Peng, L. Status Quo and Effectiveness of Bilingual Education at a University in China: A Case Study of Bilingual Higher Mathematics Course. Master's Thesis, National University of Defense Technology, Changsha, China, 2009.
54. Wang, J. "Shuangyu-waiyu" tongbu huanjingxia de jiaoxue xiaoguo shiyan yanjiu [An experimental study on the effectiveness of the synchronous bilingual-EFL teaching]. *Xinan Keji Daxue Xuebao* **2009**, *26*, 70–74.
55. Zhao, R.; Wang, J.; Liu, Z. Zhenjiuxue shuangyu jiaoxue de xiaoguo fenxi [An analysis of the effectiveness of bilingual teaching of Acupuncture course]. *Changchun Zhongyiyao Daxue Xuebao* **2009**, *25*, 639–640.
56. Deng, S.; Li, X.; Huang, H.; Li, C.; He, Q.; Chen, C. Benke husheng jichu hulixue shuangyu jiaoxue xiaoguo yanjiu [A study of the effectiveness of bilingual teaching of "Basic Nursing" for undergraduate nursing majors]. *Xiangnan Xueyuan Xuebao* **2010**, *12*, 77–78.
57. Lu, S.; Wu, J.; Cheng, M. Chuanranbingxue shuangyu jiaoxue xiaoguo pingjia yu fenxi [The evaluation and analysis of the effectiveness of bilingual teaching of Epidemiology]. *Guiyang Yixueyuan Xuebao* **2011**, *36*, 98–99, 101, 110.
58. Wang, W.; Zheng, M. Shuangyu jiaoxue zai xitong jiepouxue jiaoxuezhong de yingyong xiaoguo pingjia [The evaluation of the effectiveness of the application of bilingual teaching of Systematic Anatomy]. *Sichuan Jiepouxue Zazhi* **2011**, *19*, 80–82.
59. Yu, C. Gaozhi huli zhuanke erke huli shuangyu jiaoxue xiaoguo tanjiu [Exploring the effectiveness of bilingual teaching of Pediatric Nursing for higher vocational nursing majors]. *Yixue Xinxin* **2011**, *24*, 1142–1143.
60. Yu, L.; Han, J. Chushi yingyu shuiping dui quanyingyu shuangyu jiaoxue xiaoguo de yingxiang [The Impact of Prior English Proficiency on the Effectiveness of English-medium Programs]. *Zhongguo Waiyu* **2011**, *8*, 74–81.
61. Li, H. Gaozhi gaozhuan shewai huli zhuanke binglixue jiaoxuezhong shuangyu jiaoxue de xiaoguo tanjiu [Exploring the effectiveness of bilingual teaching of Pathematology for international-oriented nursing specialty in higher vocational colleges]. *Keji Xinxin* **2012**, *21*, 145–146.
62. Liu, Z.; Luo, Y.; Hu, H. Gaoxiao shuangyu jiaoxue shizheng fenxi [An empirical analysis of bilingual education in Chinese colleges and universities]. *Zhongguo Daxue Jiaoxue* **2012**, *5*, 58–60.
63. Xiang, Z.; Wang, W.; Wang, H.; Wang, W.; Yang, Z. Shuangyu jiaoxue zai renti jishengchongxue zhong de yingyong xiaoguo chutan [Effect of bilingual teaching in Human Parasitology]. *Kunming Yixueyuan Xuebao* **2012**, *33*, 125–127.
64. Chen, L. Neike huli shuangyu jiaoxue de shijian yu tantao [The practice and exploration of bilingual teaching of Medical Nursing]. *Zhongguo Minzu Minjian Yiyao* **2013**, *22*, 60–62.
65. Chen, X. Neikexue shuangyu jiaoxue xiaoguo ji fenxi [The analysis of the effectiveness of bilingual teaching of Internal Medicine]. *Zhongwai Yixue Yanjiu* **2014**, *22*, 70–71.
66. Pu, Y. Dijinshi shuangyu jiaoxue moshi zai gaoxiao guanggao zhuanke de yingyong yanjiu—Yi Xi'an Waiguoyu Daxue Guanggao Chuangyi ke weili [A Study of Progressive Bilingual Teaching Mode for the Advertising Specialty in Colleges and Universities: A Case Study of Advertising Creativity Course in Xi'an International Studies University]. Master's Thesis, Xi'an International Studies University, Xi'an, China, 2013.
67. Qin, D.; Ye, Y.; Xiong, Y.; Liu, J.; Li, H. Putong yixue yuanxiao youxiu benkesheng zhong kaizhan yaolixue shuangyu jiaoxue de tansuo [Exploration of bilingual pharmacology education to excellent undergraduates in general medical colleges]. *Zhongguo Gaodeng Yixue Jiaoyu* **2013**, *12*, 5–6.
68. Sha, B.; Gao, W.; Feng, H.; Jing, X. Waiyu shentou moshi zai yixue yuanxiao xibao shengwuxue shuangyu jiaoxue zhong de yingyong [Application of maintenance model in bilingual education of Cell Biology at medical colleges]. *Xibei Yixue Jiaoyu* **2014**, *22*, 907–909.
69. Li, S.; Gao, Y.; Liu, C. Wunianzhi huli benke "hulixue jichu" shiyanke shuangyu jiaoxue de xiaoguo pingjia [The evaluation of the effectiveness of bilingual teaching in the experimental course of Fundamentals of Nursing for five-year nursing undergraduates]. *Jiaoyu Guancha* **2015**, *4*, 78–79, 81.
70. Lu, X.; Yang, S.; Dong, Z.; Zhou, L.; Guan, Y. Shewai huli zhuanke yixue mianyi yu bingyuan shengwu shuangyu jiaoxue de tansuo yu shijian [Using bilingual teaching in medical immunology and pathogen biology for international-oriented nursing specialty]. *Zhongguo Gaodeng Yixue Jiaoyu* **2015**, *6*, 67–68.
71. He, L.; Ma, F.; Li, Y.; Zhou, M.; Long, P.; Gong, Y. Shenzang neike shuangyu jiaoxue chafang de shijian yu tansuo [The practice and exploration of bilingual teaching model in ward rounds of Internal Medicine of Nephrology]. *Xiandai Shengwu Yixue Jinzhan* **2016**, *16*, 546–549.
72. Han, J. Zhongyi neikexue shuangyu jiaoxue gaige tantao [Research on reform of "bilingual teaching" in Internal Medicine of Traditional Chinese Medicine]. *Zhongguo Zhongyiyao Xiandai Yuancheng Jiaoyu* **2017**, *15*, 43–44.

73. Jiang, B.; Wang, A.; Yang, P. Shuangyu jiaoxue zai yanke jiaoxue zhong de yunyong yanjiu [Application research of bilingual teaching in ophthalmology teaching]. *Zhongguo Jixu Yixue Jiaoyu* **2017**, *9*, 25–27.
74. Zhuang, W.; Zhang, X.; Huang, S.; Hong, H.; Chen, X.; Huang, T.; Ma, Z. Shantou Daxue Yixueyuan quanying yu zhongwen jiaoxue xiaoguo duibi fenxi—Yi jichu zonghe kaoshi ji OSCE weili [Comparative analysis of the effects of the English-based teaching and Chinese-based Teaching: Taking Shantou University Medical College as an example]. *Yixue Jiaoyu Yanjiu Yu Shijian* **2017**, *25*, 912–915.
75. Mi, W. Bingli shenglixue shuangyu jiaoxue kecheng gaige yanjiu [Study on the reform of bilingual teaching in Pathophysiology]. *Zhongguo Jixu Yixue Jiaoyu* **2018**, *10*, 5–6.
76. Du, L.; Lu, G. Hulixue zhuan ye zhong mei he zuo ban xue fang xiang “Hulixue Daolun” chuang xinxing shuangyu jiaoxue moshi de xiaoguo yanjiu [Study on the effect of the innovative bilingual teaching model of Introduction to Nursing in the direction of Sino-US cooperative education in nursing]. *Yixue Jiaoyu Yanjiu Yu Shijian* **2019**, *27*, 143–146.
77. He, Y.; Huang, W.; Zhang, M. Erke hulixue shuangyu jiaoxue ziyuanku de yingyong dui jiaoxue xiaoguo de yingxiang [The influence of the application of bilingual teaching resource bank on the teaching effect of Pediatric Nursing]. *Jiaoyu Xiandaihua* **2019**, *58*, 130–131.
78. Liu, A. Banianzhi guipeisheng shuangyu linchuang jiaoxue zai gushangke yingyong fenxi [Analysis on the application of bilingual clinical teaching in eight-year system standardized training students in Orthopedics and Traumatology]. *Zhongguo Zhongyiyao Xiandai Yuancheng Jiaoyu* **2019**, *17*, 1–3.
79. Long, X.; Tao, Y.; Liu, B.; Sun, Q.; Zhao, Y.; Guo, H.; Han, X. Fuchankexue shuangyu jiaoxue shijian yu sikao [Practice and reflection on bilingual teaching of Obstetrics and Gynecology]. *Jixu Yixue Jiaoyu* **2019**, *33*, 23–25.
80. Xu, L.; Xu, N.; Ren, Y. Jichu yixue zhuan ye xuesheng peiyang qingkuang de diaocha fenxi [Investigation and analysis of the training of students majoring in basic medical science]. *Jichu Yixue Jiaoyu* **2017**, *19*, 636–638.
81. Tan, R. Yixuesheng ruxue chengji dui daxue chengji de yingxiang yanjiu [A study on the effect of medical students' admission scores on their academic scores in university]. *Xueyuan* **2013**, *12*, 52.
82. Rubio-Alcalá, F.D.; Mallorquín, S. Teacher Training Competences and Subsequent Training Design for Higher Education Plurilingual Programs. In *Teacher Training for English-Medium Instruction in Higher Education*; Sánchez-Pérez, M.d.M., Ed.; IGI Global: Hershey, PA, USA, 2020; pp. 41–61.
83. Sánchez-Pérez, M.d.M. Predicting content proficiency through disciplinary-literacy variables in English-medium writing. *System* **2021**, *97*, 102463. [[CrossRef](#)]