# Dispelling the Myth of Elitism and Establishing the Evidence of Inclusion: A Case of CLIL in Online English Education Amid COVID-19 Pandemic

**Hengzhi Hu**

The National University of Malaysia (UKM), Malaysia, https://orcid.org/0000-0001-5232-913X

**Abstract**: Educational elitism is an issue to be tackled, the necessity of which has already been widely established but reinforced amid COVID-19 pandemic given that learning has been greatly disrupted; elitism in Content and Language Integrated Learning (CLIL) is a matter of debate, the settling of which still awaits more empirical studies in different educational contexts. Conducted in a particular Chinese higher education provider, this mixed-methods study examined the effects of CLIL in an online English teaching programme and indicated that when teaching practices were appropriately designed and enacted in response to learners' differing needs, students of different academic capabilities could make remarkable achievement in both language learning and content learning. This makes a contribution, though small, toward dispelling the myth of elitism in CLIL and establishes some evidence on the promotion of inclusive education in the time of COVID-19 crisis. However, a critical view should be held to elitism in either online education or CLIL with a much wider research agenda in various educational contexts to be explored.

**Keywords:** Online learning, English language education, CLIL, Inclusive education

## Introduction

The outbreak of COVID-19 pandemic since 2020 continues to pose radical challenges to education beyond the impact of the disease itself, creating enormous disruption to educational systems and aggravating the already-existing educational disparities by depriving innumerable students of their right to education. Even after over a year of upheavals and fragmented lockdowns, the pandemic situation is still evolving and giving rise to very unpredictable trends. Although governments around the world have made prompt responses with honest endeavours to promote inclusive education characterised by "a focus on (removing) systemic barriers that deny opportunities for presence, participation and achievement" in education and are keeping revising their contingency plans (Johnstone et al., 2020, p. 99), chances are that countless learners are still having disrupted and unsatisfactory learning experiences (Bozkurt et al., 2020; Reimers et al., 2021), necessitating the re-examination of the efficiency of education, especially distance education, under the circumstance of COVID-19 pandemic.



Hu, H. (2021). Dispelling the myth of elitism and establishing the evidence of inclusion: A case of CLIL in online English education amid COVID-19 pandemic. In V. Akerson & M. Shelley (Eds.), *Proceedings of IConSES 2021-- International Conference on Social and Education Sciences* (pp. 29-50), Chicago, USA. ISTES Organization.

In the field of education, elitism is interpreted by Börjesson and Broady (2016, p. 112) as a social process characterised by "highly selective recruitment of students based on scholastic merit" in programmes and marginalisation of disadvantaged learners aimed at contributing "to the intra-generational reproduction of the current dominant groups by educating and training their offspring". In the time of COVID-19 crisis, elitism seems to be aggravating rather than calming down, reinforcing educational inequalities caused by "a lack of strong pedagogic support for students from disadvantaged and marginalised spaces" (Devkota, 2021, para. 1). This essentially reflects the nature of exclusion which obstructs the full participation of all learners in educational activities (Loreman et al., 2011). Content and Language Integrated Learning (CLIL), a pedagogical approach that emphasises both language learning and content learning in a single educational programme (Coyle et al., 2010), has also long been considered to be elitist, which favours and is assumed to be the most beneficial for the learners who are evaluated to be advanced and advantaged in intelligence, academic performance and all the other socio-educational factors (Bruton, 2015; Broca, 2016; Rumlich, 2014). Opposite to this exclusive view, a number of scholars state that CLIL is open to all learners regardless of the skewed composition of learner population and is thus more egalitarian in essence (Coyle et al., 2010; Hüttner & Smit, 2014), though few empirical studies are available to further confirm this.

With this brief introduction in mind, the author of this paper wishes to focus on the context of an online CLIL programme offered by a particular Chinese higher education provider during COVID-19 lockdowns with an attempt to rectify online learning elitism and examine whether CLIL is elite or not. By adopting mixed-methods research designs, the author focuses on the comparison of the learning outcomes of the academically advanced and academically weak cohorts involved in the same online CLIL programme as well as on the teaching and learning activities in which students were engaged, aimed at answering the following questions:

1) To what extent do online CLIL instructions affect the English learning and content learning achievement of the learners who have differing academic levels?
2) What online learning context does CLIL provide for the learners?

## Literature Review

### Elitism and COVID-19: A Focus on China

Elitism is a topic frequently discussed in Chinese academia with special to education and social justice. For the past few decades, the development of education and the distribution of educational resources have been somewhat controlled and determined by elite social groups, such as "those who are rich or powerful", "hold professional positions in China" or attempt to take advantage of education to maintain their class superiority into the next generation (Wang, 2007, p. 111; see also, Jin, 2000; Wang & Yang, 2016; Wu et al., 2020). Some scholars argue that elite education characterised by the optimisation of teaching process for advanced learners via high-level educational management and significant distribution of educational resources and facilities can help to train high-quality talents with the required skills to enhance China's national power and role in the world (Hao, 2016; Liu, 2018; Xiao, 2010; Zeng, 2013), whereas a national endeavour has been made to abolish the label of elitist education and "to promote educational equity in admission criteria, resources input and

performance standards" (Zhou et al., 2007, p. 14), reflecting the nature of equality-oriented education with a focus on ensuring access and right to education for every student (Huang et al., 2016). However, it cannot be denied that elitist education still exists in all levels of education in today's China, leading to the monopolisation of elite educational opportunities by the upper class and failing to attain the ideal of inclusive education (Wu et al., 2020).

According to Jash (2020), since the outbreak of COVID-19 pandemic in China, the monopolisation of the elitist coalition seems to be aggravating. Although Jash (2020) holds a rather sceptical and questionable attitude to China's political responses to COVID-19, he does more or less spotlight the elitist issue in Chinese society. However, a critical view must be held to elitism in the field of online learning. On the one hand, the efficiency of China's prompt educational responses to COVID-19, such as the immediate organisation and offer of online courses, cannot be denied, which has ensured "the physical and mental health of teachers and students", "the supply of epidemic prevention materials and educational resources…(and) the quality of students' learning" (Xue et al., 2020, p. 1; see also, Cen et al., 2020; Gulati, 2020; Huang et al., 2020). This national endeavour, to a large extent, has ensured the continuity, quality and equality of learning. On the other hand, Xue et al. (2020) also mention that despite the effort that has been made, chances are that online learning could severely affect educational inclusion and equality and that disadvantaged students' learning could be particularly disturbed, which reflects some sense of elitism characterised by the marginalisation of less advantaged learners. Unfortunately, this assumption has only been confirmed in a few studies, such as the ones conducted by Hu et al. (2021), G. Chen et al. (2020), Y. Chen et al. (2020), Liu and Chen (2020), Ma (2020) and Zuo et al. (2021) who more or less disclose that disadvantaged learners may have unsatisfactory online learning experiences and highlight the need for teachers to re-examine their teaching practices. Nevertheless, it is necessary to revise online teaching practices during this special period in response to the need proposed by Huang et al. (2020, pp. 39-40) that effective methods to organise instructions should be worked out and employed by adopting appropriate teaching strategies and providing "support for learners with special needs…in order to provide an inclusive online learning experience for them". This is also acknowledged by Zhang et al. (2020, p. 58) who argue that there is still much ambiguity of and disagreement over the government policies with regard to how to teach and that "it remains unclear what teaching mode and pedagogy may best work for online education", which further highlights the compelling need to support academic research on online education aimed at better engaging students with learning difficulties. Currently, after over one year since the initial massive outbreak of COVID-19 virus in China, it is still safe to say that the need to examine online teaching and learning in research is not obviated but reaffirmed due to the recurrent COVID-19 outbreaks and lockdowns in some parts of China, forcing students back to online learning.

**CLIL and Elitism**

CLIL is a pedagogical approach originated from Europe and now widely utilised in bilingual programmes, which is characterised by a "dual-focused educational context in which an additional language, thus not usually the first language of the learners involved, is used as a medium in the teaching and learning of non-language

content" (Marsh, 2003, para. 3). Therefore, CLIL focuses on both language learning and content learning in a single classroom (Coyle et al., 2010). Despite the positive words describing the substantial benefits of CLIL to promote language acquisition, content learning, cognitive development and cross-cultural understanding (Coyle et al., 2009), a longstanding debate always centres on whether CLIL is selective in nature and merely beneficial for advanced learners or is inclusive of all.

Advocates maintain that CLIL is inclusive of all learners, even those who may be less advantaged or privileged, and is thus more egalitarian in essence than other kinds of bilingual programmes, such as the ones offered in private schools (Coyle et al., 2010; Hüttner & Smit, 2014). This view has been confirmed in empirical studies, such as the ones done by Pladevall-Ballester (2014) and McDougald (2015) which indicate that both high achievers and low achievers could benefit from CLIL. In contrast, other scholars assume that CLIL is selective in nature and only attracts and benefits the learners who are academically advanced and motivated (Bruton, 2011; Mehisto, 2007; Paran, 2013), and evidence shows that "many CLIL programs are de facto selective in one way or another" in various educational contexts (Bruton, 2015, p. 124; see also, Van Mensel et al., 2020). This assumption has also been confirmed in empirical studies, cautioning educators against the opening of CLIL programmes for all learners. For instance, Apsel (2012, p. 54) concludes from his study of the reasons for students to drop out of CLIL programmes that there should be a threshold of "a combination of factors including language, learning skills, content language and commitment" for CLIL admission, below which learners could easily have problems with their learning. Likewise, Zydatiß (2012, p. 26), based on the research findings that there was a strong correlation between the participants' initial academic competence with their learning outcomes in CLIL, states that "adequate levels of academic discourse proficiency…are dependent on rather high level of linguistic competence…and/or general proficiency in the working language". This finding corresponds to the one generated in Mewald's (2007) study which indicates that less advanced learners could not improve language competence in the same way as their more advanced peers involved in the same CLIL class.

There does not seem to be a one-size-fits-all answer to the question whether or not CLIL should be selective, though the fact is that it is indeed selective in practice and reflects the nature of elitism in some educational contexts. From a neutral perspective, Cañado (2019, p. 1) maintains that the "commonly harboured beliefs vis-à-vis the elitism of bilingual programmes need to be re-examined and (that) a possible future research agenda is suggested to continue advancing in this area". Although Cañado (2020, p. 16) further argues "that elitism in CLIL is increasingly a thing of the past" in the current research agenda, Hu (2021) believes that the challenges posed by COVID-19 to foreign language education have necessitated the re-examination of elitism in CLIL application. In this sense, a critical view should be held, and whether CLIL should be selective or inclusive depends on certain educational contexts and even how teachers address learners' different needs. This necessitates the organisation of empirical research in certain contexts to settle the highly context-dependent issue of elitism.

## Method

### Research Designs

This study was conducted in a particular Chinese higher educational provider, University X (pseudonym), which is famous for its CLIL programme integrating College English (CE), a compulsory course for undergraduates whose majors are not related to English studies, with American law. A sequential explanatory mixed-methods approach was employed in this study, including both quantitative and qualitative methods in order to answer the aforementioned research questions and offer "rich insights into various phenomena of interest that (could not) be fully understood using only a quantitative or qualitative method" (Liebenberg et al., 2015, p. 224).

From a total number of 85 students enrolled in this CLIL unit, 70 year-three students from the Faculty of Law were selected purposively with informed consent, and all the names appearing in this text are pseudonyms. They were selected based on the criteria that: First, they were not taking any other classes of English or American law other than this specific CLIL unit so that their academic progress could be largely attributed to the pedagogical approach used in this study; Second, they had not taken any courses related to American law before the study.

Based on the Grade Point Average (GPA) of previous English courses and law courses that the participants had completed, 38 participants were categorised into the less advanced group with relatively lower academic level (L-Group), and 32 participants were categorised into the more advanced group with higher academic performance (H-Group). Their GPA data were computed into Software Statistical Package for the Social Sciences (SPSS) 25.0. Based on the prerequisite that the Shapiro-Wilk test indicated the normal distribution of data for both the H-Group ($p > .05$) and the L-Group ($p > .05$), an independent samples $t$ test was run, with descriptive and inferential data recorded in Table 1 and Table 2 respectively. It suggested that there was a significant statistical difference between the L-Group's GPA ($M = 2.91$, $SD = .03$) with the H-Group's ($M = 3.50$, $SD = .09$), $p < .001$, with the latter having higher academic proficiency than the former.

Table 1. Group Statistics of GPA

| Group | N | Mean | Std. Deviation | Std. Error Mean |
|---|---|---|---|---|
| L | 38 | 2.910 | .033 | .005 |
| H | 32 | 3.496 | .086 | .015 |

The participants were put into two classes taught by the same teacher to ensure that all of them were instructed by the most unified methodology. In the quantitative part of this study, two instruments were utilised to measure the participants' English learning proficiency and content learning proficiency. The first one was College English Test Band-6 (CET-6), a standard nationalised test that could measure undergraduate students' English proficiency of listening, speaking, reading and writing in a valid, reliable and fair way (Zheng & Cheng, 2008). It was organised as both pretest and posttest, the papers and booklets of which were adapted from authentic test batteries. The tests were scored by professionally trained examiners in accordance with official regulations and marking rubrics to ensure the assessment validity and reliability.

Table 2. Independent Samples Test of GPA

| | Leven's Test for Equality of Variances | | t-test for Equality of Means | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | F | Sig. | t | df | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference | |
| | | | | | | | | Lower | Upper |
| Equal variances assumed | 9.844 | .003 | -38.826 | 68 | .000 | -.586 | .015 | -.616 | -.556 |
| Equal variances not assumed | | | -36.424 | 39.032 | .000 | -.586 | .016 | -.619 | -.554 |

A pilot study completed before this research indicated that the CET-6 pretest and posttest instruments had an acceptable level of reliability. The Cronbach's alpha for the objective assessment of listening and reading pretest and posttest was .93 and .87 respectively. The Cohen's Kappa was .74 and .71 respectively for the subjective writing pretest and posttest and .77 and .76 for the subjective speaking pretest and posttest.

The second instrument was American Law Knowledge Test and Portfolio Task, which was developed by the teachers from the research site to measure students' content learning proficiency and consisted of both objective multiple-choice items and subjective case analyses. As the participants had not taken any courses about American law before this study, this instrument was only administered at the end of the research to measure their content knowledge proficiency. The pilot study suggested that both the objective (Cronbach's alpha = .83) and subjective (K = .79) assessment tasks had a reasonable level of reliability.

In the qualitative part, in order to collect evidence about the learning context in CLIL, non-participant observation as a "non-judgmental description of classroom events that could be analysed and given interpretation" (Gebhard, 1999, p. 35) was done to the CLIL classes with notes taken. 10 cases were evenly recruited from the L-Group and H-Group for more in-depth study, and the observation data were triangulated with the ones collected from semi-structured interviews with the selected cases to probe into the "subjects' opinions, beliefs and feelings about the situation in their own words" (Ary et al., 2010, p. 438).

**Data Collection and Analysis**

The study was conducted online when the participants had to take online classes because of the local recurrence of COVID-19. The university's temporary closure lasted for about over two months. The study began in early 2021 and lasted for 8 weeks until the lockdown was removed. Before the study, the English proficiency pretest had been organised on an online examination system with strict invigilation. The observation data were collected during the intervention. On the first week of students returning to campus, the English proficiency

posttest and content knowledge test were administered, and the selected cases were interviewed. The gathered quantitative data were then computed into SPSS 25.0, and both descriptive and inferential data analyses were done.

In original CET-6, the total score for listening, reading and writing is 710 marks with each of them occupying 35%, 35% and 30% respectively, and the speaking test weighs 15 marks and is scored separately. The raw data of each section as well as the total scores were kept in the following analysis. The American Law Knowledge Test and Portfolio Task weighs 100 marks in total, and in this paper, only the total scores were reported. The collected qualitative data were transcribed and input into NVivo 12.0, the thematic analysis of which followed Braun and Clarke's (2006) six-step framework of familiarising the data, generating initial codes, searching for themes, reviewing themes, defining themes and producing the report. Since the selected students used Chinese in the interviews, their responses were translated to English by professional translators who deployed the back-to-back translation technique devised by Sperber (2004).

## Results and Discussion

### Improved English Proficiency

The L-Group's ($n$ = 38) and H-Group's ($n$ = 32) pretest and posttest of English proficiency data were firstly analysed through independent samples $t$ test based on the prerequisite that the Shapiro-Wilk test indicated that the data were normally distributed ($p > .05$). The descriptive and inferential data of $t$ test can be seen in Table 3 and Table 4 respectively.

Generally, the $t$ test of the pretest total score was statistically significant, with the H-Group ($M$ = 500.84, $SD$ = 23.63) reporting 41.02 points higher than the L-Group ($M$ = 459.82, $SD$ = 12.48), $p < .001$, two-tailed, $d$ = 2.23. Similarly, there was also a significant statistical difference between the subgroups with respect to the listening ($p = .011$), reading ($p < .001$), writing ($p = .005$) and speaking ($p < .001$) performance in the pretest. This indicated that the H-Group had much higher level of English proficiency at the beginning of this study than the L-Group.

After the treatment of CLIL, the $t$ test of posttest data was still statistically significant. In terms of the posttest total score, the H-Group ($M$ = 516.08, $SD$ = 24.60) reported 47 points higher than the L-Group ($M$ = 469.08, $SD$ = 10.74), $p < .001$, two-tailed, $d$ = 2.64. Also, the H-Group had higher scores of listening ($p = .001$), reading ($p < .001$), writing ($p < .001$) and speaking ($p < .001$) than the L-Group in the posttest. This suggested that there was still a gap in language proficiency between the advanced learners and the less advanced ones after the intervention.

Table 3. Group Statistics of Pretest and Posttest

| | Group | N | Mean | Std. Deviation | Std. Error Mean |
|---|---|---|---|---|---|
| Pretest-Total Score | L | 38 | 459.816 | 12.477 | 2.024 |
| | H | 32 | 500.844 | 23.629 | 4.177 |
| Pretest-Listening | L | 38 | 158.447 | 11.255 | 1.826 |
| | H | 32 | 166.5 | 14.613 | 2.583 |
| Pretest-Reading | L | 38 | 153.237 | 10.462 | 1.697 |
| | H | 32 | 173.297 | 7.748 | 1.370 |
| Pretest-Writing | L | 38 | 140.421 | 14.261 | 2.314 |
| | H | 32 | 151.625 | 17.837 | 3.153 |
| Pretest-Speaking | L | 38 | 7.711 | 0.913 | 0.148 |
| | H | 32 | 9.422 | 1.530 | 0.270 |
| Posttest-Total Score | L | 38 | 469.079 | 10.737 | 1.742 |
| | H | 32 | 516.078 | 24.603 | 4.349 |
| Posttest-Listening | L | 38 | 160.513 | 13.631 | 2.211 |
| | H | 32 | 172.188 | 13.920 | 2.461 |
| Posttest-Reading | L | 38 | 158 | 11.012 | 1.786 |
| | H | 32 | 176.906 | 8.405 | 1.486 |
| Posttest-Writing | L | 38 | 142.474 | 12.461 | 2.021 |
| | H | 32 | 157.063 | 18.555 | 3.280 |
| Posttest-Speaking | L | 38 | 8.092 | 0.761 | 0.124 |
| | H | 32 | 9.922 | 1.232 | 0.218 |

Table 4. Independent Samples Test of Subgroups' Pretest and Posttest

| | | Levene's Test for Equality of Variances | | t-test for Equality of Means | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | F | Sig. | t | df | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference | |
| | | | | | | | | | Lower | Upper |
| Pretest-Total Score | Equal variances assumed | 6.615 | 0.012 | -9.284 | 68 | .000 | -41.028 | 4.419 | -49.846 | -32.210 |
| | Equal variances not assumed | | | -8.839 | 45.179 | .000 | -41.028 | 4.642 | -50.376 | -31.680 |
| Pretest-Listening | Equal variances assumed | 2.805 | 0.099 | -2.603 | 68 | .011 | -8.053 | 3.094 | -14.226 | -1.879 |
| | Equal variances not assumed | | | -2.546 | 57.653 | .014 | -8.053 | 3.163 | -14.386 | -1.720 |
| Pretest-Reading | Equal variances assumed | 0.63 | 0.43 | -8.968 | 68 | .000 | -20.06 | 2.237 | -24.524 | -15.596 |
| | Equal | | | -9.198 | 66.98 | .000 | -20.06 | 2.181 | -24.413 | -15.707 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | variances not assumed | | | | | | | | | |
| Pretest-Writing | Equal variances assumed | 0.691 | 0.409 | -2.92 | 68 | .005 | -11.204 | 3.837 | -18.860 | -3.548 |
| | Equal variances not assumed | | | -2.865 | 59.028 | 0.006 | -11.204 | 3.911 | -19.030 | -3.378 |
| Pretest-Speaking | Equal variances assumed | 9.988 | 0.002 | -5.785 | 68 | .000 | -1.711 | 0.296 | -2.302 | -1.121 |
| | Equal variances not assumed | | | -5.551 | 48.699 | .000 | -1.711 | 0.308 | -2.331 | -1.092 |
| Posttest-Total Score | Equal variances assumed | 30 | .000 | -10.644 | 68 | .000 | -46.999 | 4.416 | -55.810 | -38.188 |
| | Equal variances not assumed | | | -10.032 | 40.861 | .000 | -46.999 | 4.685 | -56.462 | -37.537 |
| Posttest-Listening | Equal variances assumed | 0.58 | 0.449 | -3.535 | 68 | .001 | -11.674 | 3.302 | -18.264 | -5.085 |
| | Equal variances not assumed | | | -3.529 | 65.497 | .001 | -11.674 | 3.308 | -18.281 | -5.068 |
| Posttest-Reading | Equal variances assumed | 0.082 | 0.775 | -7.952 | 68 | .000 | -18.906 | 2.377 | -23.650 | -14.162 |
| | Equal variances not assumed | | | -8.137 | 67.4 | .000 | -18.906 | 2.324 | -23.544 | -14.269 |
| Posttest-Writing | Equal variances assumed | 10.199 | 0.002 | -3.913 | 68 | .000 | -14.589 | 3.728 | -22.028 | -7.149 |
| | Equal variances not assumed | | | -3.786 | 52.654 | .000 | -14.589 | 3.853 | -22.318 | -6.860 |
| Posttest-Speaking | Equal variances assumed | 12.378 | 0.001 | -7.599 | 68 | .000 | -1.830 | 0.241 | -2.310 | -1.349 |
| | Equal variances not | | | -7.309 | 49.812 | .000 | -1.830 | 0.250 | -2.333 | -1.327 |

assumed

Then, paired samples *t* tests with an α of .05 were run to examine the changes of English proficiency for each subgroup. For the L-Group, the data recorded in Table 5 indicated that there was a significant statistical difference between the pretest total score ($M$ = 459.82, $SD$ = 12.38) and posttest total score ($M$ = 469.08, $SD$ = 10.74), 95% confidence interval [-12.33, -6.19], $t(37) = -6.11$, $p < .001$, $d = .80$. Similarly, significant statistical difference between pretest and posttest could also been seen in the assessment of listening ($p = .015$), reading ($p < .001$), writing ($p = .001$) and speaking ($p < .001$). This revealed that the L-Group made considerable progress of English proficiency under the treatment of CLIL in the online context. Equally, for the H-Group, the inferential data in Table 6 indicated a statistically significant difference between the total score of pretest ($M$ = 500.84, $SD$ = 23.63) and posttest ($M$ = 516.08, $SD$ = 24.60), 95% confidence interval [-18.76, -11.71], $t(31) = -8.83$, $p < .001$, $d = .63$. In addition, a significant statistical difference could also be detected from the pretest and posttest scores of listening ($p < .001$), reading ($p < .001$), writing ($p < .001$) and speaking ($p < .001$), which demonstrated that the H-Group also made substantial progress in English learning with CLIL instructions.

Table 5. Paired Samples Test of L-Group's Pretest and Posttest

| | | Mean | Std. Deviation | Std. Error Mean | 95% Confidence Interval of the Difference | | t | df | Sig. (2-tailed) |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | Lower | Upper | | | |
| Total Score | Pretest - Posttest | -9.263 | 9.345 | 1.516 | -12.335 | -6.192 | -6.111 | 37 | .000 |
| Listening | Pretest - Posttest | -2.066 | 4.976 | 0.807 | -3.701 | -0.430 | -2.559 | 37 | .015 |
| Reading | Pretest - Posttest | -4.763 | 5.883 | 0.954 | -6.697 | -2.830 | -4.991 | 37 | .000 |
| Writing | Pretest - Posttest | -2.053 | 3.318 | 0.538 | -3.143 | -0.962 | -3.813 | 37 | .001 |
| Speaking | Pretest - Posttest | -0.382 | 0.457 | 0.074 | -0.532 | -0.232 | -5.151 | 37 | .000 |

Table 6. Paired Samples Test of H-Group's Pretest and Posttest

| | | Mean | Std. Deviation | Std. Error Mean | 95% Confidence Interval of the Difference | | t | df | Sig. (2-tailed) |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | Lower | Upper | | | |
| Total Score | Pretest - Posttest | -15.234 | 9.764 | 1.726 | -18.755 | -11.714 | -8.826 | 31 | .000 |
| Listening | Pretest - Posttest | -5.688 | 3.724 | 0.658 | -7.030 | -4.345 | -8.64 | 31 | .000 |
| Reading | Pretest - Posttest | -3.609 | 3.773 | 0.667 | -4.970 | -2.249 | -5.411 | 31 | .000 |
| Writing | Pretest - Posttest | -5.438 | 6.582 | 1.164 | -7.810 | -3.065 | -4.673 | 31 | .000 |
| Speaking | Pretest - Posttest | -0.5 | 0.440 | 0.078 | -0.659 | -0.341 | -6.429 | 31 | .000 |

**Considerable Content Knowledge Proficiency**

An independent samples *t* test was run to compare the content knowledge proficiency of the L-Group ($n$ = 38) and H-Group ($n$ = 32), with the descriptive data recorded in Table 7 and inferential data recorded in Table 8. Levene's test was non-significant ($p > .05$), indicating that the data were normally distributed. It is interesting to

note that although the average of content assessment for the H-Group ($M$ = 81.16, $SD$ = 5.08) was 1.72 points higher than that for the L-Group ($M$ = 79.44, $SD$ = 3.93), the $t$ test was not statistically significant, $t(68) = 1.56$, $p$ = .123, two tailed, 95% confidence interval [-.48, 3.92], $d$ = .37. This demonstrated that the L-Group's performance in content knowledge assessment was as good as the H-Group's at the end of this study. Based on the fact that all the learners had the same starting line of having little knowledge about the content subject prior to this study, the effect of CLIL on content learning could be considered to be equally positive for the participants who had rather differing academic proficiency.

Table 7. Group Statistics of Content Knowledge Assessment

| Group | N | Mean | Std. Deviation | Std. Error Mean |
|---|---|---|---|---|
| H-Group | 32 | 81.16 | 5.075 | .823 |
| L-Group | 38 | 79.44 | 3.934 | .696 |

Table 8. Independent Samples Test of Content Knowledge Assessment

| | Leven's Test for Equality of Variances | | t-test for Equality of Means | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | F | Sig. | t | df | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference | |
| | | | | | | | | Lower | Upper |
| Equal variances assumed | 3.205 | .078 | 1.562 | 68 | .123 | 1.720 | 1.101 | -.477 | 3.918 |
| Equal variances not assumed | | | 1.596 | 67.576 | .115 | 1.720 | 1.078 | -.431 | 3.871 |

**A Differentiated and Inclusive Learning Context**

The qualitative data suggested that differentiation, namely a teacher's responsiveness to learners' differing needs (Tomlinson, 2000), was a key feature of this online CLIL programme and that the teacher attempted to differentiate online teaching in various ways. The observation data firstly indicated that the content, namely "how the student (would) get access to the information" (Tomlinson, 2000, para. 2), was differentiated for the learners having differing levels of academic proficiency. Despite the routine lectures delivered by the classroom teacher with the aid of teacher-made PowerPoint slides that illustrated all the key language and content points, the H-Group and L-Group learners were also asked to prepare or given different learning materials for the class to achieve the same teaching objective of developing proficiency in both English and content knowledge. For instance, in the unit of Principles of American Law, the H-Group was required to read the first chapter of *An Introduction to American Law* written by Rosen et al. (2017), while the L-Group was asked to watch several videos which covered the same key points of learning (e.g., adversarial system, executive privilege, federalism, inquisitorial system) as the H-Group had but explained them in a more understandable way. Similarly, in the

unit of Criminal Law and Case Analyses, the learners were encouraged to search for different law cases that they were interested in and could understand and to bring their prepared materials to class for discussion.

Besides, the process, namely the activities in which students could engage "in order to make sense of or master the content" (Tomlinson, 2000, para. 3), was also differentiated. The classroom teacher often put students into different chatrooms on the online learning software and organised tiered activities in which all the learners worked with the same understandings but did it with different levels of support and scaffolding. An example here could be that in the unit of Criminal Law, both the H-Group and L-Group learners were required to analyse a larceny case to practice what they had learned previously. By dividing the students into different groups, the teacher could monitor their learning more efficiently and provide more scaffolding to those who had less strong academic competence. Meanwhile, students in the L-Group were given more time to finish this task, while some of the H-Group learners were given an additional case for analysis in order to well stretch them.

Another aspect of differentiation was that the learners having different levels of academic proficiency were allowed to complete differing projects to demonstrate what they had learned in this programme (Tomlinson, 2000). For example, the students were invited to present their understandings in any format, such as writing a case report or bibliography, doing a role playing between a suspect with a judge, recording a presentation, etc. When learners were allowed to decide on the formats of their assignment products, they could choose the one that they were confident with to apply what they had learned with the same ultimate learning objectives achieved.

Although the participants did not mention the term of differentiation in interviews, they did highlight how this online CLIL programme was different from other classes they were taking. Some learners from the H-Group considered this unit to be individualised. For example, Student Bai maintained:

> *We had different learning materials and tasks. For example, we were allowed to show our understanding of criminal law in many ways like role playing and presentation. This was different from other classes wherein we could only do the same learning tasks. I felt much engaged when I was given a choice for my own learning.*

This was also acknowledged by Student Li who explained:

> *We had different learning paces and levels of proficiency in this class, and there was no one-size-fits-all teaching for all of us. However, the teacher gave us a great deal of autonomy for our own learning to make it personalised. I think this was where learning occurred.*

The interviewees from the L-group gave higher praise to the differentiated instructions of this programme, and all of them noted that this programme had offered them more opportunities to learn efficiently in their own paces and learning styles. The excerpt from Student Zhao was a typical one:

> *I do not have outstanding academic performance. I was afraid at the beginning that this programme was only suitable for those overachievers, but later on, I realised that I could also make a progress*

*with the teacher providing different learning materials, processes and tasks to us…We were doing different things, but all of us were moving towards the same learning objectives…I could learn at my own pace, and indeed, I have learned a lot!*

According to Devkota (2021), elitism in online learning amid COVID-19 could derive from the lack of instructional support for those disadvantaged learners, and this essentially reflects exclusion in education rather than inclusion and highlights the need to respect individual learner's needs, eliminate marginalisation and promote engagement (Loreman et al., 2011). When elitism tends to be the norm, low student achievement could be commonplace in online classrooms in the time of COVID-19 crisis (Beck & Beasley, 2020), and this could the same case when language learners' needs could not be fully met by inappropriate instructions or pedagogical approaches (Mahyoob, 2020). Differentiation is inclusive in nature and facilitates inclusive education that engages "all learners by ensuring that each individual has an equal but personalised opportunity for learning…(and) aims at supporting educators to address the full range of learners' needs so as to overcome barriers to learning" (Boulkroun, 2020, p. 2104). Previous research findings indicate that differentiation has a positive effect on student achievement, either in on-campus education (Kotob & Arnouss, 2019) or in online learning (Beck & Beasley, 2020). This could be one of the reasons for learners' improvement in the studied CLIL programme. The implication drawn here is that when a teacher's instructions are well planned for learners having different levels of academic proficiency, the myth of elitism in online learning can be dispelled. Besides, this further reinforces that CLIL is a pedagogical approach that embraces the utilisation of different instructional techniques (Coyle et al., 2010) and that when learners' needs are well taken into account and addressed by teachers, the elitist view towards CLIL could also be abandoned.

**An Authentic and Meaningful Learning Context**

Despite an inclusive approach to online learning, the learners' progress in language learning could not be separated from the authentic and meaningful context that the teacher created, one key feature of which was an abundant amount of comprehensive input. Observation data showed that a wide variety of authentic learning materials were provided, such as TV programmes, journal articles, books, newspaper articles, video clips, etc. All the materials were updated and prepared by both the teacher and the students themselves. The participants involved in the interview also acknowledged that it was these authentic instruction materials that facilitated their learning. For example, Student Wang from the H-Group said:

*Learning materials are important, and our teacher mentioned that we needed somewhat new learning materials except for traditional textbooks. Thus, he often prepared some updated materials for us, such as news articles and videos, and sometimes required us to collect materials for the class…These materials are our likes and interests…I think they did provide us with a meaningful learning context with a lot of authentic input.*

This was also acknowledged by Student Hu from the L-Group, who maintained:

*I like the learning materials…They are more interesting and meaningful than old-fashioned textbooks.*

*They rightly offered me a chance to acquire the language closely related to my field of study and interest… I am sure that my English has improved a lot with so much input.*

Student Liu from the same group also highlighted the role that meaningful materials and input played in facilitating online English learning. Comparing the English courses, she had taken before, she asserted:

*In previous English classes, we used old textbooks a lot and found it useless and boring to learn the language that we would never use in our life. However, in this class, we have used a lot of meaningful learning materials, such as articles and audio-visual materials…I must say these learning materials and input did help me a lot with my English.*

Moreover, learners were offered a variety of opportunities and activities to practice English and apply what they had learned, such as oral presentation, case analyses, debate, role playing, etc. This indicated that this CLIL programme was also rich in comprehensive and authentic output. The interviewees further confirmed that it was these various output tasks that facilitated their learning. For instance, Student Wu from the H-Group rated these activities highly and asserted:

*These activities are different from the tedious drills we used to do in English textbooks, such as multiple-choice questions and form filling which were closely related to what we had learned about the language but not connected with my major or life. But the output activities in this unit are colourful so that I could apply what I had learned in a meaningful context related to my major.*

Similarly, Student Zhao from the L-Group also praised the diverse output activities as one of the key motivators for learning, maintaining:

*I can practice what I have learned in many different kinds of speaking and writing activities, and they have presented a valuable context related to my field of study in university. In this case, I feel more engaged in and motivated for learning.*

This was also acknowledged by Student Liu from the same group, who believed:

*The various output activities have helped me improve my language proficiency. In the process of output, I could receive feedback on my learning from my teacher and classmates…I also feel more confident thanks to the opportunities to experiment with language in different activities.*

Krashen (1985) suggests that learners could make progress in language acquisition through exposure to comprehensive input, and in his Input Hypothesis, he assumes that if a learner's current language level is regarded as 'i', acquisition occurs when they are exposed to abundant, comprehensive and natural input which contributes to level 'i + 1'. From this perspective, the productive effects of CLIL on language proficiency could be explained by the assumption that this pedagogical approach provides not only the exposure to the target language (TL) but also a quality one. This is described by Mehisto et al. (2008, p. 26) as that the natural use of language "replicates the conditions to which infants are exposed when (acquiring) their first language". Krashen (1985) also assumes that since all the learners may not be at the same level of linguistic competence, natural

input is necessary to ensure that each learner could receive some 'i + 1' input that is suitable for their current 'i' level. Again, this reflects that CLIL is characterised by the natural input of language, which is meaningful and relevant for the learners. In this sense, "language acquisition could run its course naturally under meaningful and…positive conditions" (Dalton-Puffer, 2007, p. 259). In the meanwhile, a differentiated approach to input is also activated in CLIL, which reflects Roiha's (2014) idea that differentiated CLIL is important for both academically strong and weak students in order to achieve the anticipated learning outcomes.

Language input is interwoven with language output (Saville-Troike, 2012), and in the Output Hypothesis, Swain (1985) assumes that the opportunities to produce language are highly valuable, which could help learners enhance language proficiency, gain feedback on their language accuracy, experiment with language structures and improve metalinguistic awareness. According to Meyer (2013, p. 301), "task design is at the heart of every CLIL lesson", and output is triggered by meaningful activities that could allow learners to use language in an engaging manner. This is in line with Muñoz's (2007, p. 20) argument that CLIL "learners should have numerous and varied opportunities to speak and write in different contexts and with different aims", and "to achieve this, activities and teaching methods have to be adapted to the needs and interests of the learners …(and) be meaningful." Thus, it could be said that the studied CLIL programme was well designed with a variety of learning tasks that invited learners to produce output by stimulating the use of the TL and providing learners with ample opportunities to communicate in the TL, do written practice and receive feedback. These indicators reflected effective teaching performance in a CLIL classroom (de Graaff, Koopman et al., 2007). Also, the different output activities allowed of the differentiation of products created by learners to demonstrate their levels of mastery of the learning content, and this could be an effective way for teachers "to meet students' needs in a mixed-level classroom" (Chien, 2012, pp. 281-282).

Additionally, the positive content learning outcomes cannot be separated from the discussion of interdisciplinary teaching, which "refers to the coming together of two or more subject areas…without unduly disturbing subject area boundaries" (Kansas State Board of Education, 1995, p. 40). Some participants mentioned in the interviews that the combination of CE with law had greatly triggered their interest in the subject area. For example, Student Huang from the H-Group claimed:

> *We regard this programme as a legal English one, and indeed, the teacher has placed lots of emphasis on English teaching. However, we have also gained a lot of knowledge about American law. I mean, content learning is not at the cost of language learning…Such a combination is interesting, and I can always find a purpose to learn the language and the content.*

This view was also confirmed by Student Liu in the L-Group, who believed:

> *I feel more motivated to learn. Law, which is my major, provides a meaningful context for me to develop my English proficiency. This combination is much better than isolation. Although the teacher has focused on English, content learning has also occurred effectively and naturally.*

It is interesting to note that some interviewees mentioned that this CLIL programme was language-driven, and

this was also found in classroom observation which demonstrated that the proportion of language learning outweighed that of content learning in the manner that the teacher had mostly used the content subject to theme CET. This reflects the view that although CLIL has a dual focus in nature, there could be different levels of emphasis on language and content in practices (Coyle et al., 2010). However, content learning was not at the expense of language learning. Instead, both quantitative and qualitative data showed that the participants were engaged in the learning of subject content. This could be explained from the perspective of cross-curricular teaching, which could provide "a meaningful way in which students (could) use knowledge learned in one context as a knowledge based in other contexts" and enable them "to apply, integrate and transfer knowledge" (Darn, 2006, p. 2). This is in line with Aslan's (2016, p. 1798) assumption that when language is combined with any other subject area in a meaningful and thematic way, learners would be invited to "acquire, communicate and investigate worthwhile knowledge in depth", "apply what they learn in meaningful and 'real world' contexts" and "participate and learn regardless of ability".

Harrop (2012) as well as many other scholars (Coylet et al., 2010; Hüttner & Smit, 2014) states that CLIL is an entitlement for learners of all abilities and explains this from the perspective of the relationship between language and content as well as the 'double processing' in CLIL. First of all, the fact that CLIL has a dual focus on both content and language renders the relationship between them transparent to teachers and learners. "Language is seen as a tool for learning and one that needs scaffolding and progression as much as content" (Harrop, 2012, p. 64). In this regard, CLIL brings to light the language issues in the subject area in a manner that is frequently overlooked in subject area education, facilitating teachers' awareness of learners' needs and effective instructional practices to ensure comprehension (Harrop, 2012). According to Lasagabaster (2008), this approach could address inclusion and equity in education, and if education is regarded as language socialisation of learning, CLIL could also have a potential social equalising effect. Secondly, the 'double processing', which "refers to how CLIL learners process speech in a foreign language in order to take in new information while at the same time integrating the new knowledge in an existing corpus" (Harrop, 2012, p. 64), helps learners to engage more actively with the learning materials to overcome linguistic barriers and deepen understanding of content knowledge (Dalton-Puffer, 2007). In this sense, CLIL has the potential to achieve language and content learning for all in a context that both language development and content development are scaffolded as much as possible. For learners of differing capabilities, the learning outcomes may be different, while they are generally positive and reflect inclusion in education.

## Limitation

A major limitation of this study is that it only presented a small picture of the studied CLIL programme. Within the border of Chinese academia, Hu (2021) asserts that any CLIL study should involve the examination of performance evidence, affective evidence, process evidence and materials and task evidence to extend the scope of the CLIL research agenda and deepen the understanding of CLIL application. This idea is in line with the one put forward by Coyle et al. (2010) who argue that the lack of any type of evidence from research may create suspicion about the effectiveness of CLIL. However, the study presented in this paper only focused on

performance evidence with respect to learners' language and content learning proficiency and learning process evidence integrated with some materials and task evidence, with affective evidence characterised by learners' emotions under CLIL instructions excluded from the research. In this sense, it cannot be said that this study was comprehensive enough, and more thorough research designs should be taken into account in future studies to collect and examine a wider range of empirical evidence.

## Conclusion

The inspiring findings generated from this study that learners of different academic abilities made remarkable achievements both in language and content learning in an online CLIL context shed light on the evidence against an elitist view to online learning and CLIL and establish some proof for the possibility to include and engage all learners in the classroom. At least from this study, it can be concluded that CLIL has the potential to benefit all learners and that when special attention is paid to their different needs and levels of learning, inclusion can be achieved. This offers some implication, especially to educators, that their teaching practice can have a substantial role to play in facilitating effective learning and inclusive education for all. However, a critical view should be held towards elitism either in online learning or CLIL, as this study was organised only in a specific educational context, out of which there may be rather different findings and implications. Elitism is a complex issue either in general education or language education. More empirical research studies in various educational contexts are in need to further approve of elitism or dispel it with evidence to be established on any possible effects that teaching exerts.

A final note of this paper is that in the time of COVID-19 crisis, students may have interrupted learning experiences. This is still the case in many parts of the world which are still suffering from this pandemic. When disadvantaged learners who are experiencing more barriers to learning than their advanced peers cannot receive the support they need, the disadvantage gap between them may widen. However, differentiation is the key to effective and inclusive learning, and the implication drawn from the research findings is that when teaching instructions are well organised and implemented in response to learners' different needs in a heterogeneous class, all of the students, whether they are academically strong or not, can be embraced, developed, supported and inspired to be their best with an equal chance to succeed in online learning.

## Acknowledgements

I would like to extend my special thanks to the participants of this research as well as their tutors who granted me permission to study the educational programme presented in this text. Upon the completion of this paper, a new COVID-19 outbreak occurred in my hometown, Yantai, China. Words cannot express how truly grateful I am to those who are working on the frontlines of this pandemic and putting our lives before their own.

## References

Apsel, C. (2012). Coping With CLIL. Dropouts from CLIL Streams in Germany. *International CLIL Research Journal, 1*(4), 47–56.

Ary, D., Jacobs, L. C., & Sorensen, C. K. (2010). *Introduction to Research in Education* (8th ed.). Wadswroth.

Aslan, Y. (2016). The Effect of Cross-curricular Instruction on Reading Comprehension. *Universal Journal of Educational Research, 4*(8), 1797-1901.

Beck, D., & Beasley, J. (2020). Identifying the differentiation practices of virtual school teachers. *Education and Information Technologies, 26*(1), 2191-2205.

Börjesson, M., & Broady, D. (2016). Elite Strategies in a Unified System of Higher Education. The Case of Sweden. *L'Année sociologique*, (66), 115-146.

Boulkroun, F. (2020). Short-Term and Long-Term Effects of Differentiated Grammar Instruction: A Quasi-Experimental Study. *Journal of Arabic Language Sciences and Literature, 12*(2), 2100-2124.

Bozkurt, A., Jung, I., Xiao, J., Vladimirschi, V., Schuwer, R., Egorov, G., Lambert, S. R., Al-Freih, M., Pete, J., Olcott, D., Rodes, V., Aranciaga, I., Bali, M., Alvarez, A. V., Roberts, J., Pazurek, A., Raffaghelli, J. E., Panagiotou, N., de Coëtlogon, P.,…Paskevicius, M. (2020). A global outlook to the interruption of education due to COVID-19 Pandemic: Navigating in a time of uncertainty and crisis. *Asian Journal of Distance Education, 15*(1), 1-126.

Braun, V. & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101.

Broca, A. (2016). CLIL and non-CLIL: differences from the outset. *ELT Journal, 70*(3), 320-331.

Bruton, A (2011). Is CLIL so beneficial, or just selective? Re-evaluating some of the research. *System, 39*, 523–532.

Bruton, A. (2015). CLIL: Detail matters in the whole picture. More than a reply to J. Hüttner and U. Smit (2014). *System, 53*, 119-128.

Cañado, M. L. P. (2019). CLIL and elitism: myth or reality? *Language Learning Journal, 48*(1), 1-14.

Cañado, M. L. (2020). What's hot and what's not on the current CLIL research agenda: Weeding out the non-issues from the real issues. A response to Bruton (2019). *Applied Linguistic Review*, 1-21

Cen, X., Sun, D., Rong, M., Fekete, G., Baker, J. S., Song, Y., & Gu, Y. (2020). The Online Education Mode and Reopening Plans for Chinese Schools During the COVID-19 Pandemic: A Mini Review. *Frontiers in Public Health, 8*, Article 566316.

Chen, G., Huang, H., & Meng, D. (2020). Sampling survey and analysis of online class learning situation of college students in 5 universities in Nanjing during the outbreak of COVID-19. *Chinese Journal of Rehabilitation Medicine, 35*(9), 1095-1097.

Chen, Y., Dong, Y., Li, Z., Wang, D., Wu, X., & Li, Q. (2020). Influence of Online Course on University Students Activities under the Epidemic Situation of COVID-19. *Creative Education Studies, 8*(6), 1007-1014.

Chien, C. W. (2012). Differentiated Instruction in an Elementary School EFL Classroom. *TESOL Journal, 3*(2), 280-291.

Coyle, D., Holmes, B., & King, L. (2009). *Towards an integrated curriculum – CLIL National Statement and Guidelines*. The Languages Company.

Coyle, D., Hood, P., & Marsh, D. (2010). *CLIL: Content and Language Integrated Learning*. Cambridge University Press.

Dalton-Puffer, C. (2007). *Discourse in Content and Language Integrated Learning.* John Benjamins B.V.

Darn, S. (2006). *Content and Language Integrated Learning (CLIL): A European Overview*. Izmir University of Economics.

de Graaff, R., Koopman, G. J., Anikina, Y., & Westhoff, G. (2007). An Observation Tool for Effective L2 Pedagogy in Content and Language Integrated Learning (CLIL). *The International Journal of Bilingual Education and Bilingualism, 10*(5), 603-624.

Devkota, K. J. (2021). Inequalities reinforced through online and distance education in the age of COVID-19: The case of higher education in Nepal. *International Review of Education*. Advanced online publication.

Gebhard, J. (1999). Seeing Teaching Differently through Observation. In J. G. Gebhard & R. Oprandy (eds.), *Language Teaching Awareness: A Guide to Exploring Beliefs and Practices* (chapt. 3). Cambridge University Press.

Gulati, M. (2020). Online Learning in China during COVID-19: Some Lessons. *SSRN*. Advanced online publication.

Hao, M. (2016). Constructing the Concept of Elite Education on the Platform of Higher Education Popularization. *Non-State Running Science & Technology Enterprises*, (1), 256.

Harrop, E. (2012). Content and Language Integrated Learning (CLIL): Limitations and possibilities. *Encuentro, 21*, 57-70.

Hu, H. (2021). Mapping out the Future: A Proposal of A CLIL Research Agenda in China. *Journal of English Language Teaching and Applied Linguistics, 3*(9), 1-12.

Hu, Y., Nie, J., & Gu, X. (2021). From equity of opportunity to equity of development: a comparative analysis of large-scale online education between urban and rural K-12 schools in China during COVID-19. *Distance Education in China, 5*, 13-21.

Huang, R. H., Liu, D. J., Tlili, A., Yang, J. F., Wang, H. H., et al. (2020). *Handbook on Facilitating Flexible Learning During Educational Disruption: The Chinese Experience in Maintaining Undisrupted Learning in COVID-19 Outbreak*. Smart Learning Institute of Beijing Normal University.

Huang, Z., Wang, T., & Li, X. (2016). The political dynamics of educational changes in China. *Policy Futures in Education, 14*(1), 24-41.

Hüttner, J., & Smit, U. (2014). CLIL (Content and Language Integrated Learning): the bigger picture. A response to A. Bruton. 2013. CLIL: Some of the reasons why ... and why not. *System, 44*(1), 160–167.

Jash, A. (2020). COVID-19 and Xi Jinping's Fight and Win Strategy: Implications for China. *CLAWS Journal, 13*(1), 141-156.

Jin, S. (2000). Educational System for the Elite Key Schools. *Educational Research and Experiment*, (4), 18-21.

Johnstone, C. J., Schuelka, M. J., & Swadek, G. (2020). Quality Education for All? The Promises and Limitations of the SDG Framework for Inclusive Education and Students with Disabilities. In A. Wulff (ed.), *Grading Goal Four: Tensions, Threats, and Opportunities in the Sustainable Development Goal on Quality Education* (pp. 96-115). Brill.

Kansas State Board of Education. (1995). *Kansas Visual Arts Curriculum Standards*. National Endowment for

the Arts (NFAH).

Kotob, M., & Arnouss, D. (2019). Differentiated Instruction: The Effect on Learner's Achievement in Kindergarten. *International Journal of Contemporary Education, 2*(2), 61-92.

Krashen, S. D. (1985). *The Input Hypothesis: Issues and Implications*. Longman.

Lasagabaster, D. (2008). Foreign Language competence in CLIL courses. *The Open Applied Linguistics Journal, 1*, 31- 42.

Liebenberg, J., Huisman, M., Mentz, E. (2015). Software: university courses versus workplace practice. *Industry & Higher Education, 29*(3), 221-235.

Liu, D. (2018). China urgently needs to build an elite education system. *University Education Science,* (2), 4-8.

Liu, R., & Chen, Y. (2020). Research on the Present Situation and Effectiveness of College Students' Internet-learning — Based on Network Survey Data during Epidemic Prevention and Control. *Journal of Jinhua Polytechnic, 20*(4), 21-30.

Loreman, T., Deppeler, J., & Harvey, D. (2011). *Inclusive education: Supporting diversity in the classroom* (2nd ed.). Allen & Unwin.

Ma, H. (2020). Analysis of the Current Situation and Countermeasures of Online Teaching in Higher Vocational Colleges under the Epidemic Crisis — Taking a Higher Vocational College in China as an Example. *Innovation and Practice of Teaching Methods, 3*(9), 94-96.

McDougald, J. (2015). Teachers' attitudes, perceptions and experiences in CLIL: A look at content and language. *Colombian Applied Linguistic Journal, 17*(1), 25-41.

Mahyoob, M. (2020). Challenges of e-Learning during the COVID-19 Pandemic Experienced by EFL Learners. *Arab World English Journal, 11*(4), 351-362.

Marsh, D. (2003). *The relevance and potential of content and language integrated learning (CLIL) for achieving MT+2 in Europe*. European Commission DG EAC.

Mehisto, P. (2007). What a school needs to consider before launching a CLIL program: the Estonian experience. In D. Marsh & D. Wolff (eds.), *Diverse Contexts-Converging Goals: CLIL in Europe* (pp. 61-77). Peter Lang.

Mehisto, P., Marsh, D., & Frigols, M. J. (2008). *Uncovering CLIL: Content and Language Integrated Learning and Multilingual Education*. Macmillan Education.

Mewald, C. (2007). A comparison of oral foreign language performance of learners in CLIL and mainstream classes at lower secondary level in Lower Austria. In Dalton-Puffer & Smit (eds.), *Empirical Perspectives on CLIL Classroom Discourse* (pp. 139-178). Peter Lang.

Meyer, O. (2013). Introducing the CLIL-Pyramid: Key Strategies and Principles for Quality CLIL Planning and Teaching. In M. Eisenmann & T. Summer (eds.), *Basic Issues in EFL Teaching* (2nd ed., pp. 295-313). Universitätsverlag Winter GmbH.

Muñoz, C. (2007). CLIL: Some Thoughts on its Psycholinguistic Principles. *Revista española de lingüística aplicada*, (1), 17-26.

Paran, A (2013). Content and language integrated learning: Panacea or policy borrowing myth? *Applied Linguistics Review, 4*(2), 317–342.

Pladevall-Ballester, E. (2014). CLIL subject selection and young learners' listening and reading comprehension

skills. *International Journal of Applied Linguistics, 26*(1), 52-74.

Reimers, F. M., Amaechi, U., Banerji, A., Wang, M, Barajas, F., Bell., P., Chadwick, C., Eadie, S., Gunawan, J., Haq, A. N., Henderson, P., Hinckley, K., Hinrichs, C. R., Hosain, M., Hu, B., Jaramillo, M., Kenyon, C., de Léon Mazariegos, M. J., Martinez, D.,…Yoon, V. R. (2021). *An educational calamity: Learning and teaching during the Covid-19 pandemic*. Authors.

Roiha, A. (2014). Teachers' views on differentiation in content and language integrated learning (CLIL): Perceptions, practices and challenges. *Language and Education, 28*(1), 1-18.

Rosen, D., Aronson, B., Litt, D. G., McAlinn, G. P., & Stern, J. P. (2017). *An Introduction to American Law* (3rd ed.). Carolina Academic Press.

Rumlich, D. (2014). Prospective CLIL and non-CLIL students' interest in English (classes): a quasi-experimental study on German sixth-graders. In R. Breeze, C. Martínez Pasamar, C. Llamas Saíz, & C. Tabernero Sala (eds.), *Integration of Theory and Practice in CLIL* (pp. 75- 95). Rodopi.

Saville-Troike, M. (2012). *Introducing Second Language Acquisition*. Cambridge University Press.

Sperber, A. D. (2004). Translation and validation of study instruments for cross-cultural research. *Outcome Assessment, 126*(1), 124-128.

Swain, M. (1985). Communicative Competence: Some Roles of Comprehensible Input and Comprehensible Output in its Development. In S. Gass & C. Madden (eds.), *Input in Second Language Acquisition* (pp. 235-252). Newburry House.

Tomlinson, C. A. (2000). Differentiation of Instruction in the Elementary Grades. *ERIC Digests*, 1–7.

Van Mensel, L., Hiligsmann, P., Mettewie, L., & Galand, B. (2020). CLIL, an elitist language learning approach? A background analysis of English and Dutch CLIL pupils in French-speaking Belgium. *Language, Culture and Curriculum, 33*(1), 1-14.

Xiao, C. (2010). Build a first-class country with elite education. *Dushi Jiajiao*, (11), 16-17.

Wang, C. (2007). Education and Social Inequality in China Elite Groups Perpetuating Their Privileged Status. *China Perspectives*, (3), 110-116.

Wang, X., & Yang, W. (2016). The Effects of Elitism Education System on China's Social Mobility. *Economic Survey, 33*(4), 115-120.

Wu, L., Yan, K., & Zhang, Y. (2020). Higher education expansion and inequality in educational opportunities in China. *Higher Education, 80*, 549-570.

Xue, E., Li, J., Li, T., & Shang, W. (2020). How China's education responses to COVID-19: A perspective of policy analysis. *Educational Philosophy and Theory*, 1-13.

Zeng, Y. (2013). Medical education must be elite education. *Journal of Medical Research, 42*(3), 1-3.

Zhang, W., Wang, Y., Yang, L., & Wang, C. (2020). Suspending Classes Without Stopping Learning: China's Education Emergency Management Policy in the COVID-19 Outbreak. *Journal of Risk and Financial Management, 13*(3), 55-60.

Zheng, Y., & Cheng, L. (2008). Test review: College English Test (CET) in China. *Language Testing, 25*(3), 408-417.

Zhou, N., Zhu, M., You, B., Gao, X., Wang, W., & Zhao, L. (2007). *Educational Reform and Curriculum Change in China: A Comparative Case Study.* International Bureau for Education.

Zuo, M.,Ma, Y., Hu, Y., & Luo, H. (2021). K-12 Students' Online Learning Experiences during COVID-19: Lessons from China. *Frontiers of Education in China, 16*(1), 1-30.

Zydatiß, W. (2012). Linguistic Thresholds in the CLIL Classroom? The Threshold Hypothesis Revisited. *International CLIL Research Journal, 1*(4), 17–28.