

The effect of first and second language use on question types in English medium instruction science classrooms in Hong Kong

Jack Pun and Ernesto Macaro

Abstract: The role of the first language (L1) in second language (L2) learning has been widely discussed and researched in the Second Language Acquisition (SLA) literature. Benefits of L1 use have been found for vocabulary acquisition, and in facilitating comprehension thereby alleviating student anxiety. Many teachers appear to consider L1 use as a necessary option in L2 pedagogy. Disadvantages of L1 use are argued to be that it reduces exposure to the L2 thereby not enabling more implicit forms of L2 acquisition. Considerably less research has been carried out in English Medium Instruction (EMI) classrooms, surprisingly, as these are predicated on the benefits of massive exposure to the L2. This study investigated the effect of L1 and L2 use on teacher question types and interaction patterns in science lessons in early EMI and late EMI schools in Hong Kong. Our findings show that in late EMI, where there was much greater use of the L1, teachers tended to use higher order questions and were more interactive. In early EMI, despite students having received four more years of greater exposure to English instruction than late EMI students through the use of English as the vehicular language, teachers asked lower order questions and the interaction was more monologic. The implications for policy and pedagogy are discussed.

Keywords: English Medium instruction; L1 use; questioning; Science classrooms; Second Language Acquisition

Introduction

The role of the first language (L1) in second language (L2) learning has been widely debated and researched in second language acquisition (SLA) literature. The phenomenon is also sometimes referred to as ‘codeswitching’ and more recently as ‘translanguaging’. However, we adopt what we consider to be a less value-laden term: ‘use of the L1’.

Research that has focused on the functions of teacher L1 use suggests it is used primarily to communicate meaning when there is a communication breakdown (Hobbs, Masuo, and Payne 2010; Kharma and Hajjaj 1989). However, researchers working in the area of input and interaction (Krashen 1985; Long 1981; Swain 1985) have demonstrated that through modified input and meaning negotiation, communication problems can be overcome without automatically triggering L1 use.

Classrooms where the learning of the L2 is not the primary objective but either a stated combination of content and language (as in CLIL) or where the learning of the L2 is a by-product of content learning (English-medium Instruction, EMI) have only very recently begun to emerge (Gierlinger 2015; Lasagabaster 2017; Lo 2010). Given that these classroom settings are predicated on benefits accruing from massive exposure to the L2 whilst at the same time mastering the subject content, it is appropriate to ask the extent to which the L1 is beneficial (Lin 2006; Lin and Lo 2017).

Hong Kong has adopted EMI in various guises in the past four decades. The adoption of EMI is linked to Hong Kong's history as a British colony from the mid-19th century until its return to Chinese sovereignty in 1997. As most primary schools are Chinese Medium of Instruction (CMI) and all government-funded universities adopt EMI, the Medium of Instruction (MOI) at secondary level has been a controversial topic. The debate has moved through three stages (Poon 2013): (1) the British government's *laissez-faire* policy from 1946 to 1997; (2) the compulsory mother-tongue policy that followed the handover of Hong Kong to China from 1998 to 2009 which allowed only 114 secondary schools to adopt EMI in teaching content subjects while the remaining 307 schools used CMI; (3) the fine-tuning of the MOI policy since 2010 which allows individual schools greater autonomy to decide their MOI in different

content subjects at senior levels (Legislative Council 2009). The policy is in part a response to public pressure after evidence of a decline in the English proficiency of graduates and an increase in the strongly expressed parental desire for children to become fluent in English. Under this initiative, CMI schools who met certain criteria (e.g. the students' learning ability, the teachers' capability, the requirements of individual subjects and school resources) are allowed to switch to EMI to become 'late EMI' schools (CMI from grades 7 to 9 and partial EMI from grades 10 to 12), in addition to the 'early EMI' schools where full EMI instruction was conducted in all grades.

Whilst there is some evidence that EMI has been beneficial in improving English proficiency (Lo and Murphy 2010 and see below), it is not yet clear what impact it has had on the mastery of subject content. Mastery of subject content can be explored in a number of ways. For example it can be done by directly testing students' learning or indirectly by examining the quality of the interaction taking place. Quality of interaction comprises among other things the number and frequency of higher order questions and is the focus of this study.

Literature review

Use of the L1

Results of studies which have measured the amount of L1 use by teachers vary enormously. Rolin-Ianziti and Brownlie (2002) examined the use of L1 in 1st year French courses in Australia and found a range from 0 to 18% L1 use, with an average of 8.8%. In a study of six pre-service teachers in French secondary classrooms in England, Macaro (2001) found a low amount of L1 use but a considerable range from 0 to 15.2%, with an average of 6.9%. In Japan, Kaneko (1992) found that teachers and students used between 51% and 74% L1 in senior secondary classrooms. De La Campa and Nassaji

(2009) in German classrooms in Canada found 11.3% of L1 used by teachers; the more experienced using 9.3% L1, fewer than that of novice teachers, with 13.2%. There therefore appears to be no particular pattern according to country or educational setting.

In an EMI setting, in Hong Kong, Lo and Macaro (2012) compared L1 use in Grade 9-10 lessons in EMI and MOI-switching secondary schools. There was more L1 use in the late EMI schools during the 1st year of MOI switching, with an average 98.5% of L1 use, most of the interaction occurring in L1 and the occasional English terms being used. There was less L1 use in the 2nd year, with 50.5%. In contrast, only 2.5% of L1 use was found in EMI schools.

Wannagat (2007) measured L1 and L2 use in EMI and CLIL history classrooms in Hong Kong and Germany respectively. Hong Kong students were exposed to far more teacher L1 than in Germany, at 42.6%, compared 0.3%. Students also used more L1 in EMI classrooms, with 44.4%, than in CLIL classrooms, with 1.5%.

SLA studies which have observed the functions of L1 use have reported that teachers used it for a wide range of purposes (e.g. Jafari and Shokrpour 2013; Polio and Duff 1994; Raschka, Sercombe and Chi-ling 2009; Üstünel and Seedhouse 2005), including: giving instructions about tasks, translating words and phrases, checking comprehension, providing feedback, maintaining discipline, and explaining grammar. Once again no obvious pattern of L1 use related to functions is discernible.

In EMI classrooms, Luk and Lin (2015), categorized functions into three main types: 1) 'unpacking the field' (e.g. giving grammar instruction, providing translation, remedying students' apparent lack of comprehension); 2) 'negotiating relationships' (e.g. classroom discipline; motivating students); 3) 'marking shifts of focus, topic, task or stage of the lesson' (e.g. giving instruction about a task, highlighting topic change). These functions are very similar to those in SLA lessons, with the difference perhaps

being that in content lessons words unfamiliar to students may be a good deal more abstract and are labels for more complex constructs, as in ‘condensation’ in the Luk and Lin study.

Studies which have investigated the impact of L1 use in EMI classrooms

Lin (2006) examined classroom interaction in junior EMI science classrooms and found that L1 use promoted more teacher-student interaction. Turnbull, Cormier, and Bourque (2011) also observed a positive and significant relationship between L1 use and the complexity of students’ oral outputs in science. Students in an L2-only group produced less complex utterances, whereas those in a classroom where L1 use was allowed, responded with greater complexity. The authors nevertheless advocate caution against overuse of L1.

A similar perspective was offered in a relatively early study by Butzkamm (1998) in a German secondary school history lesson. Here all the L1 use was by students when providing responses and when stuck for an L2 word. The teacher maintained the L2 throughout. Butzkamm argues that this level of L1 use demonstrates that ‘the second language is firmly in the saddle as the working language’ (p. 95) with the L1 used ‘sparingly and unobtrusively’ (p. 95) and in contrast to the ‘permissiveness which crowds out the foreign language whenever the talk becomes serious’ (p. 96).

Nonetheless in other parts of the world teachers show concerns as to whether content can be learned entirely through the L2 for students with an inadequate English level (see Macaro 2018 for an overview). In Africa, teachers reported their primary challenge was to ‘smuggle’ the L1 into EMI classrooms to overcome poor student English proficiency (Probyn 2006).

Quality Interaction and higher order questions

Researchers in the field of science education have proposed that skilful questioning by teachers can lead to meaningful learning. Questioning is also explored more generally by Dalton-Puffer (2007). Here learning requires interactive teaching styles that encourage students to express their views and ask questions, and promote a learning culture based on discussion (Yip and Tsang, 2007; Yip, Tsang and Cheung 2003). Hong Kong science teachers identified as skilful in questioning in Yip's studies (1999, 2001, 2004) used a wide range of question types (e.g. lower order, higher order). A lower-order question is one where students are asked to recall factual information from memory or explain phenomena by regurgitating previously learned content. A higher order question aims to promote students' higher cognitive skills such as analysis, synthesis and evaluation. For example, a teacher might ask a higher-order question to encourage students to synthesise their existing knowledge in order to design a method of investigation (Anderson et al. 2001). Yip's group of studies indicated that they tended to use recall questions with lower cognitive demands due to time constraints and the pressure to cover the examination syllabus (Yip 2004), as well as the low English proficiency of students, and the poor communication skills of some science teachers themselves (Yip and Tsang 2007).

One approach to judging whether a lesson offers sufficient opportunities for students to interact is the Initiation-Response-Feedback sequence (IRF) based on the seminal work of Sinclair and Coulthard (1975), in which a teacher Initiates an interaction (e.g. asking a question), the student Responds (e.g. giving an answer), and the teacher then provides some Follow-up feedback (e.g. acknowledging the answer was

correct). An extension of the IRF is where the teacher provides the kind of feedback which encourages students to give a further response (Mortimer and Scott 2003). It can therefore be argued that lessons with a more extended IRF produce greater levels of scaffolded learning than those consisting mainly of simple IRF patterns.

Using the notion of simple and extended IRF patterns, Lo and Macaro (2012) compared Arts and Science subject lessons in Grade 9 and 10 between early and late EMI schools in Hong Kong. They found that the number of extended IRF sequences observed in science classrooms was lower than those in arts classes, suggesting lower order questioning in the former. They also found that when students did interact with their teachers, these exchanges usually consisted of simpler IRF sequences (e.g. yes/no responses). This suggested that the teachers were less able to modify their questions to elicit students' additional responses after students failed to give the expected answer to the original question.

Yip, Coyle and Tsang (2007) examined the impact of EMI on science teaching in junior secondary EMI classrooms. Their observations indicated that science teachers appeared to lack the communicative competence in L2 to explain abstract and complex science concepts and were thus reluctant to use higher-order questioning and conceptual change questions. Yip (2004) examined science teachers' questioning techniques, identifying 10 types of teacher questions that evaluated conceptual changes in their students, including lower order (recalling, explanation); higher order (analysis, evaluation, synthesis); and motivational questions and questions leading to conceptual change (eliciting, challenging, extending, application). When promoting science learning, the EMI teachers tended to lower the cognitive demands by asking recall questions to accommodate students' inability to communicate science ideas in English. As a result, students only gave short answers.

Many of these researchers in science education have adopted Mortimer and Scott (2003)'s framework which is made up of five components: teaching purposes, content, communicative approach, pattern of discourse, and teacher interventions. These components are employed when examining science classroom interactions. When students are given the opportunity to use the language by engaging with subject-specific terminology, they learn by activating thinking processes to understand and apply the knowledge of science until it becomes a part of learners' repertoire (Coyle 2007). They are participating in the instructional register of science, co-constructed with their science teachers (Llinares, Morton and Whittaker 2012). Yip's studies document the role of the L1 in teachers' questioning and their findings suggest a possible relationship between the teacher's questioning technique, L1 use and the quality of science learning. A similar relationship was observed in Lo and Macaro's (2015) study where the content-subject teachers appeared to be able to modify their questions more effectively, when the L1 was used, in order to scaffold the interaction so that students could answer cognitively demanding questions. However, there is a lack of research illustrating how content-subject teachers use the L1 or the L2 to formulate questions in the context of early EMI and late EMI schools. The above discussion led us to formulate the following research questions:

Research Questions

- 1) Are there differences in L1/L2 use in early EMI and late EMI schools in Hong Kong?
- 2) Is there a relationship between greater use of the L1 and the use of higher order questions asked by the teacher?
- 3) Does greater L2 use lead to less interactive teaching?

Methodology

We adopted a mixed-method design which measured the amount of L1 use, the interaction patterns and the types of science questions asked by the teachers in the two contexts of early and late EMI.

The 485ⁱ secondary schools in Hong Kong that deliver six years of secondary education according to the Hong Kong Diploma of Secondary Education curriculum (HKDSE) form the population studied and we arrived at our sample by a stratified sampling technique according to the school's official medium of instruction (MOI), geographical location, students' academic performance, English level and school curriculum. The sampled schools were representative of Hong Kong secondary schools in that they were closely matched in socio-economic status and students' academic ability.

The participants were teachers and students from grades 10 and 11(aged 15 to 17) EMI science classrooms (physics, chemistry and biology). The research obtained ethical approval and consent forms were provided for teachers and students. Nineteen teachers (3 females and 16 males), with teaching experience ranging from 3-35 years, participated in this study. They had graduated from universities in Hong Kong with a Bachelor's degree in a relevant science discipline and a postgraduate diploma in education. All the selected teachers delivered science content that matched their disciplinary training in science. None of the participating teachers had any formal qualifications in teaching science in English or had participated in EMI teacher training. Currently, there is no English proficiency requirement for academic subject teachers in Hong Kong. The participating teachers reported that they had high levels of general English proficiency. As the lessons recorded and the examples provided in this paper show, the participating teachers were not themselves struggling to convey scientific content in English.

Lesson observation

A total of thirty-three 40-minute lessons were observed on different days and at different times of the day to enhance the validity and reliability of the data collected. All lessons were video-recorded, transcribed and translated, where necessary, from Cantonese to English by the first author. Two independent bilingual researchers performed a back translation, comparing that version against the original recordings. Coded lesson transcripts were carefully read and repeatedly compared, sorted, recoded and explored for connections among coded segments. For accuracy, about 10% of the data were independently coded by two trained bilingual raters according to the analytical and established a high inter-rater reliability of ($k > 0.8$) using Cohen's Kappa coefficient (Robson 2002).

Data analysis

The data were analysed for: percentage of interaction time, proportion of L1 and L2 use, and teachers' question types (classified as high-order questions and low-order questions, following Yip, 2004: 78). Yip, building on the work of Bloom's taxonomy (Anderson et al. 2001) and that of others in the field of science education, classifies as 'lower order' questions those that require students to *recall factual information* and questions that require students to *explain a phenomenon or process*. He classifies as 'higher order' questions those that involve:

- *analysis* by (for example) 'breaking down material into its component parts. This may involve comparing parts and identifying their relationships'
- *evaluation* by (for example) judging the value and implications of material
- *synthesis* by (for example) 'relating existing concepts to construct a new idea of formulate a new pattern (Yip 2004: 78)

In order to ensure reliability of coding for question types we carried out coding independently and established a high inter-rater reliability of ($k > 0.8$). The amount of L1/L2 use was measured by the number of words in L1 divided by the total number of words in each lesson. The data was entered into SPSS software manually and descriptive statistics were produced. Variables were not normally distributed and thus the Mann–Whitney test was used for inferential statistics.

Findings

Before describing our findings according to the research questions, we report findings for all aspects of the interaction between these two types of EMI classes for which *there were no statistical differences*. We found no differences in the amount of time spent in oral interaction as a proportion of total lesson time, no significant differences in the average length of teacher turns, nor the average length of student turns. This then sets the scene for examining our research question variables.

Our first research question was concerned with whether there were differences in the amount of L1 used between the early and late EMI lessons. Table 1 shows that teachers in the early EMI classes used on average 1.5% of L1 compared to 26% in the late EMI classes. The gap between students' use of L1 was not significant but their use of L2 was. In other words, whilst in the early EMI classes it could be said that exposure to English was the overwhelming characteristic, the same cannot be said for the late EMI lessons. Moreover, there was very wide variation in the late EMI lessons with some teachers using as much as 98% L1. Here most of the lesson was in Cantonese with generally only technical words being provided in English (see examples of interaction below). We found no statistically significant differences between grades 10 and 11 (Table 1) so it is reasonably secure to treat them together.

[Table 1 Here]

Our second research question was concerned with the types of science questions that the teachers asked and whether these were related to L1/L2 use. Table 2 provides an analysis where early and late EMI lessons are compared in terms of question types: lower-order vs higher-order questions. In our data more lower-order questions were asked in early EMI schools (41%>29%). In the late EMI schools teachers used more higher-order questions (55%>32%). Thus if we combine the finding on L1 use with the finding on question type we can conclude that there is a strong relationship between the use of the L2 and the use of lower-order questions and vice versa; it appears that greater use of English as the medium of instruction may lead to teachers restricting their use of higher-order questioning. Our third research question was concerned with the relationship between greater use of L2 and less interactive teaching. In order to illustrate this relationship we now provide excerpts from the dataset.

[Table 2 Here]

Examples from the dataset:

We provide Excerpt 1 to illustrate our general finding that the use of L1 produced the kind of interaction in general, and teacher questioning in particular, that is considered facilitative of students' science learning.

[Excerpt 1 Here]

Excerpt 1, taken from a late-EMI grade 10 class, provides examples of questioning around the topic of specific immune response. We note that virtually all the interaction is in Cantonese and only technical terms (e.g. antigen) are in English. In turn 1, the teacher begins by asking students two lower-order questions: explanation (asking for an explanation of 'non-specific') and recalling (asking them to give a label to a concept ['what is specific']). The teacher introduces the concept of immunity, and helps

students distinguish specific immune responses from general immune response. In turn 3, the teacher asks a higher-order question (evaluation) to help students evaluate the concept of pathogens, by considering its function in our human immune system. In turn 5, while the student provides an answer about antigen, the teacher continues to ask another lower-order question (explanation) to further seek the student's response. In the same turn, the teacher asks a second higher-order question (analysis). The students would need to break down the concept of antigen into its component parts. This may involve comparing parts and identifying their relationships with the human body. Thus, this higher-order question aims to encourage students to consider where these antigens belong. As we can see in this excerpt, there is a considerable amount of interaction between the teacher and students in the form of Initiation-Response-Feedback sequences, in exploring their understanding of specific immunity responses. The teacher also uses a range of interventions to promote science learning, as described by Mortimer and Scott (2003), such as focusing on a particular student's idea by zooming into a particular student response (e.g. antigen); expanding on what a student has said; further exploring the ideas; asking for clarification of student ideas; and reviewing student responses with the whole class.

We provide Excerpt 2 to illustrate the contrast with Excerpt 1. It is taken from an early-EMI grade 10 chemistry class. It is entirely in English. The teacher introduces students to the concept of Haloalkane in a long single explanation, with very minimal interaction with the students. The teacher makes a false start by asking a higher-order question (explanation). It appears that the students do not follow what the teacher has asked and the teacher ignores one student's answer and continues with his explanations. The rest of the transcript evolves in an authoritative and monologic style, with mainly comprehension checks, and teacher's long chunks of explanations of technical terms.

[Excerpt 2 Here]

Excerpt 3, taken from an early-EMI grade 11 biology class provides examples of teaching on the topic of mitotic cell division. As in excerpt 2, the entire lesson is in English. The teacher introduces the concept of cell division. He checks with students their understanding of cell division in bacteria at the microorganism level. We can see that although the teacher begins with a higher question (explanation), he subsequently asks a number of lower-order questions (recalling) to recall students' understanding of the key concepts related to cell division (e.g. bacteria, fungi, yeast, genetic variation). The use of questions is limited to the lower-order questions, mainly recall type questions, to check their understanding by provide a technical term they learned from previous lessons. The use of higher-order questions appears to be absent.

In turn 1, although the teacher asks a higher-order question (explanation), the students do not follow what the teacher has asked and the teacher ignores it and continues with lower-order questions (recalling). The lack of higher-order question probably results from the teacher not making use of students' prior experiences and existing knowledge. A consequence is that students only provide single word answers.

[Excerpt 3 Here]

Excerpt 4 is taken from a late-EMI grade 11 biology class on the topic of DNA and demonstrates examples of interactive teaching. In this excerpt, the teacher uses both English and Cantonese to develop students' understanding of the structure of chromosomes. The teacher introduces the concept of chromosomes and asks the students to think about the number of pairs of chromosomes in the human body. In turns 1, 3, 13, 15, we see the frequent use of lower-order questions (e.g. recalling) to elicit what student learned about the number of chromosomes and their origins. For example,

in turn 1 and 3, the teacher asks how many chromosomes come from father's and mother's cells. The student can only respond with the correct number.

In turns 5, 7, 8, 10, we see the teacher using a number of higher-order questions (e.g. evaluation, analysis, synthesis) to challenge students' cognitive thinking by asking them to calculate the number of chromosomes in different types of cells in sexual production. This is different from turns 1, 3, 13, 15 because the students would need to base their answers on the given information, to relate existing concepts (e.g. the 23 pairs of chromosomes in gametes) to formulate a new pattern (e.g. 46 pairs of chromosomes in zygotes). This would require students to review and resolve inconsistent views of the number of chromosomes in different types of cells (e.g. gamete vs. zygote) during sex reproduction.

From turn 7 to 12, the teacher uses Cantonese to mediate an abstract concept of haploid cells by using an analogy of the collection of stamps.

[Excerpt 4 Here]

Discussion and Conclusion

Our study sought to establish whether the use of the L1 in science led to teachers asking different types of questions (higher order or lower order). These differences in question types have been widely discussed in the science education literature with higher order questions, combined with more extended IRFs, being advocated as those which lead to greater understanding of scientific concepts. In this study we were not trying to prove causality in the direction higher order questions leads to student learning. That would need objective outcome measures. We were simply positing two things: L1 use appears to be associated with higher order questions; higher order questions and quality interaction more generally have in the past been shown to lead to better learning.

Having first established that in late EMI teachers were indeed using far more L1, we found that those teachers' use of questioning did promote more interactions, as indicated by greater number of higher-order questions. In this way, teachers appear to have been able to offer opportunities to promote higher-order thinking (e.g. evaluation, analysis and synthesis). Two excerpts from the late EMI classrooms show questioning techniques that help students in their science learning by breaking down complex concepts into their component parts, comparing and identifying their relationship. We also saw teachers asking higher-order questions to help students relate existing concepts and construct new ideas from their previously held concepts. Late EMI teachers and students were able, it seems, to interact using the rich semantic resources in their mother-tongue to help students establish a meaningful relationship between their existing and new knowledge in their cognitive structure to form an interconnected concept in Science (Yip 2004). Our findings are in line with (Lo and Macaro's) earlier finding that when schools that had used Chinese medium hitherto, but now adopted EMI, the quality of the interaction changed and in particular it became less interactive.

In our study, late EMI teachers' questions were challenging enough to encourage students to resolve inconsistent views on science topics and develop scientific concepts themselves. When the science topic was more difficult and abstract, teachers in the early EMI schools seemed to use lower-order questions to guide students to only a basic understanding of the science topic.

So one conclusion might be that flexibility in choosing the medium of instruction allows teachers to ask cognitively challenging questions to promote higher-order thinking. However, we have to ask the question why it is that even after four years of learning science through the medium of English were teachers in early EMI still using lower order questions. A possible reason may be that teachers believe that the

early EMI students' English proficiency is so limited that they cannot respond to higher-order questions and thus teachers have to use lower-order questions just to superficially cover the content of the syllabus. An alternative explanation may lie in the pedagogy itself which, from grade 7 onwards, has failed to promote progress in the learning of academic English whereas the science content learning required has progressed. In other words, the early EMI classroom began as monologic and was characterized by lower order questions and has stayed that way rather than progressed as the students' English has improved.

For the late EMI schools, the use of Cantonese still seems to be the primary language of pedagogical interaction and whilst this interaction may be more meaningful we have to ask whether, in some cases, we are really dealing with English as the vehicular language at all. True, the use of higher order questions appears to be furthering content understanding, but it is not furthering the acquisition of academic English, something which students will need in an educational setting which only offers content learning at university through the medium of English. A study by Lin and Morrison (2010) showed that when Hong Kong switched to CMI it had a significant negative impact on students' receptive vocabulary knowledge when they reached higher education with more than 17% of them having a level of English vocabulary not considered sufficient to operate at tertiary level in EMI and with only 1.4% of CMI students having a satisfactory level of productive vocabulary.

It thus seems that a solution (which runs counter to current ministry guidelines) to allow unfettered use of the L1 is *not* going to promote quality EMI. Rather, Hong Kong research needs to establish where the actual problem lies in the early years of middle school (lower secondary) and if it is really the case that teachers remain convinced that only lower order interaction is possible or whether they need strong

support to help develop students' English proficiency whereby both considerable exposure to the L2 remains the norm whilst at the same time furthering deep learning of science.

Limitations

One of the possible limitations in this study is that the number of lessons observed was relatively small. In the research design phase, the eight participating schools and the lessons observed were carefully and purposively sampled so that they could be reasonably representative and 'comparable' of the two types of EMI schools. Therefore, the two types of EMI schools studied included a range of students with similar academic ability and English proficiency. A further limitation was that we were not able to administer objective tests of science knowledge to the two groups which would have helped to demonstrate an actual causal effect of L1 use and question types.

Implications for policy and practice

The significant difference in L1 use between early and late EMI schools support the claim from Lo (2010; 2015) that most Hong Kong content–subject teachers make use of Cantonese to enable their students to learn. The L1 can function pedagogically to assist L2 learning as Lin (2006) indicated in her bilingual pedagogical approach in Hong Kong junior science classrooms. Future research should examine the role of L1 and other possible factors (e.g. teacher's Language Awareness) in determining a smooth transition from non or partial to full EMI instruction.

Given that the role of L1 in content-teaching appears to be important, future studies should evaluate different amounts of L1 use in EMI programmes. This includes measures of the use of L1 in both language learning outcomes and content learning outcomes.

The current MOI policy in Hong Kong seems to create more flexibility for schools to choose their own MOI, aiming to end the segregation between CMI and EMI streams, but a more flexible policy does not imply all the problems can be solved. The government is expected to reveal its assessment of the progress of the fine-tuning MOI policy in 2017-18, eight years after its implementation. At that time, the government should provide successful case studies of effective implementation of EMI at classroom level. These should include recognition of teachers' decisions about practical teaching interventions such as the use of L1. Professional development programmes for EMI teachers are needed so that teachers can be equipped with the necessary pedagogical skills to integrate both content and language (Othman and Mohd Saat 2009). In these professional development programmes, EMI teachers should be trained to increase students' confidence in English skills, and develop their awareness of using suitable questioning techniques in promoting higher-order thinking.

Acknowledgments

We are very grateful to the anonymous reviewers who provided valuable comments and insights in the earlier drafts of this review. Our thanks also go to María Martínez-Adrián, Junkal Gutiérrez-Mangado and Francisco Gallardo-del-Puerto for their excellent editorial advice.

Disclosure statement

No potential conflict of interest were reported by the authors.

Funding

This work was supported by Swire Educational Trust.

References:

- Anderson, L. W., Krathwohl, D. R., Airasian, P., Cruikshank, K., Mayer, R., Pintrich, P., ... & Wittrock, M. 2001. *A taxonomy for learning, teaching and assessing: A revision of Bloom's taxonomy*. New York: Longman Publishing.
- Butzkamm, W. 1998. "Code-switching in a bilingual history lesson: The mother tongue as a conversational lubricant." *International Journal of Bilingual Education and Bilingualism* 1: 81-99.
- Coyle, D. 2007. "Content and Language Integrated Learning: Towards a Connected Research Agenda for CLIL Pedagogies." *International Journal of Bilingual Education and Bilingualism* 10: 543-562.
- De la Campa, J. C., and H. Nassaji. 2009. "The amount, purpose, and reasons for using L1 in L2 classrooms." *Foreign Language Annals* 42: 742-759.
- Gierlinger, E. 2015. "You can speak German, sir: on the complexity of teachers' L1 use in CLIL". *Language and Education* 29: 347-368.
- Hobbs, V., A. Matsuo, and M. Payne 2010. "Code-switching in Japanese language classrooms: An exploratory investigation of native vs. non-native speaker teacher practice." *Linguistics and education* 21: 44-59.
- Jafari, S. M., and N. Shokrpour, N. 2013. "E-role of L1 in ESP classrooms: A triangulated approach." *International Journal of English and Education* 2: 90-104.
- Kaneko, T. 1992. *The role of the first language in foreign language classrooms*. Unpublished doctoral dissertation, Philadelphia: Temple University.
- Kharma, N. N., and A.H. Hajjaj. 1989. "Use of the mother tongue in the ESL classroom." *International Review of Applied Linguistics* 27: 223-235.
- Krashen, S. D. 1985. *The input hypothesis: Issues and implications*. Addison-Wesley Longman Ltd.
- Lasagabaster, D. 2017. "I always speak English in my classes" In *Reflections on the use of the L1/L2 in English-medium instruction*. In *Applied Linguistics Perspectives on CLIL*, edited by A. Llinares and T. Morton, 259-75. Amsterdam: John Benjamins.
- Legislative Council 2009. *Fine-tuning the Medium of Instruction for Secondary Schools (Vol. 9)*. Hong Kong. <http://sc.legco.gov.hk/sc/www.legco.gov.hk/yr08-09/english/panels/ed/papers/ed0115cb2-623-1-e.pdf>. (accessed November 20, 2013).
- Lin A.M.Y. 2006. "Beyond Linguistic Purism in Language-in-education Policy and Practice: Exploring Bilingual Pedagogies in a Hong Kong Science Classroom." *Language and Education* 20: 287-305.
- Lin, A.M.Y., and Y.Y. Lo. 2017. "Trans/languageing and the triadic dialogue in content and language integrated learning (CLIL) classrooms." *Language and Education* 31: 26-45.
- Lin, L. H., and B. Morrison. 2010. "The impact of the medium of instruction in Hong Kong secondary schools on tertiary students' vocabulary." *Journal of English for Academic Purposes* 9: 255-266.
- Llinares, A., T. Morton, and R. Whittaker. 2012. *The roles of language in CLIL*. Cambridge University Press.
- Lo, Y.Y. 2010. *What happens to classroom interaction patterns and teachers' code-switching behaviour when the medium of instruction changes? an exploratory study in Hong Kong secondary schools*. Unpublished DPhil dissertation. University of Oxford.
- Lo, Y. Y. 2015. "How much L1 is too much? – Teachers' language use in response to students' abilities and classroom interaction in CLIL." *International Journal of Bilingual Education and Bilingualism* 18: 270-288.
- Lo, Y. Y., and E. Macaro. 2012. "The medium of instruction and classroom interaction: Evidence from Hong Kong secondary schools." *International Journal of Bilingual Education and Bilingualism* 15: 29-52.

- Lo, Y.Y., and E. Macaro. 2015. "Getting used to Content and Language Integrated Learning: What can classroom interaction reveal?" *Language Learning Journal* 43: 239-255.
- Lo, Y. Y., and V. Murphy. 2010. "Vocabulary knowledge and growth in immersion and regular language-learning programmes in Hong Kong." *Language and Education* 24: 215-238.
- Long, M. H. 1981. "Input, interaction, and second-language acquisition." *Annals of the New York Academy of Sciences* 379: 259-278.
- Luk, G.N.Y., and A.M.Y. Lin. 2015. L1 as a pedagogical resource in building students' L2 academic literacy: Pedagogical innovation in the science classroom in a Hong Kong school. In *Multilingual education: between language learning and translanguaging*, edited by J. Cenoz., and D. Gorter, 16-34. Cambridge: Cambridge University Press.
- Macaro, E. 2001. "Analysing student teachers' codeswitching in foreign language classrooms: theories and decision making." *The Modern Language Journal* 85: 531-548.
- Macaro, E. (2018) *English Medium Instruction: Language and Content in Policy and Practice*. Oxford: OUP.
- Mortimer, E. and P. Scott. 2003. *Meaning Making in Secondary Science Classrooms*. Berkshire: England: Open University Press.
- Othman, J., and R. Mohd. Saat. 2009. "Challenges of Using English as a Medium of Instruction: Pre-service Science Teachers perspective." *The Asia-Pacific Education Researcher* 18: 307-316.
- Polio, C. G., and P.A. Duff. 1994. "Teachers' language use in university foreign language classrooms: A qualitative analysis of English and target language alternation." *The Modern Language Journal* 78: 313-326.
- Poon, A. Y. K. 2013. "Will the New Fine-Tuning Medium-of-Instruction Policy Alleviate the Threats of Dominance of English-Medium Instruction in Hong Kong?" *Current Issues in Language Planning* 14: 34-51.
- Probyn, M. 2006. "Language and Learning Science in South Africa." *Language and Education* 20: 391-414.
- Raschka, C., P. Sercombe, and H. Chi-Ling. 2009. "Conflicts and tensions in codeswitching in a Taiwanese EFL classroom." *International Journal of Bilingual Education and Bilingualism* 12: 157-171.
- Robson, C. 2002. *Real World Research: a Resource for Social Scientists and Practitioner-Researchers*. (2nd ed.). Oxford: Blackwell.
- Rolin-Ianziti, J. and S. Brownlie. 2002. "Teacher use of learners' native language in the foreign language classroom." *Canadian Modern Language Review* 58: 402-426.
- Sinclair, J. and M. Coulthard. 1975. *Towards an Analysis of Discourse*. Oxford: Oxford University Press.
- Swain, M. 1985. Communicative competence: Some roles of comprehensible input and comprehensible output in its development, In *Input in Second Language Acquisition*, edited by S. Gass and C. Madden, 235-56. New York: Newbury House.
- Turnbull, M., M. Cormier, and J. Bourque. 2011. "The First Language in Science Class: A Quasi-Experimental Study in Late French Immersion". *The Modern Language Journal* 95: 182-198.
- Üstünel, E. and P. Seedhouse 2005. "Why that, in that language, right now? Code-switching and pedagogical focus." *International Journal of Applied Linguistics* 15: 302-325.
- Wannagat, U. 2007. "Learning through L2 – Content and Language Integrated Learning (CLIL) and English as Medium of Instruction (EMI)." *International Journal of Bilingual Education and Bilingualism* 10: 663-682.
- Yip, D.Y. 1999. "Implications of students' questions for science teaching." *School Science Review* 81: 49-53.

- Yip, D.Y. 2001. "Promoting the development of a conceptual change model of science instruction in prospective secondary biology teachers." *International Journal of Science Education* 23: 755–770.
- Yip, D.Y. 2004. "Questioning skills for conceptual change in science instruction." *Journal of Biological Education* 38: 76–83.
- Yip, D.Y., and W.K. Tsang. 2007. "Evaluation of the Effects of the Medium of Instruction on Science Learning of Hong Kong Secondary Students: Students' Self-Concept in Science." *International Journal of Science and Mathematics Education* 5: 393–413.
- Yip, D.Y., D. Coyle, and W.K. Tsang. 2007. "Evaluation of the Effects of the Medium of Instruction on Science Learning of Hong Kong Secondary Students: Instructional Activities in Science Lessons." *Education Journal* 35: 77–107.
- Yip, D.Y., W. K. Tsang, and S. P. Cheung. 2003. "Evaluation of the Effects of Medium of Instruction on the Science Learning of Hong Kong Secondary Students: Performance on the Science Achievement Test." *Bilingual Research Journal* 27: 295–331.

Funding information

This work was supported by the Swire Educational Trust. We would like to thank the schools, teachers and students involved in the study.

Table 1. Percentage of L1 and L2 use by teachers and students by school type and by grade

		Teachers		Students		Both	
<i>Between school</i>		L1	L2	L1	L2	L1	L2
early EMI		1.5%	54%	1%	33%	6.5%	93.5%
late EMI		26%	42%	8%	7%	53%	44%
<i>p</i> -value		.001***	.008***	.079*	p<.0001****	p<.0001****	p<.0001****
<i>Between grade</i>		L1	L2	L1	L2	L1	L2
early EMI	G10	<1%	52%	<1%	33%	<1%	<100%
	G11	3%	56%	1%	33%	9%	91%
Sig.		.738	.834	.448	.753	.665	.665
late EMI	G10	42.5%	41%	6.5%	7.5%	52.5%	47.5%
	G11	21%	42%	12%	6%	56%	44%
<i>p</i> -value		.388	.601	.440	.665	.712	.268

Note: * $p < .05$; ** $p < .01$; *** $p < .001$; **** $p < .0001$

Table 2. The types of science teachers' questions by school

Types	By school	
	early EMI	late EMI
Lower order	41%	29%
Higher order	32%	55%

Excerpt 1. a late-EMI grade 10 class (English translation of the Cantonese is in brackets in italics)

The teacher asks the students to explain some scientific terms

Turn	Speaker	Text	Types of Question	Functions of the question
1	Teacher:	點解我地叫 non-specific? 點解我地咁樣叫? 這樣叫 specific? 什麼叫 specific? <i>(Why do we call these non-specific? Why would we call them like this? This would be specific? What is specific?)</i>	<i>Why do we call these non-specific? (low-order Q)</i> <i>What is specific? (low-order Q)</i>	explanation recalling
2	Student:	Target 果D所有一般病原體，所以唔係咁 specific。 <i>(Target the normal general pathogens, do this is why we called them non-specific.)</i>		
3	Teacher:	Target. 對呀，佢地 targeting the pathogens. <i>(Target, that's right, they are targeting the pathogens.)</i>		
		點解我地講佢有能力去 targeting? <i>(Why would we say that it has the ability of targeting?)</i>	<i>Why would we say that it has the ability of targeting? (high-order Q)</i>	evaluation
4	Student:	因為佢地本身有 antigen，可以被 immune system 用來分辨外來的病原體和其他有害物質。 <i>(Because they itself have antigen, which they can be distinguished by the immune system whether they have pathogens or other toxic substances.)</i>		
5	Teacher:	Antigen. 點解? 點解可以? <i>(Antigen, what does it mean? Why does it help?)</i>	<i>what does it mean? (low-order Q)</i>	explanation
		Whose antigen belongs to? 係唔係屬於身體定 pathogen? <i>(Whose antigen belongs to? Does it belong to our body or pathogen?)</i>	<i>Does it belong to our body or pathogen? (high-order Q)</i>	analysis
6	Student:	Pathogen. 因為佢地本身有 antigen，讓身體的 immune system 去分辨。 <i>(Pathogen. Because they have antigen themselves, and this can be distinguished by our immune system.)</i>		
7	Teacher:	An antigen is a type of protein.		

Excerpt 2: an early EMI Grade 10 class

Turn	Speaker	Text	Types of Question	Functions of the question
1	Teacher:	Today we will look at Haloalkane first. From our syllabus, we have only one reaction for our haloalkane, in between alcohol. We look at our book. It is on page 12. This is the haloalkane. Haloalkane is, there is a halogen atom bonded to the carbon. And then, we look at this structure first. For halogen atom, which is electronegative, it tends to draw the bonded electrons towards itself. So X, Halogen atom, carry partial negative. And then the carbon become electron deficient. So it carry partial plus. And the covalent bonds resulted to be polar, polar covalent bond. And then the carbon, this carbon atom, is said to be electron deficient carbon center, atom, or you can call electron-love, want to have more electron. And this center, carbon atom more likely attract negative charged particles. So for example, we have the negative-charged hydroxide ion. It carries negative charge and most probably it will go from left hand side and attract the carbon, rather than right hand side approach to the negative X halogen. For haloalkanes, the reactions for haloalkanes, it occurs at the carbon positive site. And then the particle going to carbon carry negative charge. Example, Haloalkanes react with hydroxide ion, and then hydroxide ion the O form covalent bond with the ion. Bond to the carbon. Why heating is required?	Why heating is required? (low-order)	Explanation
2	Student:	Slow....		
3	Teacher:	The heating is required because at room conditions, the reaction very slow. Heat it up for faster reaction. Look at the paragraph, the hydroxide ion is a aqueous solution. Haloalkane is not very good at soluble in water. Slightly soluble in water. At the beginning, there are two layers, the aqueous layer and the haloalkane layer, organic layer. And then after the reaction, if the reaction complete, the alcohol can really form hydrogen bond with the water. So at the end of the reaction, the two layers disappear to forming one layer, single layer. Because of alcohol, water-soluble. Okay?		

Excerpt 3: an early EMI grade 11 class

Turn	Speaker	Text	Types of Question	Functions of the question
1	Teacher	What do you think? What microorganisms can also grow here? Apart from bacteria. You have learned	What do you think? (low-order) What microorganisms can also grow here? Apart from bacteria. You have learned (lower-order)	explanation recalling
2	Student	Fungi		
3	Teacher	Fungi? Yes, can you give me an example of fungi? You have learned a very small, okay, unicellular fungi. What's that?	What's that? (lower-order)	recalling
4	Student	Yeast		
5	Teacher	Yeast, yes. We can grow yeast here okay? This microorganism, this colony, is derived from the division, the repeated cell division of one cell okay, into many many cells by mitotic cell division. It is the asexual reproduction of the bacteria. And the bacterial colony almost contains this number of bacteria, this amount that means about millions of that bacteria, okay, in this colony. So that you can know directly, use your naked eye, to observe this, okay?	We can grow yeast here okay? So that you can know directly, use your naked eye, to observe this, okay?	(comprehension check) (comprehension check)

Excerpt 4: late-EMI grade 11

Turn		Text	Types of Question	Functions of the question
1	T	幾多 chromosomes come from your father's sperm? (<i>How many chromosomes come from your father's sperm?</i>)	<i>How many chromosomes come from your father's sperm?</i> (lower-order Q)	recalling
2	S	23.		
3	T	幾多 come from your mother? (<i>How many come from your mother's</i>)	<i>How many come from your mother's</i> (lower-order Q)	recalling
4	S	23.		
5	T	23. So for each pair, 點解一個 come from the father and one come from the mother? (23. So for each pair, <i>why does one</i> come from the father and one come from the mother?)	<i>why does one</i> come from the father and one come from the mother? (higher-order Q)	analysis
6	S	因為要從爸爸媽媽果邊去取各 23 條的染色體去形成新細胞體 (<i>Because we need the 23 chromosomes from dad and mom to form the new cell</i>)		
7	T	因為你形成由 fertilization, i.e. the fusion of the sperm and egg nuclei. Okay, 一個從爸爸黎(<i>one comes from the father</i>), the paternal chromosome, 藍色果個, 例如, 一個從媽咪黎(<i>the blue one, for example here, and one comes from your mother</i>), the maternal chromosome. 所以人每一 pair 都有 (So in each pair, there are) blue and red long chromosomes, homologous chromosomes, homologous chromosomes. 即係話佢地有相同嘅 size 木相同形狀 (That means they have the same size and same shape) . One from the father, one from the mother. Homo means the same. 這裡有幾多(<i>Here, how many</i>) pairs are there? 有幾多 (<i>How many</i>) pairs of homologous chromosomes are there? (<i>Because you are formed from fertilization, i.e. the fusion of the sperm and egg nuclei. Okay, one comes from the father, the paternal chromosome, the blue one, for example here, and one comes from your mother</i>), the maternal chromosome. So in each pair, there are blue and red long chromosomes, homologous chromosomes, homologous chromosomes. That means they have the same size and same shape. One from the father, one from the mother. Homo means the same. Here, how many pairs are there? How many pairs of homologous chromosomes are there?	<i>Here, how many pairs are there?</i> (higher-order Q) <i>How many pairs of homologous chromosomes are there?</i> (higher-order Q)	evaluation evaluation
7	S	Two.		

8	T	Two pairs. 你睇到每一組有特別嘅 set, 例如 homologous chromosome, 我哋有郵票, 10 毫子、20 毫子, 個 7, 兩個 4, 就一 set. If we have two sets, we double it. 幾多 sets are there? (Two pairs. You can see that each contains a special set. For example, homologous chromosome, we have stamps of 10 cents, 20 cents, \$1.7, \$2.4, this is one set. If we have two sets, we double it. how many sets are there?)	how many sets are there? (higher-order Q)	analysis
9	S	Two sets. 因為要從爸爸媽媽分別人供應 23 條的染色體去形成兩 sets 新細胞體。 (Two sets. Because the new cell was formed from the 23 chromosomes originally from dad and mom's cells)		
10	T	Two. Our cells have two sets? 幾多(how many) sets do sperms and eggs have?	how many sets do sperms and eggs have? (higher-order Q)	synthesis
11	S	One sets. 因為要從爸爸媽媽分別供應 23 條的染色體, 所以係單倍體。 (One set. Because dad and mom's cells only supply 23 chromosome, so each of them should be a single set)		
12	T	One for each type. 我有五類郵票, 我每一個有一種, 我就有一 set. 如果有兩種就有兩 sets, 兩 sets 就叫 (If there are 5 types of stamps, and I have one for each type, I have 1 set. If I have two, I have two sets. Two sets, we call it) diploid cells. 如果得一個就 (If there is only one,) haploid. (One for each type. If there are 5 types of stamps, and I have one for each type, I have 1 set. If I have two, I have two sets. Two sets, we call it diploid cells. If there is only one, haploid.)		
13	T	Diploid two, haploid one. Okay? Diploid cells, haploid cells. 幾多(how many) types of haploid cells are there in humans?	how many types of haploid cells are there in humans (lower-order Q)	recalling
14	S	Only two.		
15	T	邊種 (Which kind) of haploid cells?	Which kind of haploid cells? (lower-order Q)	recalling
16	S	Sperm and Egg.		

ⁱ Source: Education Bureau, Hong Kong Special Administrative Region, Figures and Statistics on Secondary Education, <http://www.edb.gov.hk/tc/about-edb/publications-stat/figures/sec.html>.