Routledge
Taylor & Francis Group

Language and Education

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/rlae20

A comparison of content learning outcomes between Japanese and English medium instruction

Ikuya Aizawa, Heath Rose, Jim McKinley & Gene Thompson

To cite this article: Ikuya Aizawa, Heath Rose, Jim McKinley & Gene Thompson (2023): A comparison of content learning outcomes between Japanese and English medium instruction, Language and Education, DOI: 10.1080/09500782.2023.2238688

To link to this article: https://doi.org/10.1080/09500782.2023.2238688

Published online: 02 Aug 2023.

LANGUAGE AND EDUCATION
https://doi.org/10.1080/09500782.2023.2238688

Routledge
Taylor & Francis Group

A comparison of content learning outcomes between Japanese and English medium instruction

Ikuya Aizawa[a], Heath Rose[b], Jim McKinley[c] and Gene Thompson[d]

[a]School of English, University of Nottingham, Nottingham, UK; [b]Department of Education, The University of Oxford, Oxford, UK; [c]University College London, IOE Faculty of Education and Society, London, UK; [d]Department of Global Business, Rikkyo University, Tokyo, Japan

ABSTRACT

In the growing body of English medium instruction (EMI) research, few studies have directly compared the effects of medium of instruction on subject learning. This study compares direct measures of content knowledge and perceptions of knowledge acquisition for students studying Chemistry via English (n = 27) and Japanese Medium Instruction (JMI) (n = 26). Data were collected at a university in Japan where Chemistry courses were taught in both Japanese and English as part of a parallel program offering the same undergraduate curriculum in either of the two languages of instruction. An analysis was undertaken of students' learning outcomes measured by pre-post course content tests. These measurable test outcomes were triangulated with data from student interviews (n = 17) to identify differences in the learners' perceived experiences according to the medium of instruction. While the quantitative results revealed no significant overall differences in the adjusted post-test scores between EMI and JMI students, the qualitative data offered more detailed insight into the participants' perspectives of content learning, highlighting unique challenges faced only by the EMI group. Findings point to implications for educational provision in contexts where the global trend of EMI has largely been unaccompanied with research evidence on its cost-effectiveness.

ARTICLE HISTORY
Received 9 May 2023
Accepted 3 July 2023

KEYWORDS
content learning outcomes; English language proficiency; English medium of instruction (EMI); higher education; Japan; medium of instruction

Introduction

In line with global trends, higher education around the world has experienced a rapid surge in EMI programs in recent decades. Japan is no exception, and the introduction of EMI is seen as a gateway to successful internationalisation, mostly spurred by the recent government initiative of the Top Global University Project (TGUP) (see Aizawa and Rose 2019). While driving forces of EMI programs may appear increasingly attractive to stakeholders (e.g., mobility of international students and faculty members), research evidence thus far is worrying, with research highlighting concerns surrounding university students' insufficient English

CONTACT Ikuya Aizawa ikuya.aizawa1@nottingham.ac.uk School of English, University of Nottingham, Trent Building, University Park, Nottingham NG7 2RD, UK

language proficiency for EMI studies, and crucially, around its potentially disastrous effects on their academic outcomes (e.g., Pun et al. 2023). Observing this growing scepticism towards the global explosion of EMI, Macaro (2018) urged further research on the cost-effectiveness of EMI to investigate "what [students] potentially gain and what they potentially stand to lose" (p. 153). While university education, at its core, should offer students substantive disciplinary knowledge, more research evidence is needed to better understand the quality of EMI education, particularly whether "learning content through L2 English leads to at least as good learning of academic content as learning content through the students' L1" (Macaro 2018, p. 154). In the growing body of EMI research, few studies have addressed the extent to which there may be differences in learning the same content via the L1, versus learning it via English (see Macaro et al. 2018). To address this gap, the current study examines the potential costs and benefits for students learning Chemistry through EMI compared to L1 medium instruction.

Background to the study

Success in EMI

The way in which the "outcomes of EMI" are conceived varies significantly in the literature since the notion of what is "success" in higher education is highly contextual depending on local educational needs and policy objectives. Investigations of success in EMI are most often focused on evidence of knowledge gain. Some researchers operationalise EMI as a cost-effective educational approach to develop both language and content knowledge (e.g., Hu and Wu 2020). Conversely, others attribute EMI success most notably to the transformation of content knowledge, which can be measured, for example, by midterm and final test scores in EMI courses (e.g., Thompson et al. 2022). In the context of the current study, the Japanese Ministry of Education (MEXT) defines EMI as "courses conducted entirely in English, excluding those whose primary purpose is language education" (MEXT 2020, p. 58). Accordingly, the present study examined the conceptualisation of 'success in EMI' by examining students' acquisition of *content knowledge*, while acknowledging broader educational outcomes, such as perceived career prospects (e.g., Huang and Curle 2021). Recognising that educational outcomes in higher education may go further than just gaining substantial content knowledge, this study triangulated perception-based data with score gain data to explore both direct and perceived additional markers of success in EMI.

EMI outcome studies of disciplinary gains

While there have been several studies which have investigated medium of instruction differences on academic outcomes in Content and Language Integrated Learning (CLIL) classrooms at school level (see Murphy et al. 2020, for a review), research on academic outcomes of EMI at the university level has not been conducted to the same extent. One of the few studies into disciplinary gains of EMI via direct measures is Park's (2007) study, which examined the effect of EMI on content learning in South Korea. Based on true/false and open-ended questions, data were gathered from 51 EMI students enrolled in an introductory theoretical linguistics course. Although statistically significant gains in content knowledge were recorded by pre-post-tests at the start and the end of a two-hour lecture period, a

baseline group was not recruited to account for any potential initial difference between the EMI and L1 medium instruction groups.

The absence of a direct comparison group appears to be one of the most common limitations in many score gain studies, including EL-Daou and Abdallah (2019) who examined the impact of CLIL on undergraduate students' maths test performance in Lebanon (n= 21). While the study revealed that pre-test and post-test scores of a maths achievement showed disciplinary gains, it was not conclusive as to whether it was CLIL that actually led to the gains in maths knowledge without achieving a direct comparison group of L1 Lebanese medium instruction students.

Adopting a direct comparison group, Joe and Lee (2013) conducted an experiment to examine the effect of EMI on medical degree students' content learning in South Korea. A total of 61 L1 Korean students undertook both EMI and Korean-medium-instruction (KMI) lectures to compare lecture comprehension between the two conditions. While no significant differences in the post-test scores were found between the two groups, the design was limited by a lack of authenticity: rather than observing courses within the existing curriculum, the group of students had to attend one-off medical lectures specially designed for the study. One lecture was delivered in English followed directly by one delivered in Korean. The study would have achieved higher level of ecological validity if it had been part of regular EMI and KMI courses over a course of one semester.

Recognising the importance of direct measures, comparison groups and ecological validity, Dafouz and Camacho-Miñano (2016) explored Spanish students' academic performance in an EMI financial accounting course. Students' final grades (n = 383) measured on a 10-point grading scale collected over four academic years were compared using mean score differences, indicating no significant differences between the EMI and SMI students. While this was one of the most comprehensive studies to date, the outcome measure was based solely on the final grades; no qualitative data were obtained to capture potential contextual information regarding how the two groups were comparable in terms of key variables, such as prior content knowledge. Similarly, Civan and Coşkun (2016) compared the academic performance of EMI and Turkish medium instruction (TMI) students in terms of their semester point averages across nine academic disciplines, concluding that TMI was more effective than EMI. However, the study fell short of capturing comprehensive contextual information regarding how the course grades were comparable between the two learning conditions across nine disciplines.

In summary, there are few studies of student disciplinary gains. Further, of the small number of studies which have examined success in this manner, limitations include no comparable L1 learner group, a lack of authenticity in instruction and measures, the use of success or test measures that may not be comparable across the L1 and EMI groups, and no means of comparing the start points for each learner group.

Methodological concerns for EMI outcomes studies

Considering these conflicting findings, the studies reviewed above unveil several methodological concerns for evaluating the effectiveness of EMI. Consequently, this study improves on the previous literature in the following ways:

- The study adopted perception-based methods to achieve an in-depth, nuanced understanding of participants' viewpoints concerning their experience alongside content score gains obtained from direct measures.

- The study recruited JMI students as a baseline group which were comparable to their EMI counterparts to observe content learning differences.
- The study adopted a pre-post-test which was locally devised by a content expert while ensuring a high degree of ecological validity.
- The study adopted longitudinal research to follow students' content learning experiences over the course of one semester using the same pre-test and post-test while accounting for their prior content knowledge.

Largely, outcomes studies vary considerably in their research designs and methodologies (Hoare and Kong 2008), and consequently, there is yet to be a consensus in relation to research evidence into the effects of EMI. The question of how much and to what extent EMI enables or hinders content learning remains unresolved. Therefore, this study adopted a two-pronged, quantitative, and qualitative approach to unpack the complicated nature of students' content learning.

The study

The current study addresses the following two research questions:

1. To what extent does the achievement of content learning of EMI and JMI students differ?1a. To what extent does the achievement of content learning differ for students instructed via their L1 or L2?
2. How do EMI and JMI students perceive the medium of instruction to affect their content learning?

As the context of the current study was a Chemistry program within one university, a case study approach was adopted at the university course level. The reason for this focused approach also stems from Rose and McKinley (2018) observation that the implementation of the TGUP initiative is environmental and context-specific. While a case study approach allowed us to understand unique forms of EMI at the university, we acknowledge that the findings may differ from those at other universities and within other EMI programs.

Setting

This study was conducted in one of the TGUP participant universities. This institution was chosen because it offers parallel English and Japanese courses in a range of academic fields, which was the essential criterion for this study. Data were conducted in an introductory Chemistry course, which offers 70 minutes of lectures twice a week for 12 weeks; 21 lectures in total. The EMI and JMI courses were taught by L1 English and Japanese professors, respectively. The language of instruction was exclusive to each course; the EMI and JMI courses were conducted entirely in English and Japanese, respectively. This allowed us to draw comparisons between two different linguistic conditions while maintaining the content constant. Previous research suggests that content teachers may struggle to impart course content due to their lack of proficiency. As such, some may resort to code-switching or the distribution of bilingual word lists to aid the delivery of subject-specific lexicon (Tzoannopoulou 2014).

However, this was not observed in the current study, as both teachers delivered the courses in their respective L1 languages of instruction. As the course was introductory, each week, new topics such as organic chemistry and chemical equilibrium were introduced. This approach ensured that students fully grasped the foundational concepts before progressing further. The Department of Natural Sciences offers this course as a prerequisite for all first year Chemistry major undergraduate students. Due to the university's strategic plan to increase the number of EMI programs, this Chemistry course has been offered in both English and Japanese since 2017, which has in turn provided students with options to study the course in their preferred language of instruction. Accordingly, the course syllabus outlining the course's assessment criteria, learning objectives, and textbooks is the same between the two courses (one is in English and the other in Japanese).

Participants

An entire cohort of the EMI (n = 27) and JMI (n = 31) chemistry courses were initially contacted. Based on initial review of the consent form responses, five JMI students withdrew from the study, making a final total of 27 EMI and 26 JMI students, representing 100% and 84% participation rates. Table 1 summarises students' gender, L1 and IELTS scores based on the medium of instruction.

In recruiting interview participants, the study employed a maximal variation sampling strategy to "investigate a few cases but those which are as different as possible to disclose the range of variation in the field" (Flick 2009, p. 123). Participants were thus purposively selected as individual cases to offer the broadest picture of content learning experience in terms of L2 proficiency, content knowledge and language of instruction. Table 2 shows the selected cases of 13 EMI students representing a range of content knowledge and English proficiency levels. Table 2 also includes selected cases of four JMI students based on the levels of their content knowledge, enabling comparisons to be made between and within cases. Where the table contains no data in the cells, no suitable participant fitting this profile was available in the sample, perhaps due to a potential relationship between the two sampling variables of proficiency and content-learning achievement.

Content test

Academic prediction metrics such as course grades and GPAs can be methodologically problematic as they mirror numerous factors, including completion of coursework, assignments

Table 1. Information of the participants (n = 53).

Language of instruction		Gender			First language			English proficiency (IELTS)		
EMI	27	**Male**	7	26%	**Japanese**	16	59%	**6.5**	2	7%
		Female	20	74%	**Bilingual**	3	11%	**7.0**	7	26%
					English	3	11%	**7.5**	3	11%
					Other	5	19%	**8.0**	8	30%
								L1	7	26%
JMI	26	**Male**	18	69%	**Japanese**	26	100%	**5.5**	8	31%
		Female	8	31%				**6.0**	7	27%
								6.5	5	19%
								7.0	3	11%
								7.5	3	12%

Table 2. Purposive sampling of students (n = 13 EMI; n = 4 JMI).

English Proficiency (IELTS Scores)				
Content learning (Post-test scores)		**Low (1-6) (n = 4)**	**Medium (7-12) (n = 7)**	**High (13-17) (n = 6)**
EMI (n = 13)	**IELTS 6.5 (n = 2)**	Kento (Post-test score 5)	Rino (Post-test score 7.5)	
	IELTS 7.0 (n = 4)	Kei (Post-test score 4)	Aya (Post-test score 8.5)	
		Maru (Post-test score 1)	Miri (Post-test score 9.5)	
	IELTS 7.5 (n = 1)			Takako (Post-test score 12.5)
	IELTS 8.0 (n = 3)		Chinami (Post-test score 9.5)	Miki (Post-test score 15)
				Lisako (Post-test score 12.5)
	L1 English (n = 3)		Ken (Post-test score 8)	Emi (Post-test score 12.5)
			Sakie (Post-test score 7.5)	
JMI (n = 4)	**N.A. (n = 4)**	Naoko (Post-test score 3)	Shika (Post-test score 8)	Odake (Post-test score 15)
				Koike (Post-test score 16)

The classification of Pass (0–40%), Merit (40–70%) and Distinction (70–100%) was adopted from the Assessment Criteria of the university.

Table 3. Criteria and specifics of the content tests.

Criteria	Specifics
Topics	All content taught in the class from the beginning (Electrons in Atoms) to end of the term (Organic Chemistry)
Representative of the actual course tests	Test designed based on the final exam and homework problem sets
Number of items, item types	13 open ended questions (Appendix 1)
Test equivalence between the JMI and EMI course	Equivalent between the two courses; the test was first created by the EMI teacher and then translated by the bilingual JMI teacher into Japanese
Language choice made by students to answer the test	English for EMI (optional) and Japanese for JMI. No EMI students chose to use Japanese.
Scoring procedure	Test marked by the JMI and EMI Chemistry teachers for their respective courses, and scores entered into spreadsheet alongside other student demographic variables
Marking criteria	Regular course marking criteria

and attendance (Dafouz et al. 2014). Accordingly, in the present study, the EMI and JMI content teachers devised a locally developed content test as a direct proxy for academic achievement, which eliminated several uncontrolled factors involved in academic metrics. Furthermore, recognising the importance of measuring a baseline of students' prior knowledge at the onset of the course, a pre-test was adopted to account for students' prior knowledge to more accurately use post-tests data to measure gains in knowledge. Table 3 outlines the criteria used to devise the test:

Reliability and validity

The reliability of the content test items was inspected to assess the internal consistency of the measurement. The 13 post-test items yielded an Cronbach's alpha coefficient of $\alpha = .71$, which was above the cut-off line of .70.

Subsequently, construct validity of the test items was tested through item-total correlations between a self-devised test and another set of tests measuring the same sets of constructs. The paired samples correlations (n = 27 EMI) showed that the post-test scores were

positively correlated with the final course exam scores with a two-tailed test of significance, revealing a bivariate Pearson correlation coefficient of $r = 0.450$, $n = 27$, $p = .016$.

The content validity of the test instrument was separately examined based on expert ratings of each of the 13 items. The content teachers independently compared the content test against the final course exam and problem sets based on the learning aims to rate how closely they correlated. Consistent with the acceptable cut-off level of 70% for an interrater agreement, the content validity in this study was substantial with interrater agreement on 12 of the 13 items (92.3%), which added further evidence to its construct validity.

Semi-structured interviews

The interview schedule (Appendix 2) used in this study was adopted from Evans and Morrison (2011). It is comprised of four aspects of student content learning: content learning differences between EMI and L1 instruction; English language development through EMI; language-related challenges; and motivations towards EMI.

Procedures

The study was conducted over a period of one academic semester, consisting of 21 lectures in total. The pre-test and post-test were administered under the supervision of the content teachers during the first (Time 1) and last (Time 2) lecture of the semester. Students were offered up to 25 minutes to complete the test. Regarding the interview procedure, an average length of the interviews was 26.8 minutes (Mean 26.8; SD 6.4; Min 14; Max 36).

Data analysis

To investigate research question one, a Wilcoxon Signed Rank test was used to compare the pre-test and post-test scores within the groups; a Mann-Whitney U test to compare the pre-test scores between the groups; and a one-way ANCOVA to examine differences in covariate-adjusted post-test scores between the two groups. To code the interview data, the coding procedure entailed both deductive coding – directed by existing literature, and inductive coding – stemming from the empirical data, referred to as an abductive approach (McKinley 2020). A start list of themes was first constructed deductively based on previous EMI research on academic challenges (e.g., Bradford 2016; Ismailov et al. 2021; Macaro 2020) and success (e.g., Rose et al. 2020) and subsequently developed inductively from the interview data.

Findings

Directly measured content learning outcomes

The first research question examined the extent to which the achievement of directly measured content learning differed between the students enrolled in the EMI and JMI Chemistry courses. Findings are based on pre-post-test scores assessing their content knowledge. The test scores were calculated by counting the number of correct answers, ranging between 0

and 17. One EMI participant scoring 97.1% on the pre-test (an outlying value in excess of +4.08 from standard deviations) was removed. No other participant scored over 80% at pre-test, and no further cases were excluded.

Within groups comparisons: whole sample

Table 4 shows descriptive statistics for the pre-test and post-test scores for the whole sample and for the EMI and JMI samples separately. The EMI students' scores, on average, improved by 50.5% from *Mean* 1.80 (*SD* = 3.21) to *Mean* 10.40 (*SD* = 3.15); the JMI students' scores improved by 41.8% from *Mean* 2.54 (*SD* = 2.45) to *Mean* 9.64 (*SD* = 5.04). To examine the students' score changes between Time 1 and 2, their pre-test and post-test scores were compared. Due to the violation of normality, a Wilcoxon Signed Rank test was conducted, revealing statistically significant score gains for the EMI cohort, z = −4.545, p < .001 with a large effect size (r = .62) and for the JMI cohort, z = −4.463, p < .001 with a large effect size (r = .62). The median for the EMI and JMI groups on the pre-test scores increased from (Md = 0) and (Md = 2.0) to the post-test scores (Md = 12.0) and (Md = 9.75), respectively. Thus, both JMI and EMI groups improved their test scores from the start to end of their course.

Between groups comparison: whole sample post-test scores

The pre-test scores were first compared to disambiguate any pre-existing differences in content knowledge between the two groups at the start of the course. Due to the violation of normality, a Mann-Whitney U Test was carried out, revealing no significant difference in the pre-test scores of the EMI (Md = 0, n = 27) and JMI groups (Md = 2, n = 26), U = 259, z = -1.69, p = .09, r = .23. Therefore, the two groups were not significantly different in terms of their prior knowledge.

A one-way ANCOVA was subsequently conducted to examine whether there was a significant difference in the post-test scores between the two groups while controlling for their prior knowledge. While adjusting for the pre-test scores, there was no significant difference between the two groups on the post-test scores, $F(1, 50) = 1.415$, $p = .24$, partial $\eta 2 = .027$. The medium of instruction only explained 2.7% of the variance in the post-test scores. The pre-test scores explained 23% of the variance in the post-test scores, indicating that the total unexplained variance still amounted to 74.3%. Hence, the JMI and EMI groups were not significantly different in terms of their content knowledge at the end of the term when accounting for their prior knowledge.

Table 4. Descriptive statistics – EMI (n = 27), JMI (n = 26), total sample (n = 53).

	MOI	N	Min	Max	Mean	Median	SD
Pre-test scores	EMI	27	0	13.5	1.80	0.5	3.21
	JMI	26	0	7	2.54	2.0	2.45
	Total	53	0	13.5	2.16	1.0	2.87
Post-test scores	EMI	27	1	15	10.40	12.00	3.15
	JMI	26	2	17	9.64	9.750	5.04
	Total	53	1	17	10.01	9.500	4.60

Between groups comparison: L2-English only sample

The analyses above focused on an EMI sample that consisted of the L1 English (n = 7) and L2 English (n = 20) students. Subsequent analysis for Research Question 1a was conducted on L2 English students only to investigate whether there were any differences in content learning between L1 and L2 medium instructed students. After excluding both the English L1 students (n = 3) and English bilingual students (n = 4) from the sample, the same analyses were carried out with a separate data set of 20 EMI and 26 JMI students, totalling 46 students.

Within groups comparisons: L2-English only sample

Table 5 shows descriptive statistics for the test scores for the whole L2-English only sample and for the EMI and JMI samples separately. The L2-English students' scores, on average, improved by 51.5% from *Mean* 2.15 (*SD* = 3.21) to *Mean* 10.90 (*SD* = 3.15). A Wilcoxon Signed Rank test was conducted to examine score changes between Time 1 and 2, revealing a statistically significant change (n = 20), z = −3.926, p < .001 with a large effect size (r = .62). The median score for the EMI samples on the pre-test scores increased from (Md = 1) to the post-test scores (Md = 12.25). Thus, the L2 English students on average improved their test scores from the start to the end of the term.

Between groups comparison: L2-English only sample post-test scores

A Mann-Whitney U test first compared the pre-test scores, indicating no significant difference in the scores of the EMI (Md = 1, n = 20) and JMI groups (Md = 2, n = 26), U = 209, z = -1.164, p = .245, r = .172. Therefore, the two groups were not significantly different in terms of their prior content knowledge.

A one-way ANCOVA was adopted to test whether there was a significant difference in the mean post-test scores between the L2 English EMI and JMI groups while adjusting for the pre-test scores, suggesting no significant difference between the two on the post-test results, $F(1, 43) = 1.57$, $p = .217$, partial $\eta^2 = .035$. The medium of instruction only explained 3.5% of the variance in the post-test scores. The pre-test scores explained 21.5% of the variance in the post-test scores, highlighting that the total unexplained variance still amounted to 75%. Thus, the JMI and EMI students were not significantly different in terms of the post-test scores when controlling for their prior knowledge.

Based on these subsequent analyses, the finding was upheld that there was no significant difference in Chemistry knowledge according to the medium of instruction, even when analysis was confined to the students who studied through L2 English. Thus, learning

Table 5. Descriptive statistics – the EMI (n = 20), JMI (n = 26), total sample (n = 46).

	MOI	N	Min	Max	Mean	Median	SD
Pre-test scores	EMI	20	0	13.5	2.15	1	3.61
	JMI	26	0	7	2.54	2	2.45
	Total	46	0	13.5	2.62	2	3.16
Post-test scores	EMI	20	1	15	10.90	12.25	3.96
	JMI	26	2	17	9.64	9.75	5.04
	Total	46	1	17	9.64	9.5	4.78

through L2 English did not seem to compromise the EMI students' academic performance compared to that of their JMI counterparts.

Perceived content learning outcomes

While the quantitative results indicated no significant overall measurable mean differences in the adjusted post-test scores between the two groups, the qualitative data offered more detailed insight into the participants' (n = 13 EMI n = 4 JMI) perceptions of learning of Chemistry. Interview data revealed two key findings:

1. There were some differences between the EMI and JMI students in terms of their perceptions of ease and challenges with content learning, highlighting that there were some discipline-specific challenges that were exclusive to the EMI students.
2. The EMI students derived much benefit from EMI education and compensated the costs of content learning with these subsidiary benefits exclusively linked to EMI.

Each of these key findings are discussed in depth.

Discipline-specific challenges in EMI

The EMI students had difficulty with technical terms used in lectures. The teacher's use of a lot of unknown technical terms hindered their comprehension of the subject matter. Aya (IELTS 7.0, post-test 8.5) for instance had to relearn the English version of the Chemical elements in the periodic table which she had already learned at school, including "hydrogen" and "helium". She noted that "I learned Chemistry in Japanese at high school, so I have to memorise all the Chemical elements all over again in English. It's time-consuming." Similarly, Rino (IELTS 6.5, post-test 7.5) had previously learned the term "valence electrons" at school, but she did not realise when the same concept was introduced in her lecture as she did not recognise the term "valence" in English:

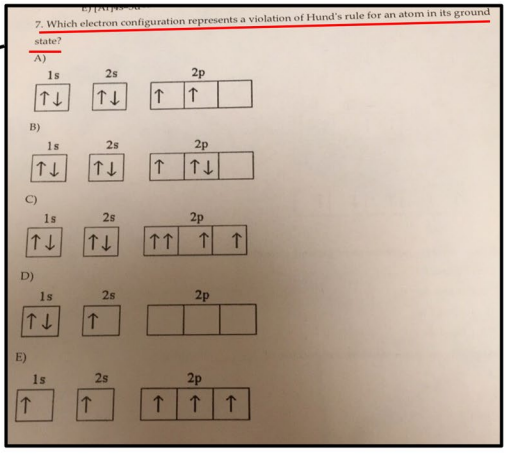
> **Rino (EMI):** I always read the textbook before my lectures but in one lecture I couldn't recognise the word "valence". It turned out that it was a basic concept covered in my high school Chemistry class. It'd be easier to tell the meaning in Japanese because this kanji corresponds to "value". The same kanji is also used in everyday life to mean "value". But the English term "valence" doesn't make sense.

In contrast, JMI student Naoko (IELTS 6.5, post-test 3) did not need to relearn any of these technical terms, claiming that "there are many technical terms I need to learn in the course. But I'm lucky that there are many words I still remember from my high school class". Another JMI student Koike (IELTS 5.5, post-test 16) was able to learn new technical terms quickly even when she studied them right before her lectures, noting that "I usually only read the textbook briefly before my lectures start but I understand almost all the technical terms when I hear them in lectures".

The EMI students also lacked a rich repertoire of vocabulary containing both scientific and everyday dual meanings used in the academic domain, such as "shell", "element", "activity", "stability", and "capacity". Maru (IELTS 7.0, post-test 1) shared an example problem

set from one of her lectures below which she could not complete due to unfamiliar terms used in the instruction:

Which electron configuration represents a violation of Hund's rule for an atom in its ground state?



Resource document shared by Maru:

Furthermore, lacking the domain specific vocabulary in English presented additional challenges for the EMI students to solve problem sets, such as calculating equations and electron configurations (e.g. $1s^2\ 2s^2\ 2p^6\ 3s^2\ 3p^1$) and determining the valencies of Chemical elements (e.g. Neon = 0, Phosphorus = 3). Kento (IELTS 6.5, post-test 5) identified the relation between his vocabulary knowledge and lecture comprehension by sharing a problem set from a task.

> **Kento (EMI):** When I solve a problem set like this, I come up with answers in Japanese first because I know all the terms better in Japanese. I don't know the word "plausible", "species", "notation". It takes more time to solve problem sets in English than Japanese.

It also became apparent that the EMI students' vocabulary-related issues stemmed from the non-transparent English language spelling system. Kento (IELTS 6.5, post-test 5) and Aya (IELTS 7.0, post-test 8.5) were unable to look up basic technical vocabulary (e.g., tetrahedral molecular) due to its difficult spelling:

> **Kento (EMI):** I had no idea what "tetrahedral" means and I didn't know the spelling either, especially the "dral" part. I also later realised that I had misplaced the "r" in the "tetra" part with "l". So I was not able to look it up in my dictionary.

> **Aya (EMI):** In one of the lectures, I didn't know how to spell "configuration" and "sodium" in "the electronic configuration of sodium" when I was taking notes. I ended up writing in katakana to substitute these English words.

In addition to technical terms, the EMI students also faced great difficulty with general academic vocabulary, reporting their frequent reliance on dictionaries to search for unfamiliar academic words. For instance, Kei (IELTS 7, post-test 4) noted that "I had to look up a word 'inhale' once when reading my textbook", which she described as "embarrassing" when she had discovered the meaning. In her view, it would have been a basic academic English word for L1 English speakers. Similarly, Miri (IELTS 7.0, post-test 9.5) noted that

if she studied in Japanese, she would have been more confident in her ability to infer the meaning of unfamiliar terms as she could analyse their Kanji radicals [components of the Japanese writing system that have semantic clues]. She provided an example that the word "hydrogen" is "sui-so" in Japanese which contains a water radical in the first Kanji character. Furthermore, Kento (IELTS 6.5, post-test 5) also had to frequently search for general academic words, such as 'involve', 'polarised', 'charged' and 'law' from the chapter 'Chemical Bonding'. Thus, Kei, Miri and Kento all had to look up the meanings of unfamiliar academic English words. Conversely, the JMI students, Odake (IELTS 5.0, post-test 15) and Shika (IELTS 6.0, post-test 6) only occasionally had to look up technical terms, rather than academic words. When doing so, it was merely to consolidate their content understanding:

> **Odake (JMI):** I don't need to use dictionaries to look up words but often use Google to search key terms to read more detailed descriptions of important concepts on Wikipedia. For example, I recently searched "Molecular Orbital Theory" because this was one of the key concepts of the section and I wanted to understand it better.

> **Shika (JMI):** I have difficulty memorising many key technical terms from each chapter, especially Chemical Bonding, but I make a list of key terms when I review my lecture notes at home. I don't have any issue understanding vocabulary used in the textbook apart from some of the difficult technical terms.

Thus, although students in general faced difficulty comprehending domain-specific vocabulary (e.g., molecular orbital theory), the EMI students expressed a greater level of challenge. Those who had previously learned Chemistry in Japanese, in particular, faced the significant additional difficulty of having to learn the English versions of school-level terms (e.g., "hydrogen", "helium") and dual meaning terms (e.g., "shell", "activity"), both of which severely impeded their performance of problem sets (e.g., calculating equations). The spelling system also impaired their performance (e.g., note taking). In contrast, the JMI students were advantaged by their prior vocabulary knowledge from school. As a result, the JMI students in most cases could easily overcome the problems through reading the textbook immediately before their lectures. Thus, the most apparent difference between the two groups appeared to be the type of vocabulary they struggled with and how frequently they relied on dictionaries.

Success in EMI

The primary aim of the Chemistry course was to achieve a sufficient level of content understanding; the EMI learners generally regarded linguistic improvement as a crucial achievement measure of EMI. Kento (IELTS 6.5, post-test 5) suggested the practical value of the EMI course as an additional platform to improve his English proficiency:

> **Kento (EMI):** I generally improve my English skills through participating in my academic English classes but the EMI course gives me extra practice opportunities. I can listen to the lectures and ask questions in English.

Despite possessing a sound foundation in the English language, Takako (IELTS 7.5, post-test 12.5) valued EMI in relation to the development of her speaking skills:

> **Takako (EMI):** I want to maintain my English level. I can practise speaking by asking questions and speaking to my classmates in English in the EMI Chemistry class.

Many of the EMI students recognised the English medium nature of the course as a driving force of their diligence as they were intrinsically motivated to learn through English, and willing to participate in the university's year abroad program. This factor, coupled with the necessity to achieve the required IELTS and GPA scores for the study abroad application criteria, served as a vital reason for their determination and perseverance:

> **Miri (EMI):** One reason why I chose this university is because I can take advantage of the exchange program to study abroad in England in my third year. I need high GPAs and IELTS scores. It's harder to study in English than Japanese but I'm studying as hard as I can to prepare for my year abroad.

While the participants expressed that the additional hours spent on EMI learning could be cumbersome and time-consuming, they also saw them as further opportunities to consolidate their disciplinary knowledge. Chinami (IELTS 8.0, post-test 9.5) was actively eager to maximise the repertoires of materials and resources. Aya (IELTS 7.0, post-test 8.5) also developed new learning strategies by studying through EMI (e.g., carrying a separate notebook to record unfamiliar words and technical concepts):

> **Chinami (EMI):** I always spend longer hours learning in English. I have to use dictionaries and extra resources. But I can learn more widely and extensively.

> **Aya (EMI):** In Japanese I always just spend five to ten minutes to skim through my textbooks, but in this EMI Chemistry class, I highlight key sentences and write down words I don't know using a separate notebook.

Thus, learning through EMI appeared to help students maintain strong motivation, diligence, and determination. The EMI students regarded the increased exposure to the content matter as opportunities to delve into more in-depth learning. One possible explanation for their increased motivation was due to the make-up of the EMI participant pool comprised of those with higher L2 proficiency and more positive attitudes towards EMI who volunteered to study through EMI.

The EMI students also enjoyed access to a wide range of learning resources, suggesting that gaining information bilingually was a distinctive feature of EMI content learning:

> **Miri (EMI):** When I do not understand concepts, I use all resources. I use textbooks from high school and additional materials from other [JMI] science courses. As long as I understand what I need to understand, I don't mind whether I use Japanese or English.

Further to this, due to the wider range of online learning resources available in English than their L1s alone, the students preferred English as the language of instruction. For example, Miki (IELTS 8.0, post-test 15) suggested that "there are [sic] more science information available written in English than Japanese on the internet". Lisako (IELTS 8.0, post-test 12.5) supplemented her studies with additional online materials and practice exercise sheets in English, noting "I like revising Chemistry concepts using a YouTube channel called Khan Academy which offers short lessons in the form of videos".

The EMI participants also identified intentions of an internationally oriented career as a crucial measure of EMI success. Rino (IELTS 6.5, post-test 7.5) was confident that successfully completing her English-medium degree would place her at an advantage in pursuing her academic career overseas:

Rino (EMI): I want to apply for a postgraduate degree in the US because there is a college which is highly rated for my discipline. Currently I learn less because of my poor English but in the long-term I don't need to study all the terms again in English.

Similarly, Lisako (IELTS 8.0, post-test 12.5) valued the international classroom landscape where her learning involved interactions with non-Japanese L1 students, serving as an additional advantage of EMI to accomplish her future employment goals. She noted that "I prefer studying in English because I like talking with my classmates from overseas. I want to work for an international organisation using English". Thus, the EMI students saw opportunities to interact with their peers and teachers in English as the additional benefits of EMI where they could develop intercultural competence and L2 proficiency.

Thus, while the EMI and JMI students had similar test outcomes, the pathways to those outcomes (challenges and success) differed substantively. The students saw various benefits of EMI despite the challenges, suggesting that success was found in more than just exam scores. To this end, the current study highlights the potential issues regarding simply computing and contrasting group mean scores. The average test results did not tell the whole story; a qualitative examination was crucial to access the experiences and perceptions of individual students and their learning experiences.

Discussion, implications, and future research

Content learning differences according to the medium of instruction

The findings of the quantitative evidence revealed no measurable differences in the development of students' understanding of Chemistry between the two groups when controlling for prior content knowledge. Such positive results concur with previous EMI research on educational outcomes (e.g., EL-Daou and Abdallah 2019; Dafouz et al. 2014), indicating that content learning through EMI in this research context did not appear to negatively affect the participants' academic performance. Nevertheless, the qualitative findings pointed to various academic challenges which were exclusive to those learning through L2 English, especially those with lower English proficiency. This suggests that EMI may negatively affect *some* participants' content learning experience in relation to L2 proficiency, whereby lower-level proficiency students may need to work harder to overcome certain academic challenges. Previous qualitative studies using interviews (e.g., Kim and Yoon 2018) indicate that students often express lower levels of satisfaction in content learning while also doubting the improvement of their overall L2 proficiency when their L2 proficiency is insufficient for EMI studies. Such findings situate this EMI study within a broader literature of EMI and CLIL research, which has shown that the L2 medium of instruction tends to benefit higher academic achievers and more proficient and motivated language learners (see Murphy et al. 2020, for a review).

That there were no differences between the two groups in terms of the average group scores of the content test is, thus, only a superficial finding. When other sources of data were investigated, more complex evidence emerged, suggesting that EMI may have a long-lasting deleterious effect on students' content learning (e.g., a reduction in the understanding of academic concepts). Indeed, many of the EMI students appeared to experience greater difficulties in reaching the same learning outcomes, suggesting that pathways (the

process) to success (the outcome) differed significantly between the two mediums of instruction. That is, EMI students needed to work substantially harder to overcome difficulties to achieve similar scholastic results.

Implications for EMI provision

Stakeholders involved in the delivery of EMI programs could evaluate the outcome of this study in different ways. From the viewpoint of the university, as no discernible differences in the content test scores between the two groups were found, this outcome may be a favourable result, especially when viewed in terms of students' academic performance. Conversely, teachers (and students) may be more interested in understanding individual students' learning experience (the process), which suggested that EMI may require some students to work harder to achieve success and overcome challenges. Both EMI students and teachers should focus on the ways in which they could alleviate the challenges associated with the process of content learning. The results highlight complexities surrounding who benefits and who loses when EMI is put into place, and that pathways to success may differ for many students as a result of the chosen medium of instruction. Either interpretation of the results points to the importance of academic support to students who are undertaking EMI.

Broadening conceptualisations of success in EMI

The current study largely adopted a narrow definition of EMI success, operationalising the outcome of EMI in relation to the score gains in students' content knowledge. To achieve a deeper understanding of the overall effectiveness of EMI, it is crucial to shed light on broader outcomes of EMI beyond content learning outcomes. Macaro's (2018) "cost-benefit" analysis of EMI posits that "a short-term, documented and carefully controlled negative effect could be acceptable if other benefits could be clearly detected" (p. 157).

While the acquisition of content knowledge is considered to be a benchmark of successful EMI in the current context of Japanese higher education (MEXT, 2018), EMI research points to a wider variety of purported benefits. Galloway et al. (2017) investigation of EMI in Japan and China, for example, highlighted that students and teachers alike attached a wide range of benefits to the phenomenon of EMI in higher education, which extended beyond language or content learning outcomes. Similarly, a survey of EMI in 52 countries revealed that the top reasons for students to undertake EMI were job opportunities and study abroad opportunities (Sahan et al. 2021). In the current study, students mirrored such perceptions, and saw post-graduation goals as long-term success markers of their EMI endeavours, rather short-term academic performance on end-of-term tests. This study highlights the fact that 'success' in EMI is complex, offering a range of additional educational benefits beyond academic measures such as test scores or course grades. Accordingly, it may be more pertinent to evaluate the value of EMI beyond language and content measures. This methodological suggestion is highlighted in Macaro's (2018) "cost-benefit" analysis of EMI, noting that "a cost-benefit assessment of EMI will need to go beyond these two objectives" (p. 292). Thus, to evaluate the trade-off between benefits and costs of EMI, future research needs to widen the conceptualisation of EMI and include a larger number of outcome measures.

Conclusion

In conclusion, while recognising the context-dependent nature of EMI implementation, the study adopted both direct and perception-based methods to achieve an in-depth, nuanced understanding of participants' viewpoints concerning EMI content learning, unveiling differences in terms of students' perceived benefits of studying Chemistry through EMI and JMI. This finding highlights that many of the benefits were distinctive to EMI education and exceeded language and content outcomes. To better assess its broader educational benefits, this study has aided in theory building of EMI outcome research. It has highlighted that benefits of EMI may not necessarily need to be investigated in an association with content gains, which are benchmarked against an L1 medium instruction counterpart. Rather the benefits of EMI should be investigated in its own right, and embrace a complexity of potential outcomes as well as risks. Future research is warranted to investigate whether EMI achieves its full potential to deliver subsidiary benefits beyond the oft-claimed language and content outcomes.

Disclosure statement

No potential conflict of interest was reported by the authors.

Ethics statement

Prior to the fieldwork, the research procedures and instruments of the project were reviewed and granted ethical clearance by Oxford University Central University Ethics Committee (CUREC). The data collection, protection, management and communication with participants had been designed in line with the guidelines devised by the British Education Research Association (BERA).

ORCID

Ikuya Aizawa http://orcid.org/0000-0003-4588-6194
Heath Rose http://orcid.org/0000-0002-6434-6663
Jim McKinley http://orcid.org/0000-0002-9949-8368
Gene Thompson http://orcid.org/0000-0002-2406-1364

References

Aizawa I, Rose H. 2019. An analysis of Japan's English as medium of instruction initiatives within higher education: the gap between meso-level policy and micro-level practice. High Educ. 77(6):1125–1142. https://doi.org/10.1007/s10734-018-0323-5

Bradford A. 2016. Toward a typology of implementation challenges facing English-medium instruction in higher education: Evidence from Japan. J Stud Int Educ. 20(4):339–356. https://doi.org/10.1177/1028315316647165

Civan A, Coşkun A. 2016. The effect of the medium of instruction language on the academic success of university students. Educ Sci-Theor Pract. 16:1981–2004. https://doi.org/10.12738/estp.2016.6.0052

Dafouz E, Camacho M, Urquia E. 2014. 'Surely they can't do as well': a comparison of business students' academic performance in English-medium and Spanish-as-first- language-medium programmes. Language and Education. 28(3):223–236. https://doi.org/10.1080/09500782.2013.808661

Dafouz E, Camacho-Miñano MM. 2016. Exploring the impact of English-medium instruction on university student academic achievement: The case of accounting. English for Specific Purposes. 44:57–67. https://doi.org/10.1016/j.esp.2016.06.001

EL-Daou B, Abdallah A. 2019. The impact of CLIL implementation on Lebanese students' attitudes and performance. GJFLT. 9(1):1–9. https://doi.org/10.18844/gjflt.v9i1.4068

Evans S, Morrison B. 2011. Meeting the challenges of English-medium higher education: The first-year experience in Hong Kong. English for Specific Purposes. 30(3):198–208. https://doi.org/10.1016/j.esp.2011.01.001

Flick U. 2009. An introduction to qualitative research. London: Sage Publications.

Galloway N, Kriukow J, Numajiri T. 2017. Internationalisation, higher education and the growing demand for English: an investigation into the English medium of instruction (EMI) movement in China and Japan. London, UK: British Council.

Hoare P, Kong S. 2008. Late immersion in Hong Kong: Still stressed or making progress? In: Fortune TW, Tedick DJ, editors. Pathways to multilingualism: Evolving perspectives on immersion education. Clevedon, England: Multilingual Matters; p. 242–263.

Hu J, Wu P. 2020. Understanding English language learning in tertiary English-medium instruction contexts in China. System. 93:102305. https://doi.org/10.1016/j.system.2020.102305

Huang H, Curle S. 2021. Higher education medium of instruction and career prospects: an exploration of current and graduated Chinese students' perceptions. J Educ Work. 34(3):331–343. https://doi.org/10.1080/13639080.2021.1922617

Ismailov M, Chiu TK, Dearden J, Yamamoto Y, Djalilova N. 2021. Challenges to internationalisation of university programmes: A systematic thematic synthesis of qualitative research on learner-centred English Medium Instruction (EMI) pedagogy. Sustainability. 13(22):12642. https://doi.org/10.3390/su132212642

Joe YJ, Lee HK. 2013. Does English-medium instruction benefit students in EFL contexts? A case study of medical students in Korea. Asia-Pacific Edu Res. 22(2):201–207. https://doi.org/10.1007/s40299-012-0003-7

Kim EG, Yoon J-R. 2018. Korean science and engineering students' perceptions of English-medium instruction and Korean-medium instruction. J Language, Identity & Educ. 17(3):182–197. https://doi.org/10.1080/15348458.2018.1433539

Macaro E. 2018. English Medium Instruction: Content and language in policy and practice. Oxford: Oxford University Press.

Macaro E, Curle S, Pun J, An J, Dearden J. 2018. A systematic review of English medium instruction in higher education. Lang Teach. 51(1):36–76. https://doi.org/10.1017/S0261444817000350.

Macaro E. 2020. Exploring the role of language in English medium instruction. Int J Bilingual Educ Bilingualism. 23(3):263–276. https://doi.org/10.1080/13670050.2019.1620678

McKinley J. 2020. Theorizing research methods in the 'golden age' of applied linguistics research. In: McKinley J, Rose H, editors. The Routledge handbook of research methods in applied linguistics (Routledge handbooks in applied linguistics). Abingdon, Oxon; New York, NY: Routledge.

MEXT. 2020. Heisei 30 nendo no daigaku ni okeru kyouiku naiyoutou no kaikaku joukyou ni tsuite [About the state of affairs regarding university reforms to education in 2018]. https://www.mext.go.jp/content/20201005-mxt_daigakuc03-000010276_1.pdf.

Murphy V, Arndt H, Baffoe-Djan JB, Chalmers H, Macaro E, Rose H, Vanderplank R, Woore R. 2020. Foreign language learning and its impact on wider academic outcomes: A rapid evidence assessment. London: Education Endowment Fund.

Park H. 2007. English medium instruction and content learning. English Language and Linguistics. 23(2):257–274.

Pun J, Fu X, Cheung KKC. 2023. Language challenges and coping strategies in English Medium Instruction (EMI) science classrooms: A critical review of literature. Studies in Sci Educ. 1–32. Advance online publication. https://doi.org/10.1080/03057267.2023.2188704

Rose H, Curle S, Aizawa I, Thompson G. 2020. What drives success in English medium taught courses? The interplay between language proficiency, academic skills, and motivation. Studies in Higher Educ. 45(11):2149–2161. https://doi.org/10.1080/03075079.2019.1590690

Rose H, McKinley J. 2018. Japan's English-medium instruction initiatives and the globalization of higher education. High Educ. 75(1):111–129. https://doi.org/10.1007/s10734-017-0125-1

Sahan KA, MikolajewskaMacaro M, Searle S, Zhou A, Veitch H. 2021. Global mapping of English as a medium of instruction in higher education: 2020 and Beyond. London: British Council.

Thompson G, Takezawa N, Rose H. 2022. Investigating self-beliefs and success for English medium instruction learners studying finance. J Educ Business. 97(4):220–227. https://doi.org/10.1080/08832323.2021.1924108

Tzoannopoulou M. 2014. CLIL in higher education: A study of students' and teachers' strategies and perceptions. In: Psaltou-Joycey A, Agathopoulou E, Mathiodakis M, editors. Cross-curricular approaches to language education. Newcastle upon Tyne: Cambridge Scholars Publishing.

Appendix 1. Content test

Content test

These questions cover concepts taught in CHM101. Complete the questions in no more than 25 minutes.

1. Give the oxidation state of each element in the following:

   (a) CO (b) $KMnO_4$

   (2)

2. Write the electronic configuration for Si.

   (1)

3. Indicate the core and valence electrons in the electronic configuration for Si.

   (1)

4. Draw the Lewis symbol for Si.

   (1)

5. Silicon reacts with fluorine. Write the chemical equation for the reaction.

   (1)

6. Draw the Lewis structure of the product in Q. 5 showing all valence electrons.

   (1)

7. Draw the molecule in Q. 6 in 3 dimensions (apply valence shell electron pair repulsion (VSEPR) theory).

   (1)

8. What is the name for the shape of the molecule in Q. 7?

   (1)

9. What is the hybridisation of the central atom in in the molecule in Q. 7?

   (1)

10. Draw the Lewis structure(s) of NO_2 and comment if more than one structure is possible.

    (3)

11. Write the Formal Charges next to each atom in the structure(s) you drew in Q. 10.

    (1)

12. Draw all valid open chain structures (isomers) for the molecular formula $C_3H_4Cl_2$.

    (2)

Appendix 2. Interview schedule

1. **Effects of medium of instruction (MOI) on content learning**

   - Does the MOI affect the level of your comprehension of the lecture and textbook?
   - Does the MOI affect your final grade and GPA?
   - Does the MOI affect the efficiency of your learning?
   - Do you think that the amount of time you spend on preparation and revision for the course would become shorter if you were learning in your first language?

2. **Effects of MOI on English language learning (only EMI students)**

   - Do you think that learning chemistry in English helps you improve your English level?
   - Which academic English skills do you improve through EMI? Reading, writing, listening, speaking?
   - Do you think that you improve your academic and technical vocabulary knowledge in English?
   - Do you think that EMI helps you improve your English skills as much as ESP courses?
   - You indicated your IELTS scores are_____. Is your English level enough to study the course in English?

3. **Challenges**

   - Do you face any challenges when learning chemistry in English/Japanese?
   - What about challenges regarding the following academic English/Japanese skills?

- writing - listening - speaking - reading - academic and technical vocabulary

   - Do you face any other general challenges when learning chemistry? Not language related-challenges?
   - When you face challenges, what form of support do you use? Reading textbooks, using dictionaries, and asking teachers?
   - Do you use any materials in your first language to help you learn chemistry better?

4. **Attitudes towards MOI**

   - Can you share your general attitudes towards learning in English? Are you in favour of EMI programs?
   - What do you think would be the benefits of EMI (or JMI) programs?
   - What do you think would be the disadvantages of EMI (or JMI) programs?
   - What was your motivation to take the EMI (or JMI) chemistry course when there was also the Japanese option?