

St Catherine's College

University of Oxford

Content Learning in Japanese and English Medium Instruction: Comparing Chemistry Learning Outcomes in Japanese Higher Education

81,347 words (including figures, notes, and tables excluding abstract, table of contents, references, and appendices)

Ikuya Aizawa

Supervised by

Dr Heath Rose

Dr Gene Thompson

Dr Jim McKinley

A thesis submitted in fulfilment of the requirement for the degree of DPhil in Education in the Social Science Division of the University of Oxford

Trinity 2021

Abstract

English Medium Instruction (EMI) is rapidly gaining momentum in line with internationalisation of higher education trends. However, this growth in provision is outpacing empirical research. Research has reported the lack of evidence on the impact of the medium of instruction on students' content learning outcomes. This study thus drew comparisons between EMI students (n=27) and Japanese medium instruction (JMI) students (n=26) in terms of their directly measured and self-perceived content knowledge acquisition. Data were collected at a research site university where the Chemistry courses were available in both Japanese and English as part of a parallel program offering the same curriculum in either of two languages of instruction. An analysis was made of students' content learning outcomes measured by a novel pre-post course content test created by the EMI Chemistry professor for the purpose of the research. These measurable test outcomes were triangulated with data from student interviews to identify differences in their perceived learning experience according to the language of instruction. These student interviews in addition to teacher interviews and lecture observations were used subsequently to discern the differences in the EMI and JMI course implementation as well as to elucidate academic challenges associated with medium of instruction. Subsequently, factors previously attested as crucial for content learning through EMI, such as proficiency, challenges, and prior content knowledge were explored to investigate the impact of these factors on students' success in content learning in the EMI program.

Methodologically, this study responds to calls for quasi-experimental research by adopting a longitudinal design of a course over three months, pairing in-depth quantitative analysis of the pre-post test scores with teacher and student interviews and lecture observations. The study found that while the quantitative findings unveiled no differences in content learning outcomes according to the medium of instruction, the qualitative evidence revealed that EMI students encountered a greater level of challenges than their JMI counterparts. The study also found that EMI students with lower English language proficiency faced considerably more onerous challenges than their more proficient counterparts when learning through EMI. Meanwhile, a range of factors also emerged as crucial facilitators for success in EMI, such as previous exposure to EMI education, motivation, revision and preparation, and learning strategies. This study will be one of the first attempts to investigate the effects of the medium of instruction on students' content learning outcomes in Japan.

Acknowledgement

There are many individuals whom I would like to thank for helping me produce this thesis. I would not have been able to conduct this study without any of their support.

Foremost, I would like to thank my supervisor, Dr Heath Rose, for his sincere support for the six years of my studies at Oxford. I could not have achieved what I have without his kindness and support.

My sincere appreciation also goes to my supervisors, Dr Gene Thompson, Dr Jim McKinley, and Dr Jessica Briggs Baffoe-Djan for offering their critical eye on my doctoral thesis.

I would also like to thank Professor Ernesto Macaro for examining my thesis at all the key milestones: the Transfer; Confirmation; and Viva.

I would also like to thank members of staff at both the Department of Education and St Catherine's College. The support network they created was vital to my student life at Oxford.

My studies at Oxford would also not have been possible without the generous support from Sakaguchi International Scholarship Foundation, Megumi Kokusai Foundation, Kanazawa Iwao Memorial Foundation, and British Association for Japanese Studies.

I would also like to thank all my friends in Oxford and Japan for sharing their knowledge to help me improve my writing and creativity. Without the thought-provoking discussion I had with them, I would not have been able to develop my research ideas.

In particular, my special thanks goes to Kedi for always encouraging me to write clearly, concisely and 'unctuously' using English idioms. Without her support, I would have opened a can of worms.

Finally, my upmost appreciation is reserved for my family for supporting my studies from Japan. Whenever I needed a hand to hold on, they took time to talk to me and give me advice.

You all have been always important figures in my life. The last six years of my study at Oxford would not have been possible without any of you.

Dedicated to all my friends and family.

Table of Contents

Abstract **2**
List of Tables and Figures **10**
List of Abbreviations **12**
Chapter 1- Introduction **13**
1.1 Background to EMI in Japan and research agenda 13
1.2 Research rationale 14
1.3 Research aims 14
1.4 Significance of the study 15
1.5 Layout of the thesis 15
Chapter 2 - Literature review **18**
2.1 Introduction 18
2.2 Terminological consideration 18
2.2.1. Defining EMI 18
2.2.2 Labelling EMI 20
2.2.3 Situating EMI on a language-content continuum 21
2.2.4 Summary 23
2.3 EMI and internationalisation of higher education 23
2.4 EMI and higher education policies in Japan 25
2.5 Reservations about EMI 28
2.6 Student challenges in EMI 29
2.7 Language support in EMI 34
2.8 English language proficiency in EMI 36
2.9 Language in science 38
2.10 Examining the outcomes of EMI 40
2.10.1 Success and outcomes of EMI 40
2.10.2 A review of academic outcomes studies of EMI 41
2.10.3 EMI outcome studies to date 43
2.10.3.1 Score gain research 43
2.10.3.2 Perception-based research 47
2.10.4 Strengths and limitations of previous outcomes studies 51
2.11 Research questions and conceptual framework 52
2.12 Chapter summary and statement of the problem 54
Chapter 3 - Methodology **55**
3.1 Introduction 55
3.2 Research questions and analytical framework 55

3.3 Methodological considerations 57
3.3.1 A mixed-methods design 57
3.3.2 A case study approach 58
3.4 Research design 59
3.5 Settings 62
3.6 Program 63
3.7 Faculty members 63
3.8 Participants 65
3.8.1 Sampling 65
3.8.2 Course level 65
3.8.3 Student level 66
3.8.4 Combining two sets of EMI cohorts 68
3.9 Data, instrumentation and procedure 69
3.9.1 Background questionnaire 70
3.9.2 Content test 70
3.9.2.1 Pretest/posttest 70
3.9.2.2 Home grown tests 71
3.9.2.3 Reliability and validity 72
3.9.2.4 Piloting the content test 74
3.9.3 Semi-structured Interview 75
3.9.3.1 Student interviews 75
3.9.3.2 Teacher interviews 76
3.9.4 The EMI Challenges Scale 77
3.9.5 Unstructured observation fieldnotes 78
3.9.6 Research procedure 80
3.10 Data analysis 81
3.11 Quantitative data analysis 81
3.11.1 Preparing data for statistical analysis 81
3.11.2 Normality, transformation and outliers 82
3.11.3 Inclusion and exclusion criteria 84
3.11.4 Descriptive and inferential statistics 84
3.12 Qualitative data analysis 86
3.12.1 Qualitative content analysis 86
3.12.2 Summary 90
3.13 Ethical considerations 90
3.14 Reflexivity of my positionality 91
3.15 Chapter summary 92

Chapter 4 - Results: Impact of the medium of instruction on chemistry content learning 94
4.1 Introduction 94
4.2 Research question 1.1 94
4.3 Within groups comparisons: whole sample 95
4.4 Between groups comparison: whole sample post-test scores 96
4.5 Between groups comparison: L2-English only sample 98
4.6 Within groups comparisons: L2-English only sample 99
4.7 Between groups comparison: L2-English only sample post-test scores 100
4.8 Medium of instruction and achievement 101
4.9 Research question 1.2 104
4.10 Content learning-related challenges 104
4.10.1 Comprehension of the course textbook 106
4.10.2 Comprehension of the course lectures 111
4.10.3 Final Grade 119
4.10.4 Adjustment to the new academic and cultural conventions of the Chemistry course 125
4.11 Success in content learning 129
4.11.1 Improved English proficiency 130
4.11.2 Improved motivation 131
4.11.3 Access to a wider range of learning resources 133
4.11.4 Long-term career advancement 134
4.11.5 Summary 135
4.12 Chapter Summary 137
Chapter 5 – Results: Context-related differences between the two mediums of instruction 138
5.1 Introduction 138
5.2 Research question 2 138
5.3 Language use 139
5.3.1 Language use in EMI 140
5.3.2 Language use in JMI 141
5.3.3 Summary - Language use 142
5.4 Peer learning 143
5.4.1 Peer learning in EMI 143
5.4.2 Peer learning in JMI 145
5.4.3 Summary – Peer learning 146
5.5 Support schemes 146
5.5.1 Support scheme: teaching assistant provision 146

5.5.2 Support scheme: The Comment Worksheet....149
5.5.3 Support scheme: The academic English language program....152
5.5.4 Summary – Support schemes....154
5.6 Educational culture....155
5.6.1 Interactive educational culture in EMI....156
5.6.2 Teacher-centred educational culture in JMI....157
5.6.3 Summary – Educational culture....159
5.7 Chapter summary....159
Chapter 6 - Results: Prior content knowledge, academic English challenges, proficiency and chemistry content learning in EMI....161
6.1 Introduction....161
6.2 Research question 3.1....161
6.3 Descriptive statistics: Proficiency, challenges, pre-test scores and post-test scores....162
6.3.1 Presenting descriptive statistics of key variables....162
6.3.2 Bivariate descriptive analysis: content knowledge achievement and proficiency....163
6.3.3 Bivariate descriptive analysis: EMI challenges and proficiency....165
6.3.4 Bivariate descriptive analysis: Content knowledge achievement and EMI challenges....167
6.3.5 Summary of descriptive statistics....169
6.4 Correlations: Proficiency, pre-test scores, challenges and post-test scores....170
6.4.1 Summary of correlations....172
6.5 Regression analysis: Proficiency, pre-test scores, challenges and post-test scores....172
6.5.1 To what extent does proficiency level predict post-test score in EMI?....173
6.5.2 To what extent does challenges score predict post-test score in EMI?....174
6.5.3 To what extent does pre-test score predict post-test score in EMI?....174
6.5.4 To what extent do proficiency and pre-test score predict post-test score in EMI?....175
6.5.5 Summary of regression analyses....176
6.6 Summary of research question 3.1....176
6.7 Research question 3.2....178
6.8 Content learning-related challenges....179
6.8.1 Comprehension of the course textbook....179
6.8.2 Comprehension of the course lectures....184
6.8.3 Final Grade....190
6.8.4 Adjustment to the new academic and cultural conventions of the Chemistry course....195

6.9 Summary of research question 3.2 200
6.10 Chapter summary 202
Chapter 7: Discussion of findings 205
7.1 Introduction 205
7.2 Research question 1 205
7.2.1 Content learning differences between the mediums of instruction 205
7.2.2 Broadening the definition of success in EMI 208
7.3 Research question 2 210
7.3.1 Key contextual factors in EMI 210
7.4 Research question 3 213
7.4.1 English proficiency and key learner factors in EMI 213
7.4.2 English language threshold in EMI 215
7.4.3 Challenges and academic success in EMI 216
7.4.4 Language-related challenges in EMI 218
7.4.5 Discipline specific challenges in EMI 220
7.4.6 Factors associated with challenges in EMI 221
7.5 Chapter conclusion 224
Chapter 8 – Conclusion 227
8.1 Introduction 227
8.2 Research question one 227
8.2.1 To what extent does the achievement of content learning of the EMI and JMI students differ? 228
8.2.2 How do the EMI and JMI students perceive the medium of instruction to affect their challenges and achievement of content learning? 229
8.3 Research question two 233
8.4 Research question three 234
8.4.1 To what extent factors, such as the EMI students' English language proficiency and perceived EMI challenges are associated with content knowledge achievement? 235
8.4.2 How do the EMI students perceive English language proficiency to affect their perceptions of challenges towards EMI study? 236
8.5 Limitations of the study 239
8.5.1 The sample size 239
8.5.2 Generalisability of the study 239
8.5.2.1 English proficiency 240
8.5.2.2 Course level 240
8.5.2.3 Self-enrollment 241
8.5.3 Duration of data collection 241

8.6 Recommendations for future research 241
8.7 Recommendations for curriculum planning in EMI 244
8.8 Conclusion 246
Appendices 248
Appendix 1: The number of classes taught in foreign languages or in English.....248
Appendix 2: Interview participants.....248
Appendix 3: Combining the two sets of the EMI groups 249
Appendix 4: Background Questionnaire Form (English translated Version).....251
Appendix 5: Challenges in your EMI classes (English version) 252
Appendix 6: EMI Student Interview Schedule (Japanese written after English) 254
Appendix 7: JMI Student Interview Schedule (Japanese written after English) 255
Appendix 8: EMI and JMI Teacher Interview Schedule (Japanese written after English) 256
Appendix 9: Content test: post-test questionnaire for pilot study purpose only.....258
Appendix 10: Marking Criteria.....259
Appendix 11: The syllabus of "Concepts in Chemistry" (EMI version) 260
Appendix 12: The syllabus of "Concepts in Chemistry" (JMI version) 262
Appendix 13: Start list of themes.....264
Appendix 14: Final list of themes 266
Appendix 15: Sample email requesting permission to conduct study 268
Appendix 16: Participant Information Leaflet 269
Appendix 17: Consent form.....271
Appendix 18: Z-score table for skewness and kurtosis.....272
Appendix 19: Histograms: EMI and JMI students' pre-test, post-test and gain scores 272
Appendix 20: ANCOVA assumption- Linearity for ANCOVA.....274
Appendix 21: Gain score analyses 275
Appendix 22: Content test.....278
Appendix 23: Tests of normality 279
Appendix 24: Matrices of pre-test score, challenges and post-test scores.....279
Appendix 25: Z-score table for skewness and kurtosis.....280
Appendix 26: Histograms of pre-test, post-test scores, challenges and proficiency..281
Appendix 27: Linearity assumption.....281
Appendix 28: Homoscedasticity assumption:.....282
Appendix 29: Normality of residuals (errors) assumption: 283
References 285

List of Tables and Figures

Table 1: Overview of EMI courses at undergraduate and postgraduate levels in Japan 27
Table 2: Research questions and analytical framework 56
Table 3: Mixed-methods research design adopted from Creswell, Plano Clark and Garrett (2008) 57
Table 4: Threats to internal and external validity 60
Table 5: Enrolment number 63
Table 6: Information of faculty members 64
Table 7: Purposive sampling of students (n=13 EMI; n=4 JMI) 67
Table 8: Sources of data in relation to time and quantity 69
Table 9: Criteria and specifics of the content tests 72
Table 10: Observation schedule 79
Table 11: Overview of quantitative data analysis 82
Table 12: Overview of qualitative data analysis 90
Table 13: Descriptive statistics – test scores for the EMI (n=27), JMI (n=26) and total sample (n=53) 95
Table 14: Descriptive statistics – test scores for the EMI (n=20), JMI (n=26) and total L2-English only sample (n=46) 99
Table 15: Top and bottom 15.8% of the post-test score distribution 103
Table 16: Interview participants 105
Table 17: Examples of analytic themes 139
Table 18: Examples of analytic themes 143
Table 19: Overview of analytic categories 147
Table 19: Overview of analytic categories 150
Table 20: Overview of analytic categories 152
Table 21: Examples of analytic themes 155
Table 22: Descriptive statistics of proficiency, challenges, pre-test, and post-test scores 163
Table 23: Descriptive statistics of post-test scores and IELTS level (n=27) 164
Table 24: Top and bottom 15.8% of post-test distribution by proficiency 165
Table 25: Descriptive statistics of challenges and IELTS level (n=27) 166
Table 26: Top and bottom 15.8% of challenges distribution by proficiency 167

Table 27: Descriptive statistics of post-test scores and challenges (n=27).....168
Table 28: Top and bottom 15.8% of post-test distribution by challenges 169
Table 29: A Spearman's rank-order correlation between proficiency, pre-test score, post-test score and challenges (n=27) 171
Table 30: Linear regression coefficients table: Proficiency and post-test scores 174
Table 31: Linear regression coefficients table: Challenges and post-test scores.....174
Table 32: Linear regression coefficients table: Pre-test scores and post-test scores .175
Table 33. Multiple regression coefficients table: Proficiency, pre-test scores and post-test scores 176
Table 34: Matrices of post-test scores and IELTS level (n=13 EMI students)***....179
Table 35: Matrices of challenges according to IELTS level.....184
Table 36: Matrices of challenges according to IELTS level.....189
Table 37: EMI participants' Chemistry course Final Grade (FG) by IELTS level....191
Table 38: EMI participants' final exam scores by English proficiency.....193
Table 39: EMI participants' Comment Sheet scores by English proficiency.....194
Table 40: Matrices of challenges according to IELTS level.....195
Table 41: Matrices of challenges according to IELTS level.....199
Table 42: Overview of EMI challenges 201
Table 43: Summary of recommendations for future research.....244

Figure 1: Continuum of EMI 22
Figure 2: Conceptual framework 53
Figure 3: The pretest–posttest non-equivalent control group design.....59
Figure 4: Gender, L1 and IELTS scores between the EMI and JMI students 66
Figure 5: Purposive sampling of student participants.....67
Figure 6: Combined EMI sample.....69
Figure 7: A list of themes discarded during first coding.....88
Figure 8: A final list of main themes for RQ1 and 3 89
Figure 9: Pre-test and post-test scores by EMI (n=27) and JMI (n=26) students.....96
Figure 10: Independence of the covariate and experimental (group) effect 97
Figure 11: Explained and unexplained variance.....98
Figure 12: Pre-test and post-test scores by EMI (n=20) students 100
Figure 13: Explained and unexplained variance.....101

Figure 14: Distribution of the participants' post-test scores (n=53) 102
Figure 15: Top and bottom 15.8% of the post-test score distribution..... 102
Figure 16: Textbook comprehension challenges by MOI (EMI/JMI) 111
Figure 17: Lecture comprehension challenges by MOI (EMI/JMI) 118
Figure 18: Problems and mediators of Final Grade by MOI (EMI/JMI)..... 124
Figure 19: Adjustment challenges by MOI (EMI/JMI) 129
Figure 20: Success in gaining content knowledge through EMI and JMI 135
Figure 21: Distribution of post-test scores by proficiency (n=27)..... 164
Figure 22: Distribution of challenges by proficiency (n=27) 166
Figure 23: Distribution of post-test scores by challenges (n=27) 168

List of Abbreviations

CLIL: Content and Language Integrated Learning

EMI: English Medium Instruction

ELA: English for Liberal Arts

EAP: English for Academic Purposes

ESP: English for Specific Purposes

ESAP: English for Specific Academic Purposes

EFL: English as a Foreign Language

HE: Higher Education

JMI: Japanese Medium Instruction

L1: First Language

L2: Second Language

MEXT: The Ministry of Education, Culture, Sports, Science and Technology

RQ(s): Research Question(s)

TGUP: Top Global University Project

Chapter 1- Introduction

1.1 Background to EMI in Japan and research agenda

In line with global trends, higher education in Japan has experienced a rapid surge in EMI programs in recent decades. In 2014, the Ministry of Education, Culture, Sports, Science and Technology (MEXT) introduced the 'Top Global University Project' or 'TGUP'− a large-scale, multimillion dollar-funded, higher education initiative, providing a total of 37 universities with substantial annual funding, with an aim to internationalise Japanese higher education (MEXT, 2014). The latest statistics from the government suggests that the participating universities are expected to increase the number of EMI courses from 19,533 to 55,928 (2.86 times) by the end of this ten-year funding period in 2023 (MEXT, 2018), indicating the large extent to which this EMI boom is expanding.

The driving forces of EMI programs (e.g. mobility of international students and faculty members) have now been well documented in research literature (e.g. Curle, Jablonkai, Mittelmeier, Sahan & Veitch, 2020). While these benefits may appear increasingly attractive to stakeholders, research evidence thus far is worrying, with research highlighting concerns surrounding university students' insufficient English language proficiency for EMI studies, and, crucially, around its potentially disastrous effects on their academic outcomes. Observing this growing scepticism towards the global explosion of EMI, Macaro (2018) urges 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 undergraduates substantive disciplinary knowledge, more research is needed to investigate the potential drawback of EMI on the quality of education received by university students, 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). Accordingly, this thesis investigates the impacts of EMI in relation to university students' learning of Chemistry in Japan.

1.2 Research rationale

In response to this research agenda, this research project seeks to investigate the impact of medium of instruction through a comparison of students' content learning outcomes between an EMI course and Japanese as Medium of Instruction (JMI) course as a comparison group. There is also a methodological component to this research agenda. Consistent with Macaro's (2018) observation of EMI research as "an evaluation of naturally occurring phenomenon", the current study investigates "an educational practice that is already happening and needs evaluation for its effectiveness" (p. 155). Rose (2019) also calls for more classroom-based research, viewing its value as "a valid reflection of real-world practices, which are not found in the sanitized experiments" (p. 899). Consequently, I adopt a quasi-experimental approach, ensuring its richness in ecological validity to capture the authenticity of an EMI course whereby students are self-assigned to two learning conditions, which Macaro explains as "voluntary enrolment" serving as a crucial factor in "outcome studies"—leading to numerous pre-existing differences (i.e. "messiness"). For example, the EMI sample in this study consists of a heterogenous group of domestic and international students with varying first languages, whereas its JMI counterpart is a homogeneous group, comprising exclusively of domestic students with Japanese as their first language. As EMI-designated research is prone to various uncontrolled factors such as these confronting researchers as methodological barriers, I follow Macaro's (2018) assertion by detailing "the enrolment process in either of the two settings" (EMI and JMI), and expatiating on the research populations utilising student and teacher interviews and lecture observations (p. 155). Thus, this research harnesses the complexity surrounding the constellation of factors driving success in EMI.

1.3 Research aims

Based on a single case of an undergraduate introductory Chemistry course in Tokyo, Japan, this thesis, focusing on students as primary stakeholders in EMI, aims to fill research gaps by investigating whether directly measured and self-perceived achievement of content learning differs between EMI and JMI programs. In this pursuit, the aim of this study is three-fold: 1) to investigate whether the learning of Chemistry through English can contribute to the acquisition of students' content knowledge in comparison to when their content learning takes place through

Japanese; 2) to identify what context-related factors students ascribe the mastery of their content knowledge acquisition in EMI; and 3) to examine to what extent key learner factors, such as EMI students' English language proficiency, prior topical knowledge and perceived academic challenges are associated with their achievement of content learning. Thus, the primary purpose of this thesis is to explore the potential impact of the medium of instruction on the academic outcomes of chemistry major university students with a range of L1 backgrounds. A secondary focus will be an examination of key factors that influence EMI students' content learning outcomes.

1.4 Significance of the study

This thesis will make significant methodological and theoretical contributions to the relatively underexplored area of educational outcomes in EMI research. Researchers argue that this rapidly expanding provision of EMI programs has outpaced an extensive body of empirical research, reporting the absence of monitoring systems or clear outcomes, which makes it difficult to assess the overall educational value of EMI, compared to its L1 medium instruction counterpart (McKinley, 2018; Rose & Galloway, 2019). Thus, this study aims to shed light on broader educational outcomes of EMI that exceed the content and language outcomes to examine to what extent students may be able to counterbalance the costs of the seemingly more onerous experience of learning through EMI with its wider educational benefits. The conclusions drawn from this research will therefore have relevant implications not only for Japanese higher education, but also for wider international EMI contexts where this boom has largely gone unmonitored without clear research evidence on cost-effectiveness of EMI. The findings will also be of value to stakeholders worldwide who are questioning the value of gaining content knowledge through EMI while also trying to promote the further expansion of EMI programs.

1.5 Layout of the thesis

The thesis is structured as follows. Chapter 2 presents an extensive critical review of literature concerning the global expansion of EMI both globally and in Japan. I begin with an overview of literature that seeks to identify the gaps between policy and practice of EMI in higher education, including various language-related challenges faced by EMI students. Subsequently, I critically analyse empirical research that has sought to shed light on discipline-specific challenges encountered by those learning

sciences through EMI to argue that students face both language and content learning-related challenges when learning content through EMI. I then highlight the necessity of language support provision for students to ensure sufficient linguistic ability for successful content learning. The chapter then continues to critique the shortcomings of existing literature on the content outcomes of EMI to inform the current research questions and conceptual framework. Chapter 2 ends with a summary of the research agenda that emerges from these limitations and justifications for the present study.

Chapter 3 of this thesis outlines the methodological approach of this study. This chapter opens with the methodological considerations for the study, including a discussion of a mixed-methods and case study approach. Then, the research questions and analytical framework will be introduced, followed by a description of the research procedure including setting, participants, pilot study, inclusion and exclusion criteria, and data collection procedures. This chapter also contains results from preliminary data analyses, concerning the normality of the test score variables and presence of outliers. Subsequently, the analytical procedure employed to examine the datasets are discussed. Finally, the chapter concludes with my engagement with ethical research.

Chapter 4 is the first empirical chapter of this study, in which I seek to answer the first research question: "Is there an impact of the medium of instruction on students' content learning?" I first introduce the quantitative research findings of this thesis regarding the directly measured content knowledge of the EMI and JMI students. Descriptive statistics will be used to explore the data and identify differences between the high and low performing students according to their medium of instruction. The focus of this section is on the mean post-test score differences between the two groups, while partialling out their pre-test scores as a covariate. This section also presents the results from an analysis of gain scores. The subsequent section presents the qualitative findings associated with the second part of the same research question regarding the self-perceived effects of medium of instruction on content knowledge between EMI and JMI students. Qualitative content analysis will be employed to explore their perceived challenges and successes based on the medium of instruction.

Chapter 5 seeks to answer the second research question of this study: "What context-related differences do the students ascribe to their achievement of content learning?" In light of the content learning differences discussed in the previous chapter, I explore

a range of contextual differences between the two mediums of instruction which students ascribed to their mastery of content learning. This chapter draws on student and teacher interviews triangulated with field notes from lecture observations to shed light on some of the plausible explanations for why the current study did not detect differences in the directly measured post-test scores between EMI and JMI students.

Chapter 6, focusing on EMI students, reporting on the quantitative and qualitative findings explicating the relationship between their pre-test scores, IELTS scores, perceptions of level of challenges in EMI, and post-test scores in attempting to answer the third and final research question of this study: "Are there relationships between the EMI students' English language proficiency, academic English challenges, prior content knowledge, and content knowledge achievement?" Firstly, the quantitative evidence will draw on descriptive and inferential statistics of the association between proficiency, challenges, pre-test and post-test scores. The second part pertains to the qualitative evidence involving 13 semi-structured student interviews to elicit a more nuanced understanding of how students with different levels of English language proficiency experience different levels of ease towards academic study in EMI.

In Chapter 7, I situate the findings of the current study within the fields of EMI and within the broader literature previously reviewed in Chapter 2. I evaluate each of the three research questions individually in turn, presenting my own interpretations in relation to the key findings in attempting to demonstrate how the current research contributes to new knowledge in the field. Chapter 7 concludes with an overview of the intended theoretical and methodological contributions of the current study.

Chapter 8 seeks to summarise and synthesise the results from the previous three findings chapters and outline my interpretations of their significance, implications and limitations. I will then evaluate the overview of the argument made in the thesis and critically discuss the extent to which the study provides satisfactory responses to the research questions. Subsequently, I present recommendations for possible areas of future research and pedagogical implications, before considering the methodological and theoretical reflections on the limitations of this study. Following that summary, I end the chapter with the contributions of this thesis to the field by contextualising its findings within the burgeoning area of EMI research.

Chapter 2 - Literature review

2.1 Introduction

In this chapter I will set out the current state of the literature surrounding EMI both globally and in Japan with respect to the learning of content in order to demonstrate how the current study contributes to new knowledge in the field. The chapter will first examine the global expansion of EMI in order to define some key constructs and position the current study in a wider context. I will then highlight gaps between policy and practice of EMI in higher education in particular by introducing various language-related challenges faced by EMI students, followed by those related specifically to learning science. This highlights the necessity of language support for students to ensure successful learning, which I explore next with an examination of L2 English proficiency in EMI. Finally, I will analyse current literature which discusses the content outcomes of EMI, followed by strengths and limitations of previous studies in order to inform the current research questions and conceptual framework.

2.2 Terminological consideration

2.2.1. Defining EMI

The use of English in higher education has given rise to the development of numerous variations of definitions that refer to the learning and teaching of content subjects through English. The definitions of EMI are often situated within the language-content spectrum, some of which lay a greater emphasis on language development or a dual focus on both content and language development. For example, Taguchi (2014) defines EMI as "English-medium education refers to curricula using English as a medium of instruction for basic and advanced courses to improve students' academic English proficiency" (p. 89). In contrast, others focus more on the development of content knowledge, such as, "settings where English is the language used for instructional purposes when teaching content subjects although not itself the subject being taught, and also a second or additional language for most participants in the setting…" (Pecorari & Malmström, 2018, p. 499).

No universally accepted definitions of EMI exist because pedagogical goals of EMI are conceived differently across diverse educational contexts (e.g. Dearden, 2015;

Vavelyuk, 2015; Fenton-Smith, Humphreys & Walkinshaw, 2017; Crandall & Christison, 2016; West, Guven & Parry, 2015). For example, Costa (2012) observes that Italian EMI lectures are "mainly meaning focused but with incursions into language focus" (p. 42), strongly suggesting that language outcomes are not explicitly on the agenda alongside content learning. Conversely, in discussing the situation in Spanish higher education, Dafouz and Camacho-Miñano (2016) describe that Spain's Ministry of Education increased the number of EMI programs at tertiary levels to improve university students' English proficiency for better graduate employability since 2014, indicating that there is a dual focus on both content and language development behind its policy introduction. Thus, individual contexts exhibit their own versions of definitions, motivations and aims for introducing EMI (Henriksen, Holmen & Kling, 2019), and the actual practice of implementing EMI in many institutions of higher education is individually unique due to "contextual, geographical, historical and political reasons" inherent in each distinctive case (Dearden, 2015, p. 14).

In this thesis, I adopt the definition proposed by Macaro (2018), which states that "the use of the English language to teach academic subjects (other than English itself) in countries or jurisdictions where the first language (L1) of the majority of the population is not English" (p. 19). This definition is highly relevant to the current context of Japanese higher education, where success within EMI is primarily associated with the development of content knowledge. This instructional goal is mirrored in the definition introduced by the Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT), who has defined EMI as "courses conducted entirely in English, excluding those whose primary purpose is language education" (MEXT, 2018). This definition states that students learn academic subjects entirely through English where the learning of content is the primary educational objective. Many of the EMI researchers in Japan (e.g. Wilkinson, 2015; Brown & Bradford, 2018) also concur with this definition that professors who teach on EMI programs in Japanese universities typically regard themselves as content subject experts rather than language experts, merely using the English language to convey content knowledge.

2.2.2 Labelling EMI

Having outlined the working definition of EMI in this thesis, I now turn to discuss different labels devised to refer to this global phenomenon to conceptually differentiate the term of EMI from other similar terms. Many members of the research community are acutely aware that there is now a bewildering array of similar terms which are used inconsistently to refer to this phenomenon. For example, Dearden and Macaro (2016) note that "the issue of terminology (e.g. EMI, CLIL) and what it represents is a complex one" (p. 456). In their subsequent systematic review of EMI, Macaro et al. (2018) assert that "the labels given to the phenomenon of EMI and their definition are inconsistent and problematic" (p. 14). More recently, Richards and Pun (2021) developed a typology of various forms of EMI programs based on ten criteria (e.g. purposes, assessment, curriculum and materials), arguing that "[EMI] is labelled in a variety of ways… the expansion of EMI worldwide has resulted in many different forms of EMI, as well as some confusion as to how they differ" (p. 1), highlighting the complexity of the term and diversity of approaches to introducing EMI.

Different literature that refers to instruction conducted through English draw on different terms (e.g. Macaro et al., 2018; Schmidt-Unterberger, 2018; Kling, 2019): such as English for Academic Purposes (EAP) (e.g. Hyland, 2016); English for Specific Purposes (ESP) (e.g. Dudley-Evans & St. John, 1998); Content Language Integrated Learning (CLIL) (e.g. Coyle, Hood & Marsh, 2010); Integrating Content and Language in Higher Education (ICLHE) (Ute & Dafouz, 2012); Content-based Instruction (CBI) (e.g. Snow & Brinton, 1988); and Content-Based Language Teaching (CBLT) (e.g. Lightbown, 2014). These terms are often synonymous and thus confusing for inexpert readers (Ament & Pérez-Vidal, 2015). It is even more confusing when some of the authors use more than one of the terms to refer to the same program in their research, with for example, Dafouz et al. (2014) adopting both EMI and CLIL interchangeably to refer to their case study English medium course in Spain (p. 224). Bradford and Brown (2017) observe that even in the research community, some researchers do not see the differences in these terms, using them synonymously. Consistent with this critique, MacGregor (2016) confirmed from her case study, which explored the perceptions of language-teaching experts in Japan, that her participants all perceived CBI and CLIL to be synonymous terms and these types of instruction could encompass varying proportions of content and language learning.

While many of these educational approaches are commonly practised in non-Anglophone educational settings due to their purported dual instructional focus of teaching both content and language knowledge (Curry & Pérez-Paredes, 2021; Doiz et al., 2011; Fenton-Smith et al., 2017), Pecorari and Malmstrom (2018) observe that some researchers interpret the geographical boundaries of EMI more widely to include Anglophone contexts in their discussion of EMI. Baker and Hüttner (2017) also argue that the strong presence of international populations in some of these Anglophone contexts make the situation no less different from that of the non-Anglophone contexts (Fenton-Smith et al., 2017; Jenkins & Mauranen, 2019). As a result, this wider definition of EMI has contributed to the emergence of further terms. For example, Dafouz and Smit (2016) devised an alternative term of English-Medium Education in Multilingual University Settings (EMEMUS) to shift the conceptualisation of EMI to a more localised and institutional context of education, capturing multilingualism in institutions where English medium education is situated.

These nuanced labels could encapsulate some educational practices more accurately than the traditional term of EMI. For example, EMEMU could be opted to serve certain Anglophone settings where a predominant population is international students and faculty members introducing a variety of first languages, in which the term of EMI may not be the most appropriate label due to the geographical restriction stated in its definition. Whatever definition is adopted, however, there is still no consensus on what labels should be used among researchers for their respective educational practices where English is used as a medium of instruction. I will now turn to discuss how I situate EMI in relation to other similar but inherently different alternatives to clarify what is and is not EMI and why the term of EMI was selected in this thesis.

2.2.3 Situating EMI on a language-content continuum

In an attempt to resolve this complexity surrounding the terminological ambiguities (e.g. equating EMI with other similar terms), I will now locate EMI within a broader educational discussion that concerns the nexus between language and content. Researchers (e.g. Schmidt-Unterberger, 2018; Brown & Bradford, 2017) advocate the necessity for establishing a clearer conceptual separation among these similar terms due to the apparent differences in what each of these approaches are inherently trying to accomplish—the instructional goals of EMI. For instance, Thompson and

McKinley (2018) offered a language-content spectrum to locate EMI on a continuum in relation to other similar terms to clarify its instructional goals (Figure 1).

Figure 1: Continuum of EMI



(adapted from: Thompson & McKinley, 2018)

In Figure 1, on the right end of the continuum is a cluster of approaches that are primarily language driven, such as English as a Foreign Language (EFL), English for Academic Purposes (EAP), and English for Specific Purposes (ESP). Located between the two ends of this continuum, approaches such as immersion, CLIL, ICLHE and CBI all share the similar learning objective of furthering both content and language knowledge. EMI, on the other hand, placed at the far left of this spectrum, does not project this dual objective, setting itself apart from the rest of the terms. In this thesis, the term EMI which is primarily content driven is thus most relevant to the context of Japanese higher education where priority is given to content development. Drawing clear boundaries between the terms to make their differences clearer, such as that depicted in Figure 1, is necessary to operationalise the key terms in this research.

However, these theoretical divisions are inevitably artificial when being put in practice. Different communities of practice attach their own version of interpretations to the balance between language and content when integrating EMI in their own teaching and learning. Although language learning is by definition not explicitly articulated, EMI does not seem to necessarily exclude language development from its overarching educational aim. Scholars (e.g. Hu & Wu, 2020; Aguilar & Muñoz, 2014; Wei, Feng & Ma, 2017) equate the development of language knowledge with a dual aim of EMI courses, arguing that EMI provides an ideal and effective learning environment for students to enhance their linguistic proficiency. Rose et al. (2020) found that in many East Asian contexts, where the expansion of EMI is often driven by top-down policy initiatives, "EMI as a policy" is centred on students' acquisition of content knowledge, but "the actual practice of implementing EMI" often revolves around a dual focus on "students' acquisition of both content and language knowledge" (p. 8). Similarly, Hu and Wu (2020) claim that students in China who elected to study through EMI expressed their perceived linguistic improvement as one

of the main factors for enrolling in the respective EMI courses over Chinese medium courses. Unterberger and Wilhelmer (2011) also argue that EMI provides implicit language learning goals whilst explicitly articulating content learning aims:

> [The absent of explicitly formulated language learning goals] does not necessarily mean that universities do not strive to improve their students' English, but that they simply perceive it as an implicit aim because it is assumed that English skills will be honed incidentally and, at this educational stage, are often considered to be sufficient. (p. 95)

Therefore, in this study, the primary focus of EMI predominantly remains the transmission of content knowledge while language learning can also be still expected to take place incidentally alongside content development (Pecorari, 2020; Wilkinson, 2015; Bradford, 2016; Brown & Bradford, 2018).

2.2.4 Summary

To summarise, EMI is defined in a variety of ways in many parts of the world, thus the complexities have arisen in scholarly attempts to devise a consensual definition. In recognition of this terminological and definitional issue, this section has offered the working definition of EMI to clarify the scope of this research, filter what literature to be reviewed in the chapter, and determine how the results will be analysed. In particular, the working definition has informed the methodological justifications of this research, such as research contexts (i.e. non-Anglophone countries), content subjects (i.e. content other than the English language itself), instructional goals (i.e. primarily content driven), and classroom language use (i.e. exclusive use of English).

2.3 EMI and internationalisation of higher education

In a globalised era in which English has become the de facto lingua franca for research, teaching and learning (Rose & Galloway, 2019; Jenkins, 2019), higher education institutions across the globe have seen a shift from English being studied as a second language (L2) along with many other school subjects to English growing to an important educational language used for learning and teaching non-language related academic content subjects (Dearden, 2015). This EMI phenomenon is growing so exponentially all over the world that it has been described by some as being a

"galloping phenomenon" (Chapple, 2015, p. 1), "pandemic" (Phillipson, 2009, p. 195), or somewhat of an "unstoppable train" (Macaro, 2015, p. 7). .

One of the main driving forces behind this rapidly growing EMI movement is the increasing forces of internationalisation, and many universities across the globe introduce EMI as a vehicle to achieve their internationalisation strategic goals. In exploring the current state of EMI policy implementation in Chinese higher education, Rose, McKinley, Xu and Zhou (2020) found that "the main driving force [behind the introduction of EMI] was to meet their internationalisation objectives" across many Chinese universities that are striving for internationalisation (p. 5). Observing this worldwide trend in higher education, Galloway and McKinley (2019) posit that internationalisation in non-Anglophone countries is often being realised by Englishisation of curricula within many higher education institutions. Similarly, Radjai and Hammond (2021) also suggest that EMI leads to the internationalisation of student population, attracting more overseas students in Japan. Internationalisation and the 'Englishization' are thus inextricably intertwined in higher education.
In recent years, an impressive body of research has been conducted to document motivations for higher education institutions around the world to adopt EMI. The British Council has commissioned a number of research reports on EMI provision in many geographical regions of the world to better understand the current situation in relation to the introduction of EMI. In particular, Curle, Jablonkai, Mittelmeier, Sahan and Veitch (2020) recently compiled a state-of-the-art global review of EMI, outlining some of the key factors which contribute to the expansion of EMI in higher education.

Much of the research reviewed in their report as well as many other policy-based studies (e.g. Ali, 2013; Aizawa & Rose, 2019; Rose et al., 2020) appear to explore the driving forces and perceived benefits of EMI based on three policy dissemination levels. First, on the national level, the expansion of EMI is associated with increasing international student revenue, research excellence and nation-building (macro level). Second, on the institutional level, EMI plays an important role in enhancing English as the language of higher education, institutional reputations, research quality and graduate employability (meso level). Third, on the individual level, EMI provides an opportunity for domestic students to improve English language competencies, gain access to wider ranges of teaching materials, communicate with overseas students and faculty members, and advance career opportunities (micro level). While these

perceived benefits of EMI identified in research are highly context sensitive and influenced by multiple factors for nations, institutions and individuals (Hultgren, Jensen & Dimova, 2015), many universities across the globe all appear to introduce EMI to raise their international profiles and meet their internationalisation objectives.

Japan is no exception, and the introduction of EMI is seen as a gateway to successful internationalisation. Many of the recent internationalisation education policies of Japan's Ministry of Education emphasised the role of EMI provision in driving the internationalisation of the domestic labour market and the global presence of Japanese companies (Rose & McKinley, 2018). In essence, the development of EMI programs is linked to the nation's global competitiveness. This predominant emphasis on internationalisation is consistent with the rising prominence of EMI in Japanese higher education.

Overall, it is now well established that internationalisation is the main driver and motivation for the introduction of EMI. Nevertheless, there has been limited research that explicitly links EMI provisions and internationalisation. For example, there are criticisms about the degree to which EMI results in improved university rankings (Hultgren, 2014) or increased international student enrolment (Curle et al., 2020). Little evidence also exists for improved student English proficiency as a result of long-term exposure to the target language (Rogier, 2012). As Galloway and McKinley (2019) alluded to previously that Englishisation can be a "threat or opportunity", more research is needed to investigate whether the opportunities outweigh the threats when EMI is introduced, which Macaro (2018) refers to as "cost-benefit" analysis of EMI. To fill this gap in literature, this thesis examines the potential costs and benefits for students from learning content through EMI, compared to its L1 medium instruction.

2.4 EMI and higher education policies in Japan

In 2014, Japan's Ministry of Education (MEXT) announced a higher education initiative, known as the "Top Global University Project" (TGUP), a 10-year, multimillion dollar investment initiative with the aim "to enhance the international compatibility and competitiveness of higher education in Japan", and to offer "prioritized support for the world-class and innovative universities that lead the internationalization of Japanese universities" (MEXT, 2014). As the most recent

higher education reform, MEXT received applications from 109 candidates and recruited a total of 37 universities as the recipients of this substantial annual funding.

Compared to previous internationalisation policies (e.g. Go Global Japan, Global 30), which had placed their primary emphasis on the increased number of EMI programs and overseas students, the TGUP has expanded on the aims of these previous initiatives in its further focus on the role of research in Japanese higher education, such as research excellence, international collaboration, increased presence of overseas faculty members, and improved position of Japanese institutions in worldwide university league tables. Thus, although the TGUP does not explicitly state a focus on the number of EMI programs per se, it has still maintained similar objectives as these previous initiatives leading to the continued growth of EMI (Rose & McKinley, 2018).

In recognition of the influence of the TGUP on the expansion of EMI, the latest figures published by MEXT (2018) showed that the TGUP funded universities increased the number of "courses in a foreign language" from 19,533 to 32,846 (1.68 times) and the proportion of "degree courses which undergraduate students can complete in a foreign language" grew from 652 to 873 (1.33 times) during the period between 2016 and 2018. The participating universities are expected to further increase the number of "courses in a foreign language" from 19,533 to 55,928 (2.86 times) and the proportion of "degree courses which undergraduate students can complete in a foreign language" from 652 to 1,226 (1.73 times) by the end of this 10-year funding period in 2023, signalling future EMI growth.

Meanwhile, MEXT has also published figures regarding the overall number of Japanese universities offering EMI programs at undergraduate and postgraduate levels annually since 2015, with the latest figures reported in 2018 (see Table 1 below). As the most recent official record including both the TGUP and non TGUP funded universities, the 2018 data involved 761 universities out of 782 universities invited to participate in the survey. The data revealed that 305 universities (i.e. 41%) offered EMI courses at an undergraduate level, which represented a 3% of reduction compared to those (324 universities, 44%) reported in 2017. Similarly, 227 universities (i.e. 36%) offered EMI courses at a postgraduate level in 2018, which represented a 2% of reduction compared to those (234 universities, 38%) reported in

2017 (MEXT, 2018). Therefore, while the TGUP universities have increased the number of EMI courses since the introduction of the initiative in 2014, the overall number of Japanese universities offering EMI courses (when including non TGUP universities) remained largely unchanged between 2015 and 2018.

Table 1: Overview of EMI courses at undergraduate and postgraduate levels in Japan

	2015	2016	2017	2018
Undergraduate	305 (41%)	309 (42%)	324 (44%)	305 (41%)
Postgraduate	229 (37%)	222 (36%)	234 (38%)	227 (36%)

Adopted from MEXT (2018)

While these official statistics have been oft-cited in numerous studies as evidence to support the exponential growth of EMI in Japan, it is worth highlighting several methodological issues inherent in how MEXT gathered their data. Firstly, MEXT offered no information about the criteria of what was being considered "universities offering English-taught classes" on their documents. It remained unclear as to whether the universities would need to provide a certain number of English-taught courses to be classified as EMI providers. We could argue that some universities could be entirely EMI universities, offering only EMI courses (e.g. Akita International University) whereas others could offer a small number of EMI course and still be considered EMI providers. Secondly, MEXT did not recognise a great deal of variability in how EMI could be defined within and between different institutions. This meant that each university could have their own definitions of EMI when responding to the MEXT's survey. Investigating this complexity surrounding how EMI was defined among different stakeholders, Aizawa and Rose (2019) found that code switching between Japanese and English was rife in many of the courses that were even labelled as fully English-taught at their case study university which was part of the TGUP initiative. Aizawa and Rose (2019) revealed that their case study university still regarded such mixed-language courses as EMI courses when in reality Japanese was used to varying degrees. Therefore, it should be noted that the way in which MEXT measured the number of EMI courses and EMI providers remained largely inconsistent. In other words, the definition of EMI was left open for individual respondents (often understood to be administration staff from university divisions, such as Centre for Teaching and Learning) at each university who determined what

should be considered EMI. Thus, the absence of such reliable national data has posed a great challenge to EMI researchers in painting an accurate picture of the true extent to which EMI has prevailed in Japanese higher education.

Despite this methodological challenge, Japanese higher education institutions, in particular the TGUP funded universities, have witnessed the burgeoning spread of EMI over the last decade, mostly spurred by the recent government initiatives (e.g. Go Global Japan, Global 30, and the TGUP). This rapid expansion of EMI may result in unexpected outcomes of EMI and exacerbate various implementation-related challenges of EMI, which I draw attention to in the next section.

2.5 Reservations about EMI

The EMI phenomenon is expanding so rapidly that keeping pace with this expeditious movement has been proven to require much effort for EMI provisional universities. For example, Macaro and Han (2019) and Galloway and Ruegg (2020) argue that introducing EMI requires structural reforms in university operation, such as building professional development courses for lecturers and supplementary English language courses for domestic students to help them adapt to the new English-speaking academic environment. Galloway and Rose (2021) observe that "Englishization" also leads to changing the role of ELT practitioners who come under increased pressure to offer more discipline-specific language support instead of the general language support. In other words, the perceived benefits of EMI in many universities are not guaranteed, and there is growing scepticism that this EMI movement might even do more harm than good when it is introduced without much careful planning and support structure (e.g. Hu, Li & Lei, 2014; Evans, 2002).

In Japan, since the introduction of the TGUP in 2014, a number of scholarly attempts have been made to take advantage of this unique "Englishized" higher education context and investigate the impact of this policy implementation on many aspects of challenges facing key stakeholders, in addition to linguistic behaviour of L2 English students and teachers, such as L1/L2 code choice and attitude, turn taking patterns, and classroom interaction (e.g. Yoshida, 2015; Moore, 2017; Kim, 2020). Brief reports about the TGUP and its subsequent impact on Japanese higher education, in particular, the improved number of international students, have also been published by different sectors (e.g. researchers, independent organisations and governments)

(e.g. Sandhu, 2015; Collins, Lee, Hawkins & Newbauer, 2016). Nevertheless, little research has been carried out to investigate how this government initiative has impacted students' educational experiences, in particular the effectiveness of EMI and its subsequent impact on students' content learning outcomes (Macaro et al., 2018). The current study, therefore, will be one of the first attempts to fill this gap and explore the outcomes of EMI in Japan.

Galloway et al. (2017) identify the dearth of such empirical evidence, arguing that EMI provision is outpacing an extensive body of empirical research. Calling for more research evidence and monitoring systems to measure the outcomes and effectiveness of EMI programs, Galloway and McKinley (2020) suggest that "internationalization education policy decisions need to be matched with research into the complexities of such rapid developments. Such research is needed to inform effective policy implementation" (p. 1). In the following section, I will now turn to discuss students' language-related and content learning-related challenges which are one of the key issues underpinning much of this growing concern about the effectiveness of EMI.

2.6 Student challenges in EMI

Attitudinal research has shown that a variety of stakeholders in the context of higher education, whether it be students, faculty members or administrative staff, express challenges when EMI programs are introduced in their institutions (e.g. Lin, 2019; Song, 2019). However, one of the oft-reported issues hindering the introduction of EMI is students' language-related challenges. A growing body of EMI research has shown that students experience language-related challenges when learning through English, which in turn causes an adverse effect on their content acquisition and mastery (e.g. Bradford, 2012; Cho, 2012; Airey & Linder, 2006; Pun & Jin, 2021).

To investigate the linguistic challenges faced by EMI students, a large proportion of research literature has explored four language skill areas, such as difficulty in comprehending lectures and professors' accents (listening) (Ali, 2020); writing essays in an appropriate academic style (writing) (Pessoa, Miller & Kaufer, 2014); delivering oral presentations and participating in tutorial and seminar discussions (speaking) (Dippold, Heron & Gravett, 2021); and understanding textbooks and reference materials with an abundance of unfamiliar vocabulary (reading) (Ekoç, 2020). There is now a rapidly growing body of empirical research into challenges in many parts of

the world. Nevertheless, researchers in many cases argue that the evidence to date is insufficient in part due to research methodology weaknesses (Macaro et al., 2018). In the following paragraphs, I turn to discuss a number of existing EMI studies which have documented challenges encountered by L2 English students in adjusting to the demands of EMI programs. The review focuses on recent literature, aiming to point out the limitations of the current research and priorities for future research.

Firstly, much of the research published to date has typically relied on cross-sectional designs, such as the use of one-off questionnaires and interviews, often failing to capture the diachronic nature of students' experience over the course of undergraduate careers (Evans & Morrison, 2011). The reliance of cross-sectional data means that researchers can only synchronically examine challenges specific to a certain data collection point in time. As a growing number of studies show that students become more accustomed to EMI study and overcome challenges (Yuksel, Soruç, Altay & Curle, 2021; Aizawa & Rose, 2020), the timing and frequency of data collection is crucial to the types of challenges researchers can observe in their investigation.

In a rare longitudinal study, Evans and Morrison (2011) followed the trajectory of first year undergraduate students entering an EMI university in Hong Kong. Based on questionnaires (n=3000) and semi-structured interviews (n=28) administered at two data collection points over the course of one academic year, the study found that certain challenges were short-term teething problems (e.g. planning and delivering presentations) whereas there were also long-term challenges (e.g. comprehending lectures) that persisted throughout the academic year. They observed that "student needs and wants change over the course of their undergraduate careers" (p. 203). In another longitudinal study, Pessoa, Miller and Kaufer (2014) tracked a cohort of L1 Arabic students (n=86) at an EMI university in Qatar over the course of four years to examine their challenges in the area of academic reading and writing skills. The study drew on interviews (n=23) conducted once per semester and analysis of a total of seven writing pieces produced over the four-year period (n=86). The study found that students faced academic writing and reading comprehension problems due to limited background knowledge, vocabulary size and reading stamina. However, interviews showed that once learning strategies were taught, they started to perceive more ease towards academic reading. The corpus-based text analysis also recorded academic writing development over the course of their study, such as increase in the use of

academic register, elaboration, and reasoning. Thus, these longitudinal studies offered detailed and illuminating evidence into the changing nature of students' language challenges in EMI. Galloway, Numajiri and Rees (2020) also observe the relative neglect of longitudinal research in the field of EMI, suggesting that "more research is needed that includes... longitudinal studies on the effectiveness of teaching university subjects through the medium of English" (p. 412).

Secondly, a large proportion of EMI studies have focused on students' general language-related challenges, rather than discipline specific challenges. It is notable that much research in the field of EMI recruits a mixture of students specialising in a range of different academic disciplines, rather than one academic discipline (e.g. Yıldız, Soruç & Griffiths, 2017; Siegel, 2020; Querol-Julián & Camiciottoli, 2019). For example, Doiz, Costa, Lasagabaster and Mariotti (2020) conducted a large-scale study to investigate university students' linguistic issues and their needs for language support. The study administered a questionnaire to EMI university students in Italy (n=145) and Spain (n=145) from a range of academic disciplines, such as: business administration; engineering; economics; history; law; marketing; and public administration. Their findings suggested that students faced challenges in the areas of academic speaking, writing, reading and pronunciation. In broad terms, their findings pointed to general academic English challenges. Similar methodological decisions to recruit students from multiple fields of study were also made in the studies of Hu and Duan (2019) in China, Kırkgöz (2016) in Turkey, and Aizawa and Rose (2020) in Japan. Even though some studies have sampled participants from one academic discipline, they still predominantly focus on students' general language-related challenges. For instance, Ali (2020) examined lecture comprehension difficulties experienced by students in an engineering undergraduate program in Oman. While the interviewees (n=12) were all recruited from the same engineering course, the types of challenges reported in their findings were unsurprisingly all language-related challenges, such as difficulties with lecturers' accents, speed of lecture delivery, note taking and unfamiliar vocabulary.

This excessive focus on language-related challenges in EMI research is explained by Macaro's (2020) observation of the field that the phenomenon of EMI to date has been studied predominantly by applied linguists rather than content specialists whose

research interests tend to lie in the development of students' academic English skills. Another possible explanation for this relative neglect of discipline specific challenges related to the difficulty in separating content from language challenges. Much research appears to use the label of "language-related challenges" as a one-size-fits-all term to broadly encapsulate all types of challenges expressed by students who study through EMI. Macaro (2020) offers an example of mathematical words, such as "homeomorphism", arguing that many EMI students could have difficulty with this word due to their limited disciplinary knowledge and/or their limited linguistic knowledge to understand each of the morphological elements, namely, "topology"; "isomorphism"; and "bi-continuous" (p. 266). Learning through L2 English is one of the unique aspects of EMI, and thus linguistic challenges are rife in many contexts. While language should be at the forefront of EMI research, more research is needed to explore discipline specific challenges which are also part and parcel of EMI study. The current study thus aims to explore students' language-related challenges as well as those specific to the learning of Chemistry in EMI.

Finally, a bulk of studies have explored students' challenges often via one single method, instead of triangulation of multiple data collection methods. In particular, many studies have not necessarily included classroom observations to account for underlaying factors affecting student learning challenges (e.g. Floris, 2014; Belhiah & Elhami, 2015; He & Chiang, 2016; Phuong & Nguyen, 2019; Pun & Jin, 2021). For instance, Lin and Morrison (2010) examined the relationship between school-university transitional challenges faced by first year students and their vocabulary size, used as an indicator of linguistic proficiency, in an EMI university in Hong Kong. The study involved vocabulary size tests to measure students' academic vocabulary size (n=762). The finding revealed that only 1.4% of students from a Chinese medium instruction (CMI) high school background reached the "satisfactory level" of productive academic vocabulary size, indicating almost all CMI students (98.6%) had an inadequate productive vocabulary size (p. 259). The authors concluded that those without this satisfactory level would have experienced difficulties in understanding lectures and coping with prerequisite reading assignments. Although the study demonstrated students' vocabulary size robustly in a large-scale study using a validated vocabulary size test (Coxhead, 2000), it should have adopted subsequent interviews or classroom observations to investigate how the

insufficient vocabulary knowledge adversely affected their academic performance. Similarly, based on one single method, Yıldız, Soruç and Griffiths (2017) employed an open-ended questionnaire to investigate student challenges at an EMI university in Turkey (n= 83). The in-depth qualitative text analysis of over 300 statements produced by students revealed a range of challenges according to six categories (e.g. difficulties related to technical terms or vocabulary). Similarly, Hua (2019) also employed an open-ended questionnaire to investigate EMI learners' experience and perspectives on their EMI learning in Taiwan (n=59). She subsequently introduced focus group, aiming to "fill knowledge gaps about their underlying causes of those reoccurred themes in the open-ended questionnaire" (p. 328). However, in both these studies, observation was absent which could have elicited more detailed and illuminating contextual information required to achieve a deeper understand of student experience and the possible source of their challenges (e.g. class size, teaching style, classroom language use, and interaction between teachers and students).

Some research has incorporated a mixture of data collection methods, including observation (e.g. Hernandez-Nanclares & Jimenez-Munoz, 2017; Zhang, 2018). For example, Jiang, Zhang and May (2019) adopted nine classroom observations, three post-observation interviews and a questionnaire survey in their challenges research in China. The study analysed post-observation interviews to successfully elicit contextual information (e.g. whether continued language support was offered alongside regular EMI courses) which greatly reduced students' challenges. Similarly, Trang and Trinant (2017) implemented a questionnaire, semi-structured interviews and classroom observations to unpack what actually transpired in the EMI lectures. The analysis of classroom interaction showed that factors, such as the absence of student-teacher interaction affected student challenges. As such, more research that includes triangulation of data collection methods is needed to offer more comprehensive contextual information reflecting student learning experience in EMI.

To summarise, this section has highlighted gaps in the previous research literature. Three research concerns have emerged. Firstly, cross-sectional data rather than longitudinal data can only capture challenges specific to a certain data collection point in time. Secondly, many studies have focused on language-related challenges instead of content specific challenges. Thirdly, adopting one single method falls short of capturing the complexity of student challenges in EMI. Thus, the current study will

address these concerns by offering as much contextual information as possible. Students' challenges in EMI are highly context dependent, such as institutions, academic disciplines, course level, student/teacher English proficiency, and student previous EMI experience. Research approaches taken should therefore also be highly sensitive not only to the meso (institute level) but also to the micro contexts (classroom level) (Curle et al., 2020). In other words, applied linguists should consult with content specialists to adopt research methodology that is most appropriate to their respective research contexts.

2.7 Language support in EMI

In recognition of the prevalence of student challenges, recent research has brought attention to the importance of language support provision to better prepare students for EMI study (e.g. Galloway & Ruegg, 2020; Galloway & Rose, 2021). For example, Thompson, Aizawa, Curle and Rose (2019) observed that an in-sessional language support program, organised concurrently alongside regular EMI courses within a bilingual business program in Japan, contributed to the development of students' academic language skills as well as other soft skills, such as self-efficacy to study through EMI. They stressed that the students attended "the language support EMI preparatory program" to address their language problems before each of their EMI lectures. Similarly, Chang, Kim and Lee (2017) in South Korea revealed that their language support program also assisted students in preparing for EMI study, in the areas of "writing", "communication" and "presentation" (p. 515). Unlike Thompson et al. (2019), many of their participants undertook these support workshops as pre-sessional programs prior to embarking on their EMI courses, seeing that their sample consisted of mostly novice EMI learners who "have never written academic research papers in English" (p. 515). These studies in addition to other prominent researchers in the field, such as Galloway and Rose (2021) all appear to suggest that there is a rising demand for integrated language support in preparation for EMI courses, but there also appears to be a variation in how the support provision is introduced across different contexts, some of which are introduced as in-sessional courses (Thompson et al., 2019), while others are pre-sessional courses (Chang et al., 2017).

Observing this variety of ways in which language support provision is introduced in institutions in many parts of the world, Macaro (2018) proposes overarching models

of EMI, namely the "selection model", the "preparatory year model", the "concurrent support model", and the "ostrich model" (p. 232). The in-sessional and pre-sessional courses in the abovementioned studies by Thompson et al. (2019) and Chang et al. (2017) can be labelled as the "concurrent support model" and "preparatory year model", respectively, while these courses also shared the feature of the "selection model" as students in their contexts were assessed their English language proficiency via English placement tests, which determined the level of linguistic support required. Conversely, the "ostrich model" is used, in contexts, where neither entrance requirement nor linguistic support is offered, which is the model that Macaro (2018) refers to as "unacceptable" (p. 233). While more research is needed to inform how EMI institutions should select and implement the most suitable model(s) according to their own local contexts and needs, Macaro (2018) suggests that each of these models have diverse consequences on their operation, such as, "teachers' daily professional lives and the interpretation of their roles" (p.231), and ultimately determine whether their EMI provision is successful.

Another decision for institutions to make is whether the support provision is offered in the form of EAP, ESP or the mixture of the two, referred to as English for specific academic purposes (ESAP) (Flowerdew, 2016). Recently it has become more apparent that the EMI research community places much priority on ESP (or ESAP), seeing its direct relevance to students' academic disciplinary needs, especially "in contexts where English proficiency is assumed to be higher" (Galloway & Rose, 2021, p. 5). Jiang, Zhang and May (2019), for instance, claim that ESP offers discipline specific language skills that students can directly integrate into their EMI studies and overcome their linguistic challenges, noting that "increasing access to ESP provision that is fine-tuned to the language issues in genuine EMI classrooms is crucial" (p. 107). Similarly, Thompson et al. (2019) found that their ESP courses integrated within their bilingual business program covered subject-specific concepts and key business vocabulary required for textbook and lecture comprehension in their regular EMI courses. Although EAP plays a crucial role in improving students' general academic English skills, ESP and ESAP in many contexts appear to be more favourable options as they offer more specific language support tailored for the disciplinary content being studied by the students and are more directly linked to specific academic needs, such as writing genres (Galloway & Rose, 2021).

To summarise, drawing on from Macaro's (2018) models of EMI, this section has discussed a variety of ways in which language support provision is integrated into universities' existing curriculum, in the form of EAP, ESP or ESAP programs. Seeing that much of the research reviewed in this section largely addressed the relationship between challenges and L2 proficiency, I will now discuss the nexus between the two, with an attempt to highlight the significance of L2 English proficiency in EMI programs in the section below.

2.8 English language proficiency in EMI

A wealth of EMI studies suggests an intersection between students' English proficiency and academic success in EMI, establishing the significance of L2 proficiency as a significant factor facilitating or hindering successful content development through EMI (e.g. Lei & Hu, 2014; Trenkic & Warmington, 2019; Hu & Wu, 2020). In emerging contexts where EMI is still in its infancy, it is particularly common that undesirable educational outcomes emerge as a result of students' limited L2 proficiency, such as reduced content coverage and inadequate content comprehension (Ali, 2013). In particular, some research shows clear differences in students' learning experience according to L2 proficiency, that is higher proficiency students only face issues associated with learning new academic subjects while their counterparts with lower proficiency experience considerably more onerous challenges in EMI programs (e.g. Lin & Morrison 2010; Evans & Morrison 2011).

Some predictive research has adopted direct measures of academic success to explore the association between L2 proficiency and academic success in EMI (e.g. Cho & Bridgeman, 2012; Martirosyan, Hwang & Wanjohi, 2015). While results remain mixed and inconsistent, some have found a statistically significant positive relationship. For instance, Sánchez-Pérez (2021) in Spain explored the relationship between disciplinary-literacy variables (e.g. text structure, cohesion, vocabulary and grammar) and students' academic grades given by lecturers based on the writing performance of EMI engineering laboratory reports (n=136), suggesting that texts with higher frequency rates of these linguistic variables were considered more proficient by lecturers than those with a smaller number of these linguistic features. The regression model containing these linguistic variables explained 48% of the variance in the students' academic grades, revealing the importance of academic

writing skills in engineering written texts. Also using regression analysis, Rose, Curle, Aizawa and Thompson (2019) revealed that of the variance in content score attainment of university students majoring International Business (n= 146), 25.7% and 36.5% was explained by general L2 English proficiency and ESP course grades, respectively. When including these linguistic variables alongside motivation in a multiple regression model, the overall model explained a total of 44.5% of content score achievement with motivation not identified as a significant predictor, indicating the significance of L2 proficiency in academic success in EMI. Based on correlation analysis, Cho and Bridgeman (2012) revealed that TOEFL scores of university students (n= 2,594) in the US explained 4% of the variance in GPA (r=.20) of who was still a significant result but the magnitude of its correlation was rather small.

Conversely, other studies exploring the predictive effects of IELTS found no correlation with academic success (e.g. Krausz, Schiff, Schiff & Hise, 2005; Woodrow, 2006). Investigating undergraduate L1 Chinese students studying via EMI at a university in South Korea (n=138), Yan and Cheng (2015) adopted multiple regression analyses to find that self-assessed English proficiency (overall, reading, writing, listening and speaking skills) had a significant but small correlation (r=.21), and was not a significant predictor of undergraduate GPA after controlling for self-assessed Korean proficiency. This indicates that English, as a third language after Korean, had limited effects on Chinese undergraduate GPAs in South Korean universities. In light of these conflicting findings, it is worth noting that English proficiency in these prediction studies was measured via different types of proficiency tests and self-assessed instruments, and outcome measures (often GPA or final course test scores) were also highly contextualised (often including components of numerous courses). Accordingly, such measures were not comparable across different educational contexts.

To summarise, there is a paucity of empirical research that measures the impact of L2 proficiency on EMI academic outcomes while taking other factors into account, such as “the level of prior knowledge of the academic concept” (Macaro, 2020, p. 267). This is mainly because previous research has measured success largely based on perception-based methods, resulting in qualitative enquiries rather than quantitative or mixed methods enquiries (Macaro et al., 2018). In the same vein, existing quantitative findings have mostly drawn from correlational designs exploring the relationship

between proficiency and academic performance rather than controlling for possible effects of other variables (Yan & Cheng, 2015). Although L2 proficiency is paramount to higher levels of academic success, these mixed findings have pointed to an array of complex interactions between factors. Thus, the current study explores this relationship, by examining whether, and to what extent, L2 proficiency predicts academic outcomes in EMI programs when controlling for other factors, such as prior content knowledge and the magnitude of challenges, which is sought to further offer the more solid and wholistic picture of the effectiveness of EMI.

2.9 Language in science

While much of the research reviewed in section 2.6 above has predominantly focused on general language-related challenges which are commonly expressed in all fields of study, research into science education suggests that science as an academic discipline introduces a specialised and technical language that makes it unique but all the more challenging (e.g. Marshall, Gilmour & Lewis, 1991; Adams, 2003). For example, Schleppegrell (2007) identifies linguistic structures integrated into in mathematics that are distinctive from how language is typically used in other academic discourses, indicating that these discipline-specific forms present challenges to students (e.g. dense noun phrases, mathematic symbolic notation). Norris and Phillips (2003) also argue that students first need to be literate in the discourses of science (i.e. language in the learning of science) to become knowledgeable about scientific concepts (i.e. science content knowledge). In the learning of chemistry, Cooper and Stowe (2018), for example, observe that high frequency words (e.g. 'water', 'light', 'small') are often substituted with technical words (e.g. 'aqua', 'photo', 'micro'). Many of these words are then combined to make compound words, such as "photosynthesis" or "micromolar". Chemistry also draws on numerous semiotic (meaning-creating) features to express meanings that ordinary language cannot express, such as chemical symbols and graphic representations (e.g. the structural formula of chemical compounds). These examples show that science has its own ways of handling language to construct knowledge, and students need to learn to understand the language to participate in the learning of science.

Investigating secondary school pupils' comprehension of scientific vocabulary, some seminal studies in England (Cassels & Johnstone, 1983) and Australia (Gardner,

1972) found that many of their L1 English participants encountered difficulty in understanding non-technical words which contained scientific meanings, such as "evacuate", "sequence" and "accumulate". Seeing that learning the language of a new discipline is part and parcel of learning the new academic discipline, Yore (2008) argues that even L1 English students still need to learn "an English-dominated domain of science" as their second or third languages (p. 9).

While language issues in science are not confined to L2 learners, an increasing body of EMI research into science education suggests that the magnitude of their challenges may become more severe when learning occurs through L2 English. For example, a large-scale study in Papua New Guinea involving L2 English students in Grades 7 to 12 and their first year of university (n=2,111) was carried out by Marshall, Gilmour and Lewis (1991) to investigate their comprehension of 45 academic English words frequently used in science. The students' comprehension of these 45 words were examined in the following four formats: "synonym"; "sentence"; "science context"; and "non-science context". The study found that only 11 out of the 45 words achieved the satisfactory performance (i.e. 70% or more participants gave the correct response) in all four formats. It was particularly interesting that the participants had difficulty with certain words not only in the "science context" but also in the "non-science context". For example, only 49% and 14% of the participants gave the correct responses to words, such as "consistent" in the "science context" and "non-science context", respectively. The authors explained that the participants might have confused "consistent" with other words, such as "constituent" in the sense of "part of". While this study did not recruit an L1 English comparison group to see whether their performance would have been higher in the "science context" and "non-science context", the study still demonstrated that many of their L2 English students had particular difficulty with English words that were graphologically or phonetically similar, such as "consistent" with "constituent", and "component" with "opponent".

More recently, adopting the same vocabulary list from Marshall et al. (1991), Vladusic, Bucat and Ozi (2016) in Croatia explored a total of 82 university students' understanding of technical words, representations, symbols, and general academic words used in chemistry. The findings of their study were consistent with that of Marshall et al. (1991), revealing that many of the L1 Croatian students faced language comprehension problems because many of the similar sounding technical words

contained different meanings, and many of the non-science words were used in the science discourses containing different meanings (i.e. dual meaning vocabulary words). In a perception-based study, Kim, Kweon and Kim (2017) adopted a questionnaire to explore challenges faced by L2 English engineering students in South Korea (n=524). The study reported that less than 2% and nearly 60% of the participants were satisfied with their learning quality through EMI and Korean-medium instruction (KMI), respectively, indicating that most students favoured KMI over EMI. Their subsequent interviews revealed that many of the participants faced more difficulty in understanding key engineering concepts and solving problem worksheets when learning through EMI than KMI. Similar findings have also been found in other science education studies, such as Marais and Jordaan (2000) in South Africa, Song and Carheden (2014) in the US, and Chan (2015) in Hong Kong, all pointing to both the general language-related challenges (e.g. the use of conjunctions, such as "therefore") and challenges specific to the learning of science (e.g. components of chemical equations, such as [NO_2], and $2NO_2$).

In summary, the findings of the studies reviewed in this section all appear to indicate that students learning science through L2 English encountered challenges both in terms of their science content knowledge and in their command of the English language. Given the distinctive way science operates, this study seeks to tease out the perceived challenges of students learning Chemistry and deepen understanding of the discipline-specific and linguistic challenges of science learning in EMI.

2.10 Examining the outcomes of EMI

2.10.1 Success and outcomes of EMI

As the number of EMI programs have rapidly increased around the world, so have concerns regarding the outcomes of EMI. However, 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 the local educational needs. The current literature broadly discusses "success" in terms of language acquisition, content acquisition, or the acquisition of both language and content. Some researchers (e.g. Dafouz & Camacho-Miñano, 2016) regard EMI as a cost-effective educational approach to develop both language and content knowledge. Conversely, others attribute EMI most notably to the development of content knowledge through

English. Zaif, Karapınar and Yangın Eksi (2017) operationalised success solely based on academic course grade scores in an EMI business administration course in Turkey. Thompson, Takezawa and Rose (2021) adopted composite midterm and final test scores from an EMI finance course as predictors of academic success in an international business program in Japan, indicating that success was measured via content exam scores.

Seeing that educational outcomes in higher education are often considered as more than just gaining substantial content knowledge, researchers have triangulated perception-based data with score gain data to explore a perceived benchmark of success. To unpack perceived benefits of EMI, Rose, Curle, Aizawa and Thompson (2019) in Japan (n=146), and Xie and Curle (2020) in China (n=100) examined success in EMI in their international business programs. Their interview participants interpreted EMI to be successful if they could learn how to apply and develop new knowledge in the chosen academic subject (Xie & Curle, 2020) and improve long-term career prospects (Rose et al., 2019). Recently, more studies have measured outcomes of EMI in terms of career prospects (e.g. Huang & Curle, 2021; Altay & Erçin, 2020). Therefore, these studies have attributed success to the development of more "soft" skills through EMI.

As discussed in section 2.2, the policy objectives of EMI in Japan do not explicitly aim to develop language proficiency; yet students are expected to reach satisfactory levels of disciplinary knowledge. Accordingly, this current study operationalises success in EMI in terms of students' mastery of content knowledge while also recognising many other substantial benefits of EMI as being perceived by students, including language development.

2.10.2 A review of academic outcomes studies of EMI

A growing body of research has evaluated the outcomes of EMI in relation to students' academic performance. For the most part, studies have shown either positive or negative evidence. Positive studies reveal that there are no clear discernible differences in the academic outcomes between EMI and L1 medium students or that EMI students achieve even higher academic outcomes than their L1 medium counterparts (e.g. Joe & Lee, 2013; Dafouz, Camacho & Urquia, 2014). Negative studies suggest that EMI causes a detrimental impact on students' academic performance (e.g. Kim, 2002; Klaassen & Graaff, 2001). Despite the mixed findings,

there is increased concern about the scarcity of this research avenue (e.g. Park, 2007; Byun et al., 2011). Macaro et al. (2018) observe that "the research evidence to date is insufficient to assert that EMI benefits language learning nor that it is clearly detrimental to content learning" (p. 36). The findings of their literature review points towards an urgent need for further "cost-benefit" research to investigate the academic outcomes of EMI (Macaro, 2018). In this section, I turn to discuss previous studies which have looked into the effects of EMI on students' content learning.

In order to review previous outcome research as systematically as possible to determine what studies to review, I established a set of inclusion and exclusion criteria. Following the review protocols developed by Macaro et al. (2018) and Pecorari and Malmström (2018) in their systematic review, the first phase of selection involved conducting separate literature searches in databases, including Web of Science, SCOPUS, the British Education Index, and ERIC using the following keywords (both AND/OR was used to separate the phrases): "English medium instruction"; "English medium of instruction"; "English as a medium of instruction"; "CLIL"; "higher education"; "effects"; "effectiveness"; "outcomes"; "content learning"; and "academic achievement". In this first round, only articles indexed in any of the four databases and published after the year of 2011 were included in the search, which marks the surge of academic interest in EMI research according to Macaro et al. (2018).

The second phase of selection involved careful examination of the titles, abstracts and research questions of the articles sourced in the first round of selection to determine what studies to include in the review. The following criteria were created to narrow down the set of studies which 1) examined the effects, effectiveness and outcomes of EMI; 2) focused on academic outcomes instead of other types of outcomes (e.g. English language, motivation, career prospects); 3) recruited students studying academic subjects other than English language (e.g. EAP or ESP) through English; 4) took place in higher education settings; 5) adopted empirical data collection methods (e.g. interviews, pre-post-test, observations); and 6) framed research as CLIL but the actual instructional practice had no explicit emphasis on language learning outcomes.

I also excluded non-empirical studies, such as conference proceedings, book reviews, reports, and masters' and PhD thesis, which meant that only peer-reviewed journal

articles were included in the review. Subsequently, articles that were not initially found in the four databases, but nevertheless satisfied all the above criteria, were also added to the final review, including five studies published before the year of 2011.

Using these principles, the two rounds of search generated unique records of 23 studies which all focused on the outcomes of EMI in relation to student' content learning in the context of higher education. All the studies were documented on an excel sheet, with the title, year of publication, author(s), journal, country, research methods, and summary of findings. Overall, among these 23 studies, there were 12 score gain studies and 11 perception-based studies. In the following sections, I now discuss the findings of these selected articles according to score gain and perception-based research to identify the current limitations and priorities for future research.

2.10.3 EMI outcome studies to date

2.10.3.1 Score gain research

With respect to methods adopted to evaluate the effects of EMI, Macaro et al. (2018) found that among 83 HE studies analysed in their review, only seven adopted objective tests. Of these, five employed standardised testing instruments and two adopted test instruments devised by researchers specifically for their projects. Only a small proportion of empirical evidence is based on objective instruments. The following studies reviewed in this section examined the effects of EMI in relation to score gains measures (e.g. final grades, pre-post-tests). Among the 23 outcome studies reviewed, 12 studies were based on some form of objective research instruments. Of the 12 studies, 10 and 2 studies offered positive and negative evidence, respectively.

Among the scant empirical research into disciplinary gains of EMI via objective measures is Park's (2007) study, which examined the effect of EMI on content learning in South Korea. Without a direct comparison group, data were gathered from 51 EMI students enrolled in an introductory theoretical linguistics course (e.g. syntax, morphology, phonetics). Their gains in disciplinary knowledge were recorded by pre-post-tests. The tests were developed specifically for the study in a short answer or true/false format, which were distributed at the start and the end of a two-hour lecture period. The findings revealed that statistically significant gains in content knowledge were obtained between time one and two, with the improvement not being significantly related to the students' English proficiency. The study also recorded students' self-assessed improvements in their English knowledge, in the areas of

reading, listening and vocabulary. Although the content and language gains were successfully documented, the study did not achieve a baseline group to account for any potential initial difference between the EMI and L1 medium instruction groups.

The absence of a direct comparison group is one of the most common limitations in many of the score gain studies reviewed, including EL-Daou and Abdallah (2019) who examined the impact of CLIL on undergraduate students' maths test performance in Lebanon (n=21). The descriptive statistics and paired samples test of pre-test and post-test scores of a maths achievement test administered as part of their regular curriculum showed a statistically significant difference in mean test scores at the start and end of the "40 instructional periods" (p. 6), indicating disciplinary gains among the CLIL students. The participants' reflection logs also recorded students' beliefs and feelings about their learning experience, suggesting that CLIL enhanced their motivation to learn new technical words and academic concepts more easily than learning through their L1. Despite this positive evidence, 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.

In the same vein, Yang (2015) adopted a mixed methods approach to investigate the impact of CLIL on students' academic outcomes in Taiwan. The questionnaire and interview findings suggested that students from 13 different academic disciplines (n=29) perceived that their content learning was somewhat ineffective compared to learning through L1 Chinese. Meanwhile, the study also demonstrated a moderate but statistically significant positive correlation between their language gains in four skill areas measured via the General English Proficiency Test and content performance measured via academic grades from content courses ($r=.572$, $p=.001$), suggesting that those with higher English proficiency could achieve satisfactory content knowledge. In other words, once reaching a certain linguistic threshold, they could learn content through EMI satisfactorily. Nevertheless, the study only recruited EMI students, and it was inconclusive as to whether CLIL had any impact on their academic performance.

Recognising these limitations, Joe and Lee (2012) conducted a study in South Korea with a direct comparison group to examine the effect of EMI on medical degree students' content learning, and relationships between comprehension of EMI lectures and general English knowledge. A total of 61 L1 Korean students undertook both

EMI and Korean-medium-instruction (KMI) lectures to compare lecture comprehension between the two conditions. The study employed 14 T/F and seven cloze-type tests to compare the disciplinary knowledge at points one and two, also recording students' general English knowledge that was determined from their self-reported scores on a standardised English test (TOEFL) (time one only). The pre-post test questions were administered in the language of instruction; the items introduced in English in the sessions were provided in English and those taught in Korean were answered in Korean. The study provided positive evidence that: 1) the level of EMI lecture comprehension was not related to whether English or Korean was used; 2) their final grades based on the post-tests significantly improved for both groups and no significant difference was found between the two groups; and 3) students' English proficiency did not influence the degree of their lecture comprehension nor final grades. Despite obtaining a comparison group, the design was limited by a lack of authenticity: rather than observing courses within the existing curriculum, the group of students in the study had to attend one-off medical lectures specially designed for the study. One lecture was delivered in English which was followed directly by one delivered in Korean. The study would have achieved higher level of ecological validity if it had been part of the regular EMI and KMI courses over a course of one semester.

Much of the EMI outcomes research reviewed so far has pointed to the importance of objective measures, direct comparison groups, and ecological validity. One of the most comprehensive studies to date to consider these methodological limitations is a study conducted by Dafouz et al. (2014) in Spain. Their study explored the effect of EMI on students' academic performance verified in the results of coursework and final exam grades when compared to their counterparts' discipline learning in Spain. It sampled 106 L1 Spanish students from three different academic disciplines (history, accounting and finance). The authors also took into consideration some crucial variables that could affect learning achievement, such as English proficiency and university access grades for both EMI and Spanish-medium-instruction (SMI) courses, to make the two groups comparable. The results illustrated that both EMI and SMI cohorts yielded similar academic achievement, indicating that the language of instruction did not compromise students' learning. However, comparing the final grades was problematic, noted by the authors themselves that "differences as to the

precise criteria applied in the final evaluation of students" performance seem to be largely dependent on individual teacher decisions' (p. 229). Hence, whether the academic achievement was accurately measured remains questionable. In addition, the authors suggested that qualitative data was needed to clarify the assessment criteria across the three subjects, stating that "further research needs to look from a more qualitative perspective into the precise nature of the coursework activities and the assessment practices" (p. 232). Accordingly, further research is warranted to qualitatively inform how student grades are decided and what differences might exist between two groups.

Expanding on Dafouz et al. (2014), Dafouz and Camacho-Miñano (2016) explored Spanish students' academic performance in an EMI financial accounting program. 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. However, the study was again based solely on the final grades; no qualitative data were obtained to inform any contextual information regarding how the EMI and SMI groups were comparable in terms of key variables, such as prior content knowledge. Civan and Coşkun (2016) compared academic performance of EMI and Turkish medium instruction (TMI) students in terms of their semester point averages across nine different academic disciplines in a university in Istanbul, concluding that TMI was more effective than EMI. However, their study also did not provide sufficient contextual information regarding how the course grades were comparable between the two learning conditions across nine disciplines. The same methodological issue about the absence of qualitative data was also identified in Hernandez-Nanclares and Jimenez-Munoz (2017) who compared 172 EMI and 482 Spanish medium students solely based on course grades obtained via final examination assessments.

More recently, Yang, O'Sullivan, Irby, Chen, Lin and Lin (2019) conducted a case study of medical degree students to document differences in content learning between EMI and Chinese medium instruction (CMI) programs in China. Data consisted of 1) EMI and CMI students' entrance exam scores collected in 2009 and 2010; 2) graduation exam scores collected in 2015 and 2016; and 3) students' and teachers' perceptions of the program via focus groups and questionnaires. The study found no significant differences in both the entrance and graduation test scores between the

EMI and CMI students, indicating that EMI had no adverse effects on their content learning. However, instead of using the entrance exam scores as an indicator of prior knowledge, the pre-test should have been the same test as the post-test to measure the score gains between the two time points while controlling for the prior content knowledge. Also, no contextual information was collected to confirm that the students undertook the entirety of their six-year medical degree in English, or to account for any potential contextual differences between the CMI and EMI degree programs.

Finally, of the 12 score gain studies reviewed, only one study conducted by Guo, Tong, Wang, Min and Tang (2018) incorporated classroom observation to record contextual and institutional information, such as admission policies and instructional practices to ensure ecological validity. The study observed both the EMI and CMI lectures of the television art program in China to ascertain classroom-related contextual differences, including student and teacher language use and classroom activities, which were shown to influence the degree of students' disciplinary gains. Although classroom observation is a critical method to measure the fidelity of EMI implementation (e.g. language use) and speculate on potential contextual, linguistic, institutional and demographic variables (e.g. Costa & Coleman, 2013), most of the studies reviewed did not necessarily include observations. Other quasi-experimental studies reviewed also fell short of capturing comprehensive contextual variables via lecture observation (e.g. Zaif et al., 2017; Rashed, Hashimoto & Oyabu, 2018). Recognising this methodological limitation, the current study draws on rich sets of qualitative data, including observations and interviews to shed light on various contextual factors that are believed to be highly specific to the research context.

2.10.3.2 Perception-based research

Research into the effects of EMI in various contexts have largely been based on stakeholders' self-assessed views surrounding the improvement of content knowledge (Macaro et at., 2018). The previous perception-based research into the effects of EMI has generated mixed findings. Macaro et al. (2018) offer an account for these conflicting findings, noting that "no explanation can be arrived at for these different self-assessed findings" (p. 33) and "these [inconsistent] findings are not surprising" without "using objective tests rather than self-report" (p. 35). Among the 23 outcome studies reviewed in this section, 11 studies measured students' academic outcomes

based on perception-based research instruments. Of the 11 studies, three and eight studies offered positive and negative evidence, respectively.

One perception-based study was conducted by Huang (2015) in Taiwan who employed questionnaires containing 25 six-point Likert-scale items (N=157). Participants rated nine statements each concerning "learning achievement", "motivation" and "anxiety" on a six-point scale (1 = very ineffective, 6 = very effective), such as "the EMI courses help to improve my English listening and speaking ability" (p. 75). Among the two other constructs explored, students rated "learning achievement" at the highest mean score of 4.24 (composite mean scores of all the nine items). Despite this positive result, it was inconclusive as to whether their questionnaires actually measured students' disciplinary achievements because an item such as "as a whole, the EMI courses are helpful to me" could easily be interpreted as participants understanding the 'learning achievement' of EMI with respect to other outcomes such as career advancement and L2 development, highlighting an issue concerning the questionnaire used. Other positive evidence was obtained by Kang and Park (2005) who investigated the effectiveness of EMI at an engineering university in South Korea via a 17-item questionnaire. Upon establishing a strong correlation between students' language proficiency and opinions towards EMI (n=366), the study found that those with higher proficiency expressed positive opinions about the effectiveness of understanding lectures and textbooks through EMI. Other EMI effectiveness studies, such as Fidan and Soruç (2018) in Turkey also yielded positive evidence via questionnaires, showing a higher degree of motivation among EMI students compared to those in the Turkish medium instruction group.

When other methods were incorporated to achieve a closer examination of students' learning experience, more negative evidence has surfaced, suggesting that some students were unable to achieve sufficient content knowledge through EMI. Lei and Hu (2014) conducted a series of interviews with six EMI and four Chinese-medium instruction (CMI) undergraduates undertaking a Business Administration program in a university in southwestern China, concluding that, compared to their CMI counterparts, EMI students were neither satisfied with the richness of content taught in English, nor with the improvement of their English ability. Their interview findings indicated that all the EMI participants criticised the shallowness of the academic content, and highlighted that learning conditions which have previously been found

(Hynninen, 2012; Smit, 2010) to contribute to L2 learning (e.g. explicit grammar teaching, and interaction between students and teachers) were noticeably absent from the EMI in comparison to the CMI classroom. This finding complements other findings (e.g. Kung, 2013) and mirrors Chapple's (2015) assertion that "the idea that merely taking a content class taught in English will lead to substantial content gains is dubious" (p. 4). Another study by Başıbek et al. (2014) in Turkey based on the results of a survey of 63 lecturers at two EMI universities also echoes this negative finding, revealing that although lecturer participants state that 'EMI would contribute to [students'] English proficiency levels as well as their fields of study', they also believe that 'learners are not proficient enough to learn subject matters in English' and note that 'lecturing in Turkish provides deeper and clearer understanding in terms of the content of the lesson' (p. 1823). Negative findings were also identified in the studies of Sert (2008) in Turkey, and Kim and Yoon (2018) in South Korea, who adopted interviews as a principal data collection method.

In a study that found EMI to be less effective, Airey and Linder (2006) conducted in-depth qualitative research in Sweden to examine the association between the medium of instruction (i.e. English or Swedish) and the learning experiences of 22 undergraduate physics students at two Swedish universities. The participants were enrolled in bilingual programmes where lectures were offered in both languages. Lectures were video-recorded first and subsequent semi-structured interviews were conducted with all the participants to gain an in-depth understanding of their EMI experience using selected excerpts of the video-recording during stimulated recall. The researchers concluded that although students initially indicated no difference in their learning experiences in Swedish and English, it became clear, after stimulated recall, that there were some crucial differences. For example, all the students initially did not think that Swedish or English was an important factor to determine the level of their lecture comprehension. However, stimulated recall highlighted various linguistic problems resulting from their insufficient linguistic knowledge (such as increased workload, inability to understand lectures). In addition, students also initially expressed that they were making progress in gaining content knowledge regardless of English or Swedish, but the researchers revealed that Swedish was more efficient. Overall, the study offered a valuable insight into the usefulness of in-depth qualitative method, such as stimulated recall and interviews.

Based solely on lecture observations, Hu and Duan (2019) carried out a cross-disciplinary study of the effects of teaching medium and disciplinary background on teacher questions and student responses at a Chinese university. Ten EMI and ten Chinese medium instruction (CMI) content courses were observed and analysed, comprising of ten hard and ten soft academic disciplines. Results showed that the majority of lecturer questions and student responses in the EMI lectures were cognitively and linguistically simpler than the CMI lectures. In other words, many of the interactions "focused on rote memorization and comprehension of factual information" (p. 314). As high-order thinking skills would require higher quality of student-teacher interaction, the study concluded that EMI hardly achieved its purported dual aim of developing both content and language. This negative finding could also be related to the research context. For instance, a majority of the sampled EMI courses were introductory level courses which did not require much classroom interaction, and that the participants of the study, on average, only obtained IELTS 5.5, which was not sufficient English proficiency for successful EMI learning.

In another study which found a negative impact of EMI, Cho (2012) investigated the perceptions of the implementation of EMI at a science and engineering university in South Korea, highlighting the adverse effects of EMI on content teaching and learning. Questionnaires were first administered to 41 faculty members, 439 undergraduate students and 403 postgraduate students. Follow-up interviews were conducted with seven professors, ten undergraduate students and eight postgraduate students. The results showed that due to the teachers' and students' low English proficiency, EMI was ineffective in delivering content. For example, students asked fewer questions, and student-teacher interaction or among students were repressed. The interview data also highlighted the absence of structured language support provision, insufficient lecture comprehension, and low confidence among students. More than 68% of the postgraduate and 60% of the undergraduate students indicated that their class participation was lower when learning in English. Similar negative findings were also identified in the studies of Hellekjær (2010) in Norway, and Tseng, Pilcher and Richards (2020) in Taiwan, pointing to the potential ineffectiveness of EMI.

Overall, it is noteworthy that a higher proportion of perception-based studies generated negative evidence than the score gain studies. Of the 12 score gain studies,

10 studies offered positive evidence whereas only two of the 11 perception-based studies offered positive evidence. In other words, when indirect methods, in particular interviews are incorporated, a darker picture of EMI implementation appears to emerge as students can elaborate on their academic challenges associated with EMI study. In light of these conflicting findings, the studies reviewed above unveil some methodological issues to consider when evaluating the effectiveness of EMI. In the following paragraphs, I will discuss these key points that have emerged.

2.10.4 Strengths and limitations of previous outcomes studies

To summarise what has been identified in this section, I will discuss limitations of previous research and suggest how the current study addresses these limitations. Firstly, previous studies highlight the importance of mixed methods research that incorporates the triangulation of direct and indirect data collection methods. In the 23 studies reviewed, some of the research used exam results and final grades as well as a combination of indirect and direct data collection methods (e.g. Park, 2007; Yang, 2015; EL-Daou & Abdallah, 2019), whilst others used only interviews or questionnaires that were more impressionistic, merely drawing from stakeholders' opinions (e.g. Kang & Park, 2005; Cho, 2012; Hu & Duan, 2019). Accordingly, the current study adopted perception-based methods to achieve an in-depth, nuanced understanding of participants' viewpoints concerning EMI implementation alongside score gain results of disciplinary knowledge obtained from objective test measures.

Secondly, among the 23 studies reviewed, 12 studies involved a direct comparison group of EMI with instruction conducted in the respective L1 of the context under investigation (e.g. Rashed, Hashimoto & Oyabu, 2018; Dafouz et al., 2014; Yang et al., 2019; Zaif et al., 2017). In one of the studies, an L1 comparison group was adopted but created only for the purpose of their study to emulate an experimental design, resulting in the low degree of ecological validity (Joe & Lee, 2012). The current study therefore recruited Japanese medium instruction students as a baseline group that were comparable to their EMI counterparts to observe the effects of EMI while ensuring the high degree of ecological validity.

Thirdly, of the 12 score gain studies, five cross-sectional studies (e.g. Dafouz et al., 2014; Civan & Coşkun, 2016; Dafouz & Camacho-Miñano, 2016) adopted direct measures of GPAs and course grades but fell short of capturing students' prior content

knowledge. Yang et al. (2019) conducted longitudinal research accounting for prior content knowledge, but used different pre-test and post-test, meaning that they could not record actual gains in content test scores. Thus, the current longitudinal study followed the students' content learning experience over the course of one semester using the same pre-test and post-test while accounting for their prior content knowledge.

Finally, of the 12 score gain studies, only one study adopted classroom observation as a means to collect detailed contextual information alongside comparisons of content test scores (Guo et al., 2018), indicating that comprehensive contextual factors were rarely examined. Thus, there is a paucity of empirical research that adopts observation in EMI contexts. Recognising that quasi-experimental research introduces various contextual, institutional, demographic and linguistic variables, the current study used observation as well as other forms of qualitative data to capture the underlaying quasi-experimental factors that could affect students' content learning outcomes.

Overall, outcomes studies vary considerably in their designs and methodologies (e.g. Hoare & Kong, 2008; Yip, Tsang & Cheung, 2003), and consequently, there yet to be a general consensus in relation to research evidence into the effects of EMI. Macaro (2018) observes this sheer absence of conclusive evidence as "a sad state of affairs for a global phenomenon that is accelerating at such a fast pace" (p. 183). The question of how much and to what extent EMI enables disciplinary gains thus still remains unsolved and thus will be investigated in this thesis.

2.11 Research questions and conceptual framework

In conjunction with the literature set out above and gaps identified in the extant literature, that the following research questions have been devised alongside a conceptual framework for the study. The conceptual framework is outlined in Figure 2 below, highlighting the relationships between key learner, contextual and institutional factors established in previous research. Indicated in the framework are focal points (numbers 1–3), which correspond to the research questions in this study:

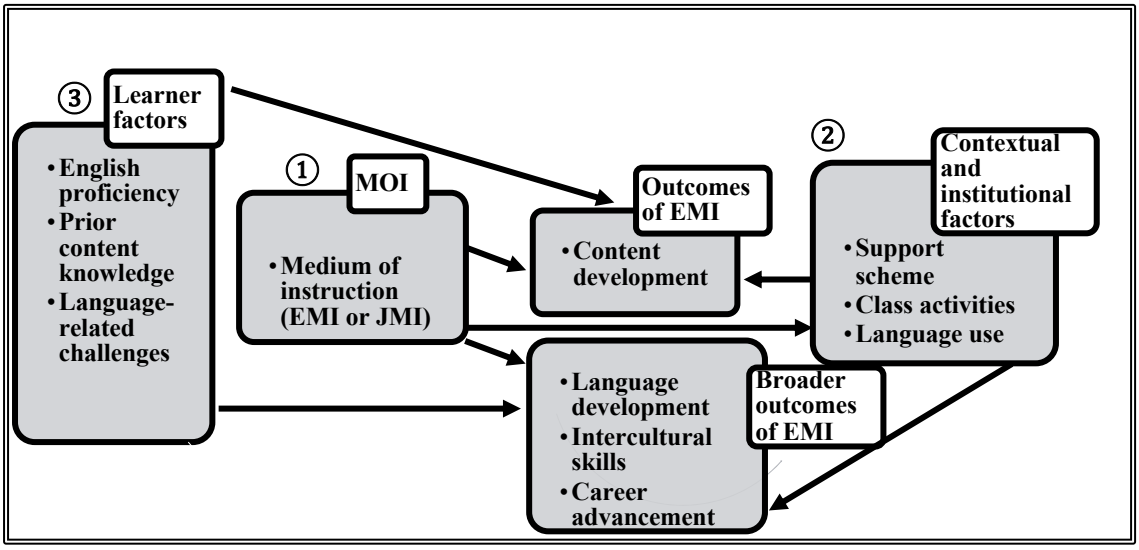
- RQ1: Is there an impact of medium of instruction on undergraduate students' content learning?

- RQ 1.1: To what extent does the achievement of content learning of the EMI and JMI students differ? (Directly measured content knowledge)
- RQ 1.2: How do the EMI and JMI students perceive the medium of instruction to affect their challenges and achievement of content learning? (Perceived content knowledge)

- RQ2: In what ways do the two mediums of instruction differ? What context-related differences do the students ascribe their challenges and achievement of content learning?

- RQ3: Are there relationships between the EMI students' English language proficiency, academic English challenges, prior content knowledge, and content knowledge achievement?
  - RQ 3.1: To what extent factors, such as the EMI students' English language proficiency, prior content knowledge and perceived EMI challenges are associated with content knowledge achievement? (Directly measured relationship)
  - RQ 1.2: How do the EMI students perceive English language proficiency to affect their perceptions of challenges towards EMI study? (Perceived relationship)

Figure 2: Conceptual framework



The central hypothesis of this study is adopted from Macaro's (2018) "cost-benefit" analysis of EMI that "learning content through L2 English leads to at least as good

learning of academic content as learning content through the students' L1" (p. 154). The conceptual framework (Figure 2) thus places the mastery of content knowledge acquisition as the main outcome variable. It also integrates language development as well as other potential outcomes that go beyond the focus on content or language development (e.g. career advancement), what Macaro (2018) calls "a much broader cost-benefit analysis" (p. 185). The key variables included in this framework are those that were found to be influential in the review of previous EMI studies carried out above: medium of instruction; English proficiency; prior content knowledge; and language-related challenges. Nevertheless, EMI content learning has been shown to be affected by a range of other factors, such as high school teaching medium, motivation, previous EMI experience, learning strategies, teaching styles, classroom activities, language use, and institutional support programs. Thus, these institutional, contextual and learner factors shown to contribute to content improvement were also considered as they arose in the study. While investigation into content outcomes of EMI is merely one aspect of the broad cost-benefit analysis of EMI, the current study aims to "better theorize the cost-benefit ratio of EMI" (Macaro et al., 2018, p. 66) by responding to the still largely unheeded call in the field of EMI for "careful theoretical analysis and consistent layers of good quality research" (Macaro, 2018, p. 153). Thus, this study aims to tease out the differences between learning occurring through EMI and the L1 medium instruction in terms of the development of content knowledge.

2.12 Chapter summary and statement of the problem

In summary, the existing body of research reviewed in this chapter has provided valuable empirical accounts of the effectiveness of EMI on students' content learning in a number of higher education contexts around the world. However, this chapter has shown that this account is both theoretically and methodologically insufficient. To echo Macaro (2018), there is no appropriate research evidence of actual measurements of content acquisition in sufficient quantities to confirm whether EMI negatively affects content outcomes (p. 183). To this end, this chapter has shed light on strengths and limitations in the literature, informing the research questions of the current study and a conceptual framework within which this study is located. These research questions and methodological considerations will be discussed in more detail in the next chapter.

Chapter 3 - Methodology

3.1 Introduction

This chapter presents an overview of the research methodology and outlines the methodological framework of the study. First, research questions, analytical framework and methodological considerations are outlined. This will be followed by a brief summary of the research design including setting, data collection, participants, pilot study and research procedure. The chapter continues the discussion of my engagement with quantitative and qualitative data analysis, outlining the analytical approach to the research questions. Finally, recognising a growing awareness of ethical issues in education research, the chapter concludes with a discussion of ethical considerations in relation to the study's compliance to ethical procedures.

3.2 Research questions and analytical framework

The critical review of research literature in the previous chapter has indicated that there has been a growing body of evidence into academic outcomes of EMI in global contexts (e.g. Dafouz & Camacho-Miñano, 2016; Hu & Duan, 2019; Zaif, Karapınar & Yangın Eksi, 2017). However, it has also pointed to a number of methodological concerns, which were highlighted in Macaro's (2018) comprehensive review of EMI. First, few studies have adopted a pre-test post-test design without partialling out students' baseline content knowledge. Second, many of the studies have sought to explore perceived rather than directly measured students' content learning outcomes. Third, an absence of mixed-methods research (e.g. triangulation of test scores and classroom observations) fell short of capturing the complexity of contextual factors which may influence students' outcomes. Thus, based on short-comings in previous research the following research questions have been devised, which are presented alongside the data set and analysis procedures used to answer them (Table 2):

Table 2: Research questions and analytical framework

	Research Questions	Instruments & Data	Analysis
1	Is there an impact of the medium of instruction on students' content learning?		
1.1	To what extent does the achievement of content learning of the EMI and JMI students differ? (**Directly measured content knowledge**)	・Pre-test scores ・Post-test scores ・Gain scores (post-test–pre-test)	・Descriptive statistics ・A Mann-Whitney U Test ・A Wilcoxon Signed Rank Test ・A one-way ANCOVA ・A t-test of gain scores (between-subject)
1.2	How do the EMI and JMI students perceive the medium of instruction to affect their achievement of content learning? (**Perceived content knowledge**)	・EMI student interviews ・JMI student interviews	・Qualitative content analysis
2	In what ways do the two mediums of instruction differ? What context-related differences do the students ascribe their achievement of content learning? (**Contextual factors**)	・EMI and JMI student and teacher interviews ・Unstructured observation fieldnotes	・Qualitative content analysis
3	Are there relationships between the EMI students' English language proficiency, academic English challenges, prior content knowledge, and content knowledge achievement?		
3.1	To what extent factors, such as the EMI students' English language proficiency and perceived EMI challenges are associated with content knowledge achievement? (**Directly measured relationship**)	・Pre-test scores ・Post-test scores ・IELTS scores ・The EMI Challenges Scale	・Descriptive statistics ・Spearman's correlation ・Linear regression analyses ・Multiple regression analyses
3.2	How do the EMI students perceive English language proficiency to affect their perceptions of challenges towards EMI study? (**Perceived relationship**)	・EMI student interviews	・Qualitative content analysis

3.3 Methodological considerations

3.3.1 A mixed-methods design

The current study adopts a mixed-methods design because quantitative or qualitative methods alone do not suffice to address the overall research questions, which require both statistical and nuanced explorations of the data. Mixed methods research integrates both quantitative and qualitative methods within one study to address complex and multi-faceted questions which may not be answered by one of the methods when employed alone (Creswell, Fetters & Ivankova, 2004). A mixed-methods design achieves a triangulation of the methods to corroborate research validity by offsetting the shortcomings of the single method. Quantitative research is criticised for often failing to record in detail contextual factors, whereas qualitative research is prone to its subjective interpretation and focused locality in a specific context In this study, I converged multiple sources of quantitative and qualitative data to accomplish greater validity to achieve a more comprehensive understanding of the research findings. The current classroom-based study necessitated the use of qualitative methods, such as observations and interviews to account for context specific factors to interpret the statistical results with narrative details, including the classroom learning conditions and participants' voices. Furthermore, the relative scarcity of mixed-methods studies in the existing EMI literature also warranted the use of a mixed-methods design. In terms of the order in which quantitative and qualitative data are used within a single study, Creswell, Plano Clark and Garrett (2008) introduced different types of mixed-methods designs. Based on the classification of the sub-designs, I have summarised the temporal sequence of each of the methods in relation to the research questions (Table 3).

Table 3: Mixed-methods research design adopted from Creswell, Plano Clark and Garrett (2008)

Research question	Sub-designs	Description	Integrated interpretations
RQ 1	Embedded Design	Qualitative data (**EMI and JMI student interviews**) were gathered within a quantitative design (**content tests**) to follow up on the quantitative results.	Quantitative Qualitative
RQ 2	Qualitative Design	Factors influencing the students' content knowledge acquisition were examined using	Qualitative

		qualitative data (**observations, teacher and EMI and JMI student interviews**)	
RQ 3	Explanatory Design	First quantitative data were gathered and analysed (**IELTS scores, content test scores, and challenges**), which were complemented by qualitative data (**EMI student interviews**) to explain the findings from the first stage.	Quantitative Qualitative

As indicated in Table 3, I employed an embedded design for research question one where quantitative and qualitative methods were implemented simultaneously, and an explanatory design for research question three where quantitative and qualitative methods were analysed separately in different stages. In contrast, research question two was primarily led by a single use of qualitative methods, adding explanatory dimensions to numerical evidence (Creswell & Plano Clark, 2007). Thus, the data were compared and merged to take advantage of the breadth and depth of each of the methods to arrive at more validated conclusions (Greene, Caracelli, & Graham, 1989).

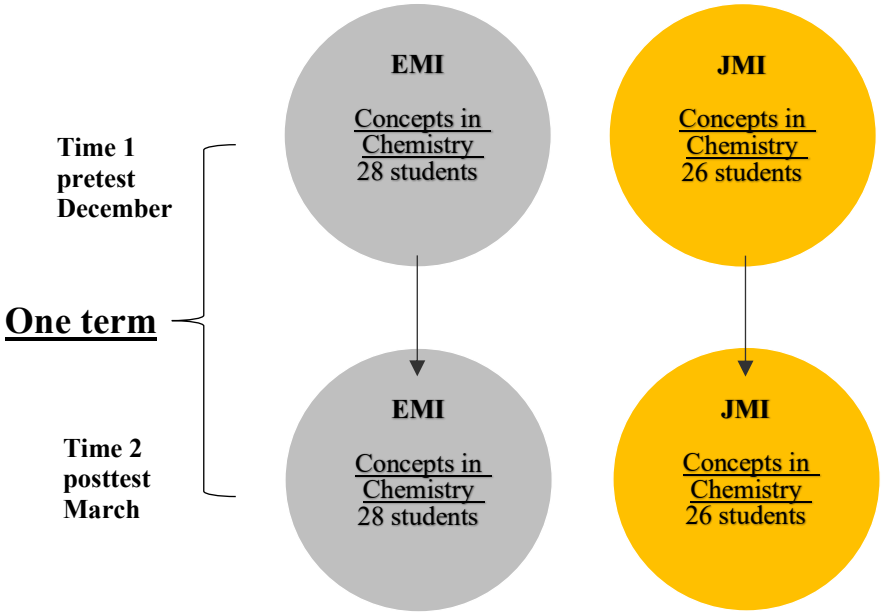
3.3.2 A case study approach

As the study examines the specific EMI provision at a single university, some methodological considerations of a case study approach are worth exploring. Bell (1993) describes the case study as the study of one case which provides a rich look into one specific aspect of an issue in some depth to enhance our understanding of a person, process and group. The focus of the current study is a Chemistry program within one university. The reason for this focused approach stems from Rose and McKinley's (2018) observation that the implementation of the Top Global University Project (TGUP) initiative in Japan is environmental and context-specific. A case study approach enabled me to illustrate certain qualities which were unique to the university, necessary for understanding unique forms of EMI implementation. The methodological framework underpinning the current study is hence positioned as a case study at the institutional and course level, necessitating that case studies offer a 'thick description' (Yin, 2013) of the specific research context. Recognising the context-dependent nature of the case study, it is prudent to establish transferability of the findings to other contextually similar or different EMI programs.

3.4 Research design

The current research adopted the pretest-posttest non-equivalent control group design. According to Bryman (2012), this design evaluates the effect of an independent variable (i.e. EMI) on a dependent variable (i.e. post-test scores), and is often practised by social science researchers due to logistical and ethical reasons. All participants completed the pre-post-test (see section 3.9.2 for instrumentation) at time one and two (see Figure 3 below). The test score differences between the two time points were assessed within the groups to explore the score changes. The post-test scores while adjusting for the pre-test scores were then compared between the groups to examine whether one group demonstrated a larger gain than the other group.

Figure 3: The pretest–posttest non-equivalent control group design



Unlike true-experimental designs, quasi-experimental designs lack the key feature of random assignment. In the study, the researcher did not randomly assign the participants to the two conditions. Instead the individual students were free to select their preferred group. Consistent with Macaro's (2018) observation that "details of the enrolment process … would give crucial information about the research populations" (p. 155), this enrolment process might have affected the demographic of students who studied through EMI. Indeed the EMI professor stressed in the first lecture of the term that the students should aim to improve content knowledge, not English language knowledge, which also discouraged those who were less confident in English to

register for the EMI course. Thus, it was possible that the enrolment process introduced selection bias among the EMI students. The EMI course consisted of more participants attaining higher levels of English proficiency than its JMI counterpart.

Without fully discarding the possibility of confoundedness, Grant and Wall (2009) warn that quasi-experimental research, such as this study, cast some doubt on its internal validity (causal interpretability). That is, the observed post-test score changes between the groups could be attributed to the independent factor (i.e. EMI) and to other causal factors (e.g. the pre-existing differences in student demographic factors). Cook and Campbell (1979) also discuss threats to external validity (generalisability). The results of the current study, due to myriad of context-specific factors (e.g. participants, courses, university and semester), should only be generalised to the current subset of the entire population being represented by the sample researched.

Knapp (2016) outlines various threats to internal and external validity (p. 468). Table 4 below summarises some of the threats to the validity which were relevant to the present study and how the researcher addressed and minimised them through the data collection process. While there was no intention to establish a causal relationship based on the whole population, it is still important to be aware of possible underlying reasons, other than the language of instruction, why the test score gains might have occurred between the groups (West, 1985).

Table 4: Threats to internal and external validity

Threat	Description	Example	Solutions and Considerations
1 History	During the semester, there might have been events other than the Chemistry program that took place that could be the cause of the change in the outcomes measure.	・English for specific purposes (ESP) ・Other disciplinary courses covering similar academic content ・Extra self- or group-study sessions	Observations and interviews were adopted to identify any differences in events between the two groups that could be considered history threats influencing the students' post-test performance.
2 Testing	As the pre-test and post-test were the same test, the participants were familiar with the questions, and therefore the post-test became	・Higher post-test scores at time two because of the opportunity to take the pre-test at Time one.	・The students were not informed that there would be a post-test at the end of the semester.

	easier. The improvements in scores could be due to a practice effect rather than a treatment effect.		・The three month-interval between Time 1 and 2 was considered long enough to mitigate a practice effect.
3 Statistical Regression	As the groups were below the average score in the population of interest, their pre-test average must improve on the post-test regardless of the exposure to the treatment. This threat would also be an issue when the participants were well above the average.	・The participants' baseline content knowledge was either below or above the average of the population mean (e.g. the lowest 10% of the population based on their pre-test scores).	The pre-test was administered at the start of the term. These test scores were compared against each other and not found to be significantly different from each other. Although it was not plausible to compute the average score in the population of interest, the participants shared the same degree of baseline pre-test knowledge between 2018 and 19.
4 Threats to External Validity	Due to non-random assignment of the participants and context-specific factors (participants, university and timing), it was implausible to generalise the results of the present study to the target population.	・The EMI course already being formed at the university attracted more motivated, linguistically competent, academically inclined students than their L1 instruction counterparts.	Detailed information about the participants' backgrounds were gained through the questionnaire and interviews (e.g. English proficiency level, study abroad experience, and past EMI experience). This enabled the researcher to make a narrower generalisation which was more specific to the current sample - the one it was studied.

*1, 2 and 3 are threats to internal validity and 4 is a threat to external validity.

Despite the methodological drawbacks of quasi-experimental designs, these threats to validity can still be minimised by strengthening the research design and accounting for plausible alternative explanations for changes in the outcome measure. Reynolds and West (1987) suggest the use of covariance analysis to reduce the plausibility of alternative explanations. I adopted the students' pre-test scores as a covariate which examined whether the two groups were comparable prior to the study and investigated whether the treatment effects were still detected after adjusting for the covariate. Similarly, Bryman (2012) claims that potential threats in question can also be ruled out so long as researchers "[seek] out further information that will help to discount some of the rival interpretations of a causal link that arise from the absence of a true

experimental design" (Bryman, 2012, P. 50). Accordingly, I adopted interviews and classroom observations to identify everyday classroom behaviour of students and teachers pertaining to "what happens on the ground" (Samoff, 1994, P. 144).

3.5 Settings

This study was conducted in one of the Top Global University Project (TGUP) participant universities in the greater metropolitan area of Tokyo. This university is a small-scale, private and prestigious institution that has a total student body of 3,000 undergraduate and 200 postgraduate students. In total, the university offers 14 undergraduate departments with 31 majors and five graduate programs with more than ten area of concentrations. This institution was chosen for the following reasons. Firstly, the university has been annually funded by the Japanese Ministry of Education since 2014 as part of the TGUP, thus proposing clear figures to increase the number of EMI courses from 16.5 % to 40.0% during the period between 2013 and 2023 (Appendix 1). Secondly, the university values its strong tradition of bilingual and internationally orientated education, offering a number of parallel English and Japanese courses in a range of academic fields. Parallel courses are two courses offered by a single department which share identical or very similar content offered in the different language of instruction, which was the essential criterion for my study. Thirdly, it was identified by Rose and McKinley (2018) as a university that placed an equal emphasis on the admission of April and September entry students, and thus offered potential to explore students from both domestic and overseas high school systems. The university every year maintains a total number of approximately 300 overseas students which amounts to 10% of its total student body, achieving a heterogeneous student body with varying degrees of English language proficiency. The level of students' L2 proficiency was one of the crucial demographic variables explored in this study. Finally, the case university had a long history and stable reputation for providing uniform and consistent EMI provision (Aizawa & Rose, 2019). Based on a case study of another TGUP funded university in Tokyo, McKinley (2018) concluded that EMI provisions can fluctuate dramatically within a single institution, especially within new programs. Brown (2017) also argues that the newly-forming EMI universities are more likely to face a great deal of variability in the delivery and quality of their programs than their TGUP supported counterparts who have longer history of teaching through EMI.

3.6 Program

This study was conducted in the Winter semester of 2018-19 and 2019-20 in an introductory Chemistry course entitled "Concept in Chemistry". This course offers 70 minutes of lectures twice a week for 12 weeks from December to March. It consists of 21 lectures in total. 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 identical between the two courses (one is in English and the other in Japanese) (Appendix 11, 12). Table 5 below shows the enrolment number over the past ten-year reporting period, showing the typical class-size of the program (Mean 48; SD 8.4).

Table 5: Enrolment number

Year	Number	Year	Number	Year	Number
2006	73	**2011**	48	**2015**	42
2008	43	**2012**	39	**2016**	48
2009	48	**2013**	44	**2017**	E 15, J 31
2010	56	**2014**	47	**2018**	E 18, J 31

3.7 Faculty members

Table 6 shows information about the content teachers who participated in the study. During the fieldwork, the content teachers played a pivotal role in collecting the data with the researcher in the following areas:

- Approving the researcher's visit to the research site as a gatekeeper
- Creating and submitting the university's research ethics application
- Recruiting the student participants
- Devising, translating, administering and marking the content test
- Conducting the lectures
- Organising research meetings with the researcher
- Sharing contextual information (e.g. attendance, final grades, GPAs)
- Organising faculty meetings

Table 6: Information of faculty members

Name	MOI	Position	Gender	CEFR	Overseas studying and teaching experience
Ko	EMI	Professor	Male	L1 English	Teaching in Japan over 20 years
Den	JMI	Professor	Male	C1 (high)	Taught in the US over 5 years
Kawa	JMI	Professor	Male	C1 (high)	N.A.

To ensure the anonymity of content teachers, their details (gender, name, position) are modified to disguise their identity (see section 3.13 for ethical considerations).

The current study as an investigation of real-world practices is prone to the influence of various factors. Of these, teacher-related factors are considered instrumental in influencing student learning. Education research exploring different sectors, for example, primary (Wanders, Veen, Dijkstra & Maslowski, 2020), secondary (Egalite & Kisida, 2018) and higher education (Alcott, 2017) all underpin that teachers' pedagogical ability (e.g. teaching methods, pedagogical styles and teacher engagement) accounts for a large proportion of variance in student academic success. In particular, Chalmers's (2019) recent Oxford University Press report offers various types of linguistic support strategies teachers can implement to facilitate their multilingual learners' language and academic development at EMI school settings, indicating the significance of EMI teachers' pedagogical competence.

It is no surprise that EMI outcome studies reveal that teaching varies between L1 and EMI programs. For example, Park (2007) observed that an experienced bilingual instructor applied different pedagogical methods depending on teaching EMI or Korean medium courses. Similarly, Dafouz, Camacho and Urquia (2014) reported that their participant lecturers did not reach a consensus on assessment criteria. Researchers generally have no control over the operationalisation of teaching, assessment and curriculum, which in turn increases the influences of teacher factors.

Nevertheless, this study embraces such differences as part of its ecological validity—that is these differences are part of students' real-world educational experiences, thus the data need to be interpreted as part of this context. Rather than viewing them as confoundedness, I adopted classroom observations and teacher interviews to closely monitor fidelity to condition in terms of the pre-determined teaching plans between the EMI and JMI teachers. Additionally, the professors themselves organised faculty

meetings regularly to discuss their teaching practices and reconcile the assessment criteria.

3.8 Participants

3.8.1 Sampling

As the data collection of the study involved the selection of small samples from one particular university course, it was not the central aim to make inferences about the population of the students studying academic subjects through EMI or JMI at the university (Lavrakas, 2008). Instead it sought to adopt purposeful sampling to achieve the selection of information-rich cases which were "prespecified up front…. according to a preconceived, initial set of criteria" (Mills, Durepos & Wiebe, 2010, p. 838) rather than random sampling, such as stratified or cluster samples (Flick, 2009). The case selection in the study was based on maximum variation sampling (Patton, 2002) at multiple levels (course level and student level) exploring more than one unit of analysis.

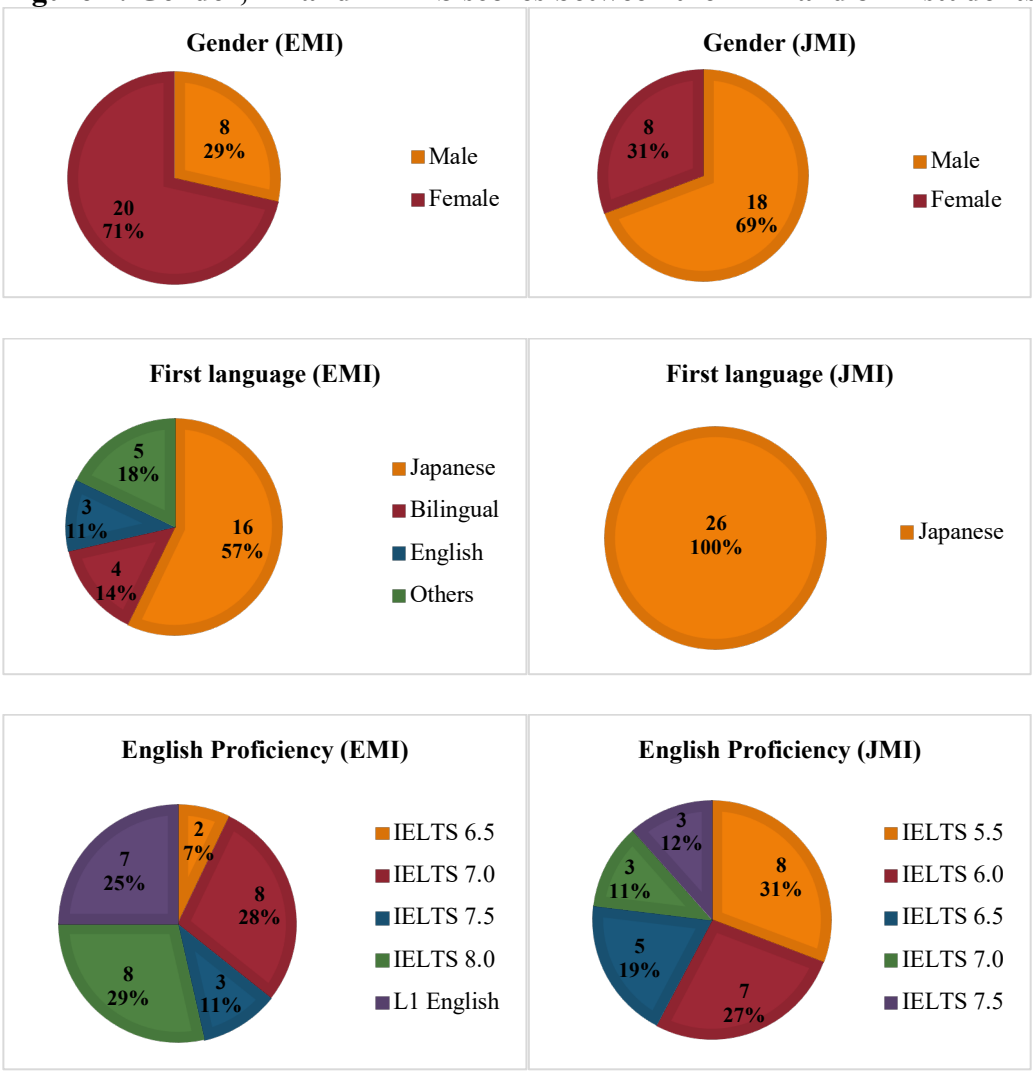
3.8.2 Course level

The first stage of case selection (course level) involved choosing EMI and JMI courses that were comparable in terms of students' baseline content knowledge. The introductory Chemistry program was chosen as it was available in either of two languages of instruction and consisted of homogeneous groups attaining the uniform level of prior knowledge. The JMI and EMI content teachers assisted in the distribution of the participant information sheet (Appendix 16). During the first lecture of the term, all students enrolled in the EMI (n=28) and JMI (n=31) courses were asked to express their willingness to participate in the study. Based on initial review of the consent form responses, I had to exclude five JMI students (who withdrew from the study) and thus arrived at the grand total of 28 EMI and 26 JMI students, which represented 100% and 84% participation rates, respectively. The pie charts (Figure 4 below) summarise the participants' gender, L1 and IELTS scores according to the mediums of instruction.

It should be clarified that the EMI and JMI samples included one and two non-Chemistry major second-year students, respectively. These students however were indistinguishable from their first year counterparts in terms of their pre-test scores and

the self-rated prior-knowledge reported on the background questionnaire, and thus remained in the study.

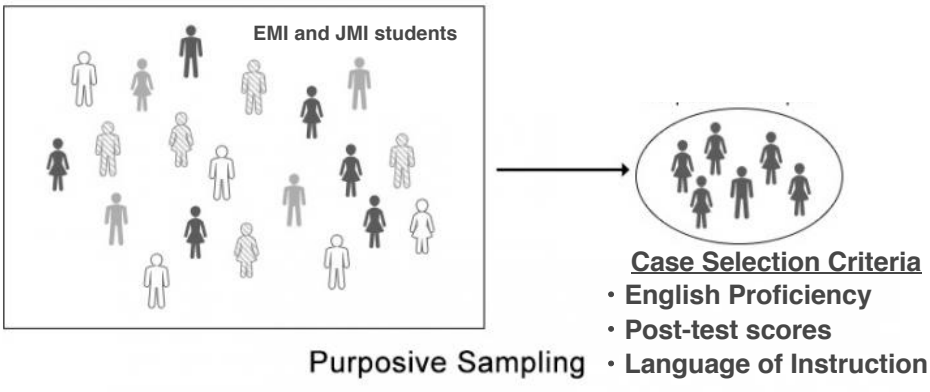
Figure 4: Gender, L1 and IELTS scores between the EMI and JMI students



3.8.3 Student level

The next stage of case selection (student level) employed a maximal variation sampling strategy (Blaikie, 2000) which "only investigates a few cases but those which are as different as possible to disclose the range of variation in the field" (Flick, 2009, p123). Participants are included from a range of cases in order to correspond to "maximum variation and differentiation" (Patton, 2002, p. 243). To recruit interview participants, I selected cases that satisfied characteristics considered crucial to be represented in the sample (see Figure 5).

Figure 5: Purposive sampling of student participants



I purposively selected students as individual cases that would offer the broadest picture of content learning experience in terms of L2 proficiency, content knowledge and language of instruction. Students registered their email addresses on the background questionnaire. I contacted those who were interested to arrange interviews. Table 7 shows the selected cases of 13 EMI students representing a range of content knowledge (measured by the post-test scores) and English proficiency levels (measured by IELTS scores). Table 7 also includes the selected cases of four JMI students who studied through their L1 on the basis of the levels of their content knowledge, enabling comparisons to be made between and within cases.

Table 7: 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)	Hikaru (Post-test score 5)*	Rina (Post-test score 7.5)*	
	IELTS 7.0 (n=4)	Kiley (Post-test score 4) Marii (Post-test score 1)*	Ayaka (Post-test score 8.5)* Minori (Post-test score 9.5)*	
	IELTS 7.5 (n=1)			Tikamon (Post-test score 12.5)
	IELTS 8.0 (n=3)		Charlotta (Post-test score 9.5)	Miyazaki (Post-test score 15)* Lisa (Post-test score 12.5)
	L1 English (n=3)		Maxwell (Post-test score 8) Saika (Post-test score 7.5)	Emma (Post-test score 12.5)
JMI (n=4)	N.A. (n=4)	Nao (Post-test score 3)	Shizuka (Post-test score 8)	Oda (Post-test score 15)* Kyoka (Post-test score 16)*

* Those who participated in both the first and second round of interviews

** Low (0-40%), Medium (40-70%) and High (70-100%) adopted from the university's grading criteria

In accordance with relevant studies (e.g. Lin & Morrison, 2010) which underpin associations between students' L2 proficiency and success in EMI learning, it was speculated that the participants who achieved higher levels of English proficiency would perceive their learning experience less taxing than their counterparts with lower English levels. Previous research (e.g. Turhan & Kırkgöz, 2018) also indicates relationships between students' L2 proficiency and other variables (such as previous experience learning through EMI, challenges and motivation). Evans and Morrison (2011) found that university students from English based school systems could overcome their challenges more effortlessly than their counterparts from L1 medium schools. Accordingly, this study selected 17 interviewees representing different cases to explore a range of variables that were shown to influence their content learning.

3.8.4 Combining two sets of EMI cohorts

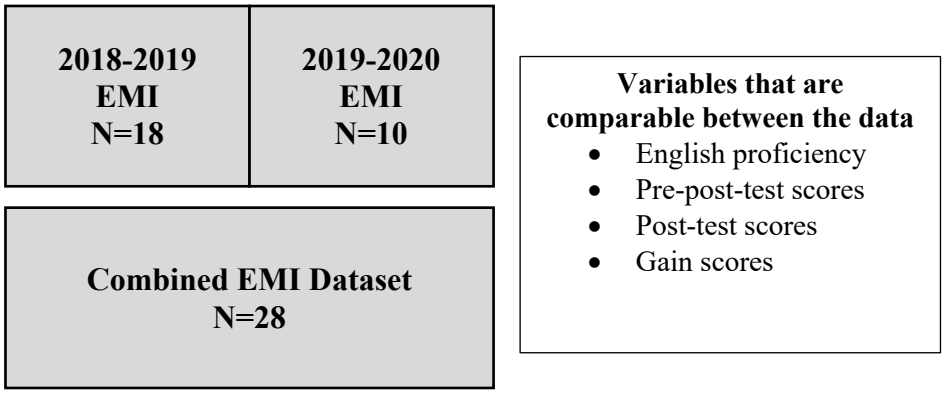
In the current study, the EMI samples were collected over the course of two consecutive academic years. A method for combining separate data sources to achieve larger datasets has been increasingly employed in social science research (e.g. Citro, 2014; Lohr & Raghunathan, 2017). Elliot, Raghunathan and Schenker (2020) note that "larger sample sizes and thus greater information and precision can be obtained by combining data from two or more independent surveys"; indeed, relying purely on a single survey often leads to research suffering from its inadequate responses. Classroom-based case research with small sample sizes such as this study can see particular benefit from data conversion. The increased number of responses ensures that the researcher can ascertain clear patterns in the results generated over a period of time, in turn increasing the reliability of the research instruments. However, others (e.g. Schenker & Raghunathan, 2007) raise concerns that data involving separate sets of participants over several data-collection points (as opposed to a single set of datasets) may introduce additional sampling errors and increased nonresponse biases (see Brummet, Masterton & Smith, 2014 for further discussion of sampling issues). Renssen and Nieuwenbroek (1997) recommend that two sets of data should only be calibrated as a single source when obtaining a set of variables which are comparable across the datasets, measuring the same quantities. Roberts and Binder (2009) outline variables which should be comparable across separate datasets as follows:

- Types of participants (demographic variables)
- Sources of the participants' information (research instruments)

- Research contexts (contextual variables)
- Sampling designs (data collection procedure)

Accordingly, the current study conducted analyses of a Mann-Whitney U test and an independent samples t-test to carefully examine the comparability of the two data sources. The results (see Appendix 3) showed that the participants in the two datasets were not significantly different in terms of the crucial variables outlined in Figure 6 (i.e. English proficiency, pre-post-test scores and gain scores) that were used to match records from the new and old EMI datasets. Thus, the two sets of data were seen as comparable and merged into a single set to be used for the subsequent analyses. Conversely, the JMI sample only involved data gathered in the first data collection phase as the JMI teachers withdrew from the study in the second academic year.

Figure 6: Combined EMI sample



3.9 Data, instrumentation and procedure

Table 8 below shows the timing of the data collection and sources of data collected in the study. Descriptions and justifications of the instruments used are first presented, which are then followed by explanations of the data collection procedure.

Table 8: Sources of data in relation to time and quantity

Time	Data	Quantity
TIME1 The start of the semester 6th Dec	Pre-content test	28 EMI; 26 JMI
	Background questionnaire	28 EMI; 26 JMI
BETWEEN TIME1&2 (December to March)	Observations fieldnotes	10 EMI; 10 JMI
	Student interviews	13 EMI; 4 JMI; Total 456 minutes (7.6 hours); Mean 26.8; SD 6.4; Min 14; Max 36

TIME2 **The end of the semester** **5th March**	Post-content test	28 EMI; 26 JMI
	The EMI Challenges Scale	28 EMI
	Teacher interviews	1 EMI; 1 JMI
	Student interviews	6 EMI; 2 JMI; Total 102 minutes (1.7 hours); Mean 12.8; SD 3.2; Min 9; Max 19

3.9.1 Background questionnaire

The background questionnaire was distributed during the first lecture of the semester (Time 1) and obtained the participants' demographic information: L2 proficiency; major (e.g. Chemistry, Biology); study abroad experience; L1; high school types; and prior EMI learning experience (see Appendix 4 for the full list of variables collected). During the final lecture of the semester (Time 2), the students also reported the average time spent in hours of self-study per week (Appendix 4). The course attendance, the final exam scores and the course grades were all recorded by the content teachers and shared with the researcher at the end of semester.

3.9.2 Content test

3.9.2.1 Pretest/posttest

I employed a pre-post-course test to assess the students' content knowledge. Several studies have adopted direct measures of academic success (e.g. test scores), instead of perception-based instruments (e.g. interviews). A seminal study by Graham (1987), for example, discussed the use of student grade point average (GPA) as a potential indication of academic success in EMI university settings. However, academic prediction metrics such as course grades and GPAs are methodologically problematic, mirroring numerous factors, including completion of coursework, assignments and often course attendance (Dafouz et al., 2014). The current study, accordingly, adopts locally developed pre-post-course tests, which are more of a direct proxy for academic achievement, eliminating the large number of uncontrolled factors involved in academic metrics. Although GPA remains widely used as an indicator for academic success (see Yan & Cheng, 2015; Rose, Curle, Aizawa & Thompson, 2019), researchers (e.g. Perkins & Saris, 2001; Yamarik, 2007; Chin 2002; Wagenaar, 2002) recognise the importance of measuring a baseline of student content knowledge at the onset of the course, endorsing a pretest-posttest design. In the same way, Neuman (1989) measures the content knowledge of sociology undergraduates, assessing a gain

in their knowledge from the start to the end of the semester. In a more recent study, Delucchi (2014) recorded students' baseline performance in a social statistics course, noting that "students enter courses with unequal knowledge, skills, and educational experiences" (p. 232). Thus, I decided to adopt pre-tests to establish the students' prior knowledge and post-tests as requisite to measure their content knowledge gains.

3.9.2.2 Home grown tests

For the current study, a pre-post-test was locally devised by the Chemistry professors at the university. Although tests developed locally achieve less powerful statistical validity and reliability than their standardised counterparts (Benjamin & Pashler, 2015), their relevance to the specific learning objectives is crucial for classroom-based research such as this study. Some studies use low-stakes tests that have little impact on students' course grade and require little preparation and revision in advance, whilst others adopt high-stakes tests such as final exams which are part of the regular curriculum that results in higher level of student motivation. Some researchers (occasionally in corroboration with content teachers) (e.g. Park, 2007) create specific tests to assess score gains for their research, whereas others simply adopt standardised tests or measures, such as entrance examinations or validated language tests. Numerous studies assessing university students' learning (e.g. Delucchi, 2014; Park, 2007; Price & Randall, 2008) adopt home-grown tests since most learning objectives at a tertiary level are distinctive to each course. Researchers (Keenan-Takagi, 2000; Wilson & Sloane, 2000; Mery, Newby & Peng, 2011) argue that home grown tests are particularly reliable to measure gains in student learning for the following reasons:

- The ability to measure what is precisely taught in a specific course;
- The flexibility to use a range of test formats (e.g. multiple-choice items) accommodating local needs;
- The increased likelihood that the data are useful for the participant instructors

For the present study, the EMI Chemistry professor devised a content test to ensure high-level representation of the final course exam, together with the regular curriculum and learning objectives. These tests require criteria already developed in the learning objectives (Martell & Calderon, 2005) by using the same rubric (Gerretson & Golson, 2005) devised by the content teachers. The following criteria were shared between the researcher and the professors to devise the test (see Table 9):

Table 9: Criteria and specifics of the content tests

Criteria	Specifics
Topics	The test only covered content taught in the class from the beginning (e.g. Electrons in Atoms) to end of the term (e.g. Organic Chemistry) (see Appendix 11, 12 for the course syllabus).
Representative of the actual EMI and JMI course tests	To represent the regular course assessments and curriculum, the content test was designed based on the final exam and (where appropriate) homework problem sets.
Length, number of items, item types	The test comprises of 13 open ended questions (see Appendix 22 for the content test)
Administration	The students were given 25 minutes at the end of their first and final lecture of the semester to take the content test under the supervision of the professors and the researcher. The test was conducted under the same condition (closed book) as the final course exam (see section 3.9.6 for the detailed explanation): "This is a closed book exam. You may use a calculator and English-Japanese language dictionary, but no other reference materials."
Test equivalence between the JMI and EMI course	I ensured that the content test was equivalent between the JMI and EMI class, conducted in Japanese or English. The test was first created by the EMI teacher and translated by the bilingual JMI teacher into Japanese. I checked the comparability of the JMI and EMI versions before the administration.
Language choice made by students to answer the test	The test was conducted in the chosen language of instruction (English for EMI and Japanese for JMI). The EMI students were however permitted to provide their answers in Japanese if language was being an issue for their performance. However, none of the participants had to resort to the using Japanese.
Scoring procedure	The test consisted of a total of 17 points and was marked by the JMI and EMI professors for their respective course and returned to the researcher. The researchers then recorded the scores using a spreadsheet alongside other student demographic variables.
Marking criteria	As the pre-post-test was created based on the final course test, the regular marking criteria used in the course were adopted by the teachers (see Appendix 10).

3.9.2.3 Reliability and validity

Although it was not the central aim to devise a standardised instrument that could be used in wider research contexts, it was still crucial to assess the reliability and validity of the content test used in the study. Researchers warn that a self-devised instrument that has not been tested for its reliability and validity may result in a lack of "external credibility for accountability purposes" (e.g. Gratch-Lindauer, 2003, p. 28).

Reliability

The reliability of the content test items was first inspected using Cronbach's alpha coefficient. Reliability assesses the internal consistency of the measurement (Gravetter & Wallnau, 2000). Cronbach alpha coefficient of .70 indicates acceptable internal consistence. The Cronbach's alpha coefficient of the 13 post-test items yielded an α value of α=.71, which was above the cut-off line (Blunch, 2013).

Although all items were worthy of retention, resulting in a decrease in the alpha value if removed, the only exception to this was question #6 shown below. The deletion of this item from the analyses would increase the alpha value to α=.75 (i.e. a very small change in α value according to Blunch, 2013), but the EMI content teacher suggested that this item should be retained. The correlation coefficient of this item (r=.345) was still more than r=.30, indicating that it belonged to the same construct. This decision was also supported from the content expert standpoint that question #6 was an instrumental question in linking the previous #5 and subsequent #6 questions, assessing the understanding of the same concept, as shown below.

#5. Silicon reacts with fluorine. Write the chemical equation for the reaction.
#6. Draw the Lewis structure of the product in Q. 5 showing all valence electrons.
#7. Draw the molecule in Q. 6 in 3 dimensions (apply valence shell electron pair repulsion (VSEPR) theory).

Validity

To examine the validity of the content test, construct validity was assessed. Construct validity refers to the degree to which an instrument measures the construct of learning outcomes that it is intended to measure (Beanland & Schneider, 1999). Based on Mery et al. (2011), construct validity was tested through assessing the item-total correlations between the self-devised test (i.e. the content test) and another set of tests measuring the same sets of constructs (i.e. the final course exam). The paired samples correlations (n=28 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 $r = 0.450$, $n = 28$, $p = .016$.

The validity of the test instrument was also examined based on the expert ratings of each of the test items. The JMI and EMI content teachers independently compared the content test against the final exam and problem sets based on the course's key learning objectives (Appendix 11, 12) to rate how closely they correlated. Consistent with the acceptable cut-off level of 70% for an interrater agreement (Cameron, Wise & Lottridge, 2007), the content validity in this study was substantial with interrater agreement on 12 of the 13 items (92.3%), which added evidence to its construct validity.

3.9.2.4 Piloting the content test

The locally developed content test was piloted in November 2018 prior to the main investigation. The test was administered to natural science major students at the university who had previously taken the course (n=11) to simulate the post-test condition. The Chemistry professors asked students from previous academic years to participate in the pilot study. The pilot study was conducted to inspect the following aspects of the content test (Earl & Torrance, 2000; Price & Randall, 2008).

- Clear wordings and sentence structures
- No double-barrelled items used in the instructions
- The appropriate test level for the target demographic
- No priming and practice effects

A short open-ended questionnaire (Appendix 9) and brief interviews were conducted as part of the pilot study to elicit general feedback from the students. Accordingly, the wording for the instruction sentences, where the students expressed confusion, were revised, although the questions themselves were unchanged. It was also pointed out that the students who had no prior disciplinary knowledge would not require the full 25 minutes, which helped to modify the administration process of the pre-test in the main study (see section 3.9.6 for administration of the test). According to Trochim, Donnelly and Arora (2015), researchers using a pre-post-test design, should obtain an appropriate test level for a chosen sample to minimise a regression threat and accurately observe the expected score changes (discussed further in section 3.4 above). In the study, to achieve the appropriate test level, I inspected the questions that had been shown to be too easy or too difficult by the pilot study participants. Accordingly, two out of 13 questions were found to be too difficult as none of the students could answer them correctly. The content teachers, nevertheless, confirmed

that these questions were made intentionally more difficult than others to discriminate between students attaining varying degrees of knowledge, thus these items were retained following their advice (i.e. the item discrimination index, see Salkind, 2010), and safe-guarded against a ceiling effect.

Finally, I had also planned to inspect priming and practice effects; that is, score differences between the pre-test and post-test were due to participants' previous exposure to the same test. Following Weiler, Moyle, Torland, Wolf and Witsel (2014), it was planned to compare the mean post-test scores between those who completed the pre-test against those who only completed the post-test. Those statistically significant items could be subject to priming. However, the insufficient number of participants recruited for the pilot study did not allow this statistical analysis to be conducted.

3.9.3 Semi-structured Interview

Direct evidence (e.g. test scores) offers measurable student learning outcomes, but it is not conducive to *why* questions, such as 'why do some students learn more than others?' (Ross, Young & Sturts, 2012; Martell & Calderon, 2005). Indirect evidence of student self-concepts (e.g. interviews), on the other hand, can provide more in-depth information about their content learning processes and trajectories. This study adopted indirect evidence as an essential outcome measure to investigate why some students were able to make far greater progress than others even when starting with similar pre-test scores. I adopted interviews to draw out their perceived learning experience while identifying underlying factors affecting their content learning. The semi-structured form adhered to the topical focus, using a set of pre-determined questions. Yet I was still able to generate context-specific knowledge through conversations by adjusting how questions were posed in response to the participants' answers (Gray, 2018).

3.9.3.1 Student interviews

The interview schedule (Appendix 6, 7) used in this study was adopted from Evans and Morrison (2011) who investigated EMI challenges and Park (2007) who examined the effects of EMI on student content learning. It centres on four aspects of student content learning (i.e. content learning differences between EMI and L1 instruction, English language development through EMI, language-related challenges, and motivations towards EMI). I had used this same interview schedule in the

following EMI studies I was involved in (e.g. *success* for Rose, Curle, Aizawa & Thompson, 2019; *self-efficacy* for Thompson, Aizawa, Curle & Rose, 2019; and *challenge* for Aizawa, Rose, Thompson & Curle, 2020). In these studies, the data collection involved semi-structured interviews with undergraduates enrolled in an EMI business program in another TGUP university in Tokyo. These studies demonstrated the importance of interviews in combination with questionnaires and test scores for proposing underlying factors affecting student academic success (e.g. prior disciplinary knowledge, and teacher-student relationships).

These studies also enabled me to pilot the instrument prior to my main PhD study. Consequently, I underlined key phrases in each question, so that a cursory glance during the interviews reminded me of the entire questions without reading them word by word. This ensured a smooth conversation flow. I also reorganised the order of the questions to achieve more coherent transitions between the different sections. The section about the students' demographic information (e.g. study abroad experience, L1 and the number of past and present EMI courses) was also removed as I was already able to obtain the same information from the background questionnaire. Furthermore, the protocol was also tested through a small-scale pilot study involving three Chemistry major second-year students at the research site university. The pilot interview data allowed the researcher to develop an inductive list of themes to conduct an initial phase of coding of the interview data for the subsequent qualitative analysis. For this reason, the pilot study recruited the participants with the similar disciplinary and demographic backgrounds as those recruited in the main study.

3.9.3.2 Teacher interviews

Besides students, interviews were also conducted with the JMI and EMI content teachers. The interview schedule (Appendix 8) centres on teachers' views on: teaching practices; content and language learning; and challenges faced by the students. The interviews had two purposes with regard to the research questions. Firstly, they sought to offer contextual information regarding the differences in teaching practices between the JMI and EMI courses. Although on the face of it, both courses shared identical learning objectives, it was crucial to gain the professors' situated knowledge and views affecting their teaching practices. The interviews offered a better understanding of how teaching was operationalised in practice in

terms of language use, classroom discussions, pedagogical styles, pop quizzes, student expectations, faculty meetings, and teaching progress.

Secondly, the teacher interviews substantiated the findings of the student interviews by adding another dimension to the professors' views on content learning. I especially focused on their perceptions on how the magnitudes of the students' success and challenges were affected by changing the language of instruction. This enabled me to verify the data collected from other data sources and gain a more accurate portrayal of students' EMI learning experience observed by the professors. The triangulation of teacher—student interviews is instrumental in measuring student learning outcomes (Carter, Bryant-Lukosius, Dicenso, Blythe & Neville, 2014).

3.9.4 The EMI Challenges Scale

A 45-item questionnaire (Appendix 5) referred to in this study as the English Medium Instruction (EMI) Challenges Scale was adapted from Evans and Morrison (2011) regarding the students' perceptions of difficulty for tasks in the four skills in academic English: writing (15 items), speaking (10 items), reading (10 items) and listening (10 items). The participants indicated the degree of ease via a five-point Likert scale ranging from 1 (very difficult) to 5 (very easy). This questionnaire has been widely adopted to investigate students' EMI challenges in global settings (e.g. Rose, McKinley, Xu & Zhou, 2020 in China; Aizawa, Rose, Thompson & Curle, 2020 in Japan; Kasamak, Sahan & Rose, 2020 in Turkey) and tested for its constructs confirming that the assumed factors were transferable across different research contexts. Recently, a large-scale study by Aizawa et al. (2020) (n=264) used exploratory and confirmatory factor analysis to suggest that all the underlying factors in their questionnaire loaded on to the expected factors based on a priori assumptions of Evan's and Morrison's (2011) dimensions, confirming that each of the four constructs can be analysed independently.

In the current study, the questionnaire was localised based on the specific nature of the Chemistry course. The localisation process involved removing its entire writing component and two speaking items regarding "oral presentations" as the current introductory level course had no writing or presentation components. The adapted questionnaire was then tested for internal reliability via Cronbach's alpha. The alpha value showed the questionnaire to reach acceptable reliability, α=.812 (reading);

α=.828 (speaking); and α=.801 (listening), all exceeding .70 coefficient of appropriate reliability (Dörnyei & Csizér, 2012). Hence, the localised version of the EMI Challenges Scale was administered, comprising of speaking (8 items), reading (10 items) and listening (10 items).

3.9.5 Unstructured observation fieldnotes

As most research into the effectiveness of EMI has adopted a quasi-experimental approach, reflecting on real-world practices, observations are a crucial means to capture the natural climate of instructional activities affecting the process of "lived through" learning experiences (Curdt-Christiansen, 2020, p. 346). For instance, Zhao and Dixon (2017) adopted observations to "present a richer picture of EMI classroom practices" and "analyse the discourse and interactions within the classrooms themselves" (p. 208). In a study investigating the effects of EMI in Hong Kong, Din, Tsang and Cheung (2003) used observations to "identify the causes that may have led to the differential science achievement of the EMI and CMI student". They argued that "observation of science lessons can help us understand the realism of classroom contexts in EMI and CMI and compare the classrooms in relation to the different medium of instruction used." (p. 326). Zhiguo (2018) also supplemented semi-structured interviews with observations to highlight discrepancies between participants' accounts and behaviours to build a more detailed and accurate understanding of the classroom climate.

As such, I adopted observations as a medium to elucidate as many underlying factors as possible (e.g. instructors' pedagogical styles and the types of activities introduced) that probed into the causes of potentially differential learning outcomes between the two courses. Since I did not intend to quantitatively measure certain aspects of the classroom characteristics according to a pre-determined coding scheme (e.g. the proportion of time allocated to certain learning activities or student–teacher interactions), unstructured observation was deemed more suitable. The unstructured form enabled me to flexibly develop as many detailed memos as possible to exploratorily illustrate how teaching and learning differed between the two mediums of instruction and elicit contextual information using fieldnotes. I was also able to distance myself from preconceived notions of what would be different between the two courses without being restrained by protocols. I observed ten lectures each for the JMI and EMI courses in which new learning objectives were introduced. Table 10

below outlines the timing of my observation schedule in relation to each learning objective obtained from the professors. As illustrated in the images provided below, I sat at the back of the lecture room, taking observation notes on what I saw, listened, felt and experienced in relation to the characteristics of the two courses.

Images: EMI and JMI classroom observations



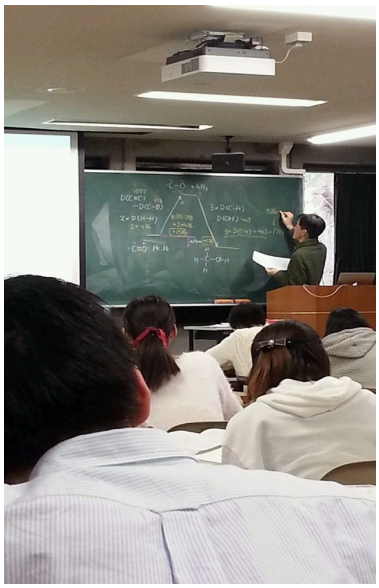


Table 10: Observation schedule

No	Timing	Learning Objectives
1	6th Dec 2018	Introduction
2	11th Dec 2018	Electrons in Atoms
3	18th Dec 2018	Periodic Table and Atomic Properties
4	8th January 2019	Chemical Bonding I: Lewis theory, covalent bonding
5	15th January 2019	Chemical Bonding I: polarity, resonance, molecular shapes, bond order length and energy
6	24th January 2019	Chemical Bonding II: valence bond theory, hybridisation, multiple bonds
7	5th February 2019	Chemical Bonding II: molecular orbital theory, delocalisation, bonding in metals
8	19th February 2019	Acids and Bases: Arrhenius, Bronsted, PH
9	26th February 2019	Strong/weak acids and bases ions, Lewis Structure, organic Chemistry
10	5th March 2019	Final Exam

3.9.6 Research procedure

The current study was conducted over a period of one academic semester from December to March, consisting of 21 lectures in total. Regarding administering the content test, the pre-test and post-test were administered under the supervision of the EMI and JMI professors during the first (Time 1) and last (Time 2) lecture of the semester, respectively. The students were offered up to 25 minutes to complete the tests, but were allowed to leave the room early in 'Time 1' given that some of them had no prior content knowledge. The pre-tests were conducted in the first introductory lecture where the professor explained the course objectives. At the end of this first lecture, I briefly explained the aim of the research, responded to the students' questions and distributed the pre-tests. During 'Time 2', the post-test was administered as part of the final course exam. The students were given the full 25 minutes to complete the tests. No students were allowed to leave the room early in 'Time 2'. The same data collection procedure was applied in both the EMI and JMI courses.

Regarding the interview procedure, the students shared their email addresses with me on the background questionnaire at 'Time 1'. I contacted those who were interested to arrange an interview during the first three weeks of the semester. The average length of the interviews was 26.8 minutes (Mean 26.8; SD 6.4; Min 14; Max 36). The interviews were held after the chemistry class in the same classroom. I used Japanese with the L1 Japanese (n=11), bilingual (n=2) and international (n=2) students and used English with the international (n=1) and L1 English (n=1) students. All interviews were recorded with an audio recorder with their written and oral consent. At the end of the semester ('Time 2'), the EMI (n=6) and JMI (n=2) students agreed to participate in the second round of interviews. The 'Time 2' interviews aimed to elucidate: 1) the students' overall learning experience throughout the semester; 2) their perceived post-test performance; 3) how their challenges manifested at 'Time 1' persisted and dissipated during the semester; and 4) their perceived benefits of learning content through EMI and JMI. The average length of the second round of interviews was 12.8 minutes (Mean 12.8; SD 3.2; Min 9; Max 19). The interviews with the EMI and JMI professors lasted approximately 30 minutes and were conducted in English and Japanese, respectively.

Regarding the administration of the EMI Challenges Scale, given that the questionnaire was designed to explore student challenges distinctive to EMI content learning, it was distributed only to the EMI students at 'Time 2', right after the administration of the post-test. This timing allowed the participants to reflect on their semester's worth of learning and challenges when completing the questionnaire.

In summary, during the first lecture of the semester (Time 1), the researcher offered a brief description of the study and distributed the background questionnaire along with the pre-test. During the semester, the lecture observations and student interviews were regularly conducted (see Table 10 for sources of data in relation to time). During the final lecture of the semester (Time 2), the students completed the EMI Challenges Scale, final course exam, and post-test (see section 3.9.2). In the meantime, the researcher also conducted the student and teacher interviews and collected additional student demographic information from the content teachers (e.g. course grades and attendance)

3.10 Data analysis

In this section, data analysis undertaken in this study will be discussed. With a mixed methods approach, the quantitative and qualitative data gathered were analysed independently according to each of the research questions. First, the data entry and cleaning process performed prior to quantitative analysis is described, followed by an explanation of normality and outlier inspection. I will then discuss descriptive and inferential statistics performed on variables (e.g. the content test scores, IELTS scores, and challenges scores). Regarding qualitative analysis, coding process and methods of data analysis will be explained. The qualitative strand offered a narrative dimension to the numerical data, adopting a thematic analysis of the data including student interviews, teacher interviews and observation fieldnotes.

3.11 Quantitative data analysis

3.11.1 Preparing data for statistical analysis

The data preparation and cleaning process involved organising and recoding the different types of data from multiple sources into one dataset on SPSS: the pre-test scores; post-test scores; gains scores; challenges scores; and student demographic variables (e.g. gender, IELTS scores, attendance, first language, final course exam scores, GPA, and study abroad experience). I first coded the continuous data (e.g. test

scores) into the spreadsheet in relation to each of the participants. The gain scores were computed by subtracting the pre-test scores from the post-test scores. The IELTS band scores were treated as numerical values measured on a discrete and continuous scale in increments of 0.5 from a minimum score of 1.0 to a maximum score of 9.0. To assign numerical value to all of the proficiency levels in the dataset, the L1 English speakers were recoded as 9.0 to be consistent with the competence defined by Common European Framework of Reference for Languages (CEFR). Next, I assigned numerical values to ordinal and categorical data. The students' mediums of instruction (EMI and JMI) and sex (male and female) were dichotomous variables and coded as 0 and 1, respectively. As for the challenges scores, a composite mean ("a mean of means") was formed for each of the academic English skill areas (e.g. speaking, listening) (Evans & Morrison, 2011, p. 203). Based on these new datasets assembling all the quantitative variables, descriptive and inferential statistics were performed. Table 11 illustrates the data sources and statistical analysis in relation to each of the research questions.

Table 11: Overview of quantitative data analysis

Research questions	Data sources	Data analysis
1.1 Directly measured content knowledge	・Pre-test scores ・Post-test scores ・Gain scores (post-test–pre-test)	・Descriptive statistics ・A Mann-Whitney U Test ・A Wilcoxon Signed Rank Test ・A one-way ANCOVA ・A t-test of gain scores (between-subject)
3.1 Directly measured relationship	・Pre-post-test scores ・IELTS scores ・The EMI Challenges Scale	・Spearman's rank correlation coefficient ・Linear regression analysis ・Multiple regression analysis

3.11.2 Normality, transformation and outliers

Before conducting statistical tests, the data were tested for normality of distribution. Normality of the dependent variables was inspected using numerical and visual methods (Sullivan & Artino, 2013). A z-score of skewness and kurtosis was first computed by dividing the skewness and kurtosis measures by the respective standard error (see Appendix 18 for a z-score table). This procedure reinforced the subjectivity of visual inspection by offering clear cut-off points, and was suitable for small to medium sized samples ($n < 300$) (Field, 2018). In small samples ($n < 50$), such as this study, the critical value for rejecting the null hypothesis is an alpha level 0.05, meaning that z-scores above $|1.96|$ indicate non-normality. Additionally, the formal normality test of Shapiro-Wilk test was applied, which is considerably more powerful

to detect deviations from normality than the Kolmogorov-Smirnov test for the current relatively small sample size (Field, 2018). A non-significant result (sig. value of more than .05) indicates normality (see Appendix 23).

An eyeball test also assessed the normality of the score distribution by inspecting the actual shape of histograms (see Appendix 19 for histograms). A normal distribution curve should be symmetrical with neither positively/ negatively skewed nor kurtotic shapes. Based on these inspections, the pre-test scores for both the EMI and JMI students presented a positively skewed distribution, indicating that a large proportion of the JMI and EMI students' scores clustered at the lower end of the distribution. The post-test and gain scores, on the contrary, appeared to be reasonably normally distributed.

Although transforming data can mathematically modify each participant's data value to meet the assumptions of the parametric tests (Tabachnick & Fidell, 2007; Briggs Baffoe-Djan & Smith, 2020), the pre-test scores measured on a ratio with an absolute zero in this study would have distorted the underlying meaning of the measurement and violated the assumption of transformation. Consequently, given that a violation of the normality assumption restricted the selections of statistical tests used for the subsequent analyses (Field, 2018), I adopted a Wilcoxon Signed Rank test (a nonparametric alternative to the paired sample t-test) to compare the pre-test and post-test scores within the groups (within groups comparison of the pre-post-test scores); a Mann-Whitney U test (a nonparametric alternative to the independent t-test) to compare the pre-test scores between the two groups (between groups comparison of the pre-test scores); and a one-way ANCOVA to compare the post-test scores between the groups (between groups comparison of the post-test scores).

Potential outliers were also detected through transforming the test scores to z-scores. A common threshold for screening for outliers for continuous variables is that they are greater than ±3.29 standard deviations removed from the sample mean (Tabachnick & Fidell, 2007). Accordingly, one outlying value in excess of +4.08 was detected in the dataset. In the subsequent section below where I discuss inclusion and exclusion criteria, the justification of the removal of this outlying value from the sample is explained.

3.11.3 Inclusion and exclusion criteria

As the research essentially examined changes in the students' academic performance, participants scoring over 80% on the pre-test (the equivalent of grade A in relation to the university's exam evaluation criteria) were excluded from the analysis. As a result, one EMI participant scoring 97.1% on the pre-test was removed. The questionnaire data clarified that she had already studied Chemistry in another module prior to the current semester, which had advanced her disciplinary knowledge more than her peers. No other participant scored over 80% at pre-test, and no further cases were excluded. Thus, a total of 53 students participated in the study, comprising of 27 EMI (51%) and 26 JMI (49%) students.

3.11.4 Descriptive and inferential statistics

Descriptive statistics

Descriptive statistics were first calculated for the outcome variables involving the pre-test, post-test and gain scores (Range, Mean, Median, SD) according to the independent variables (EMI=27 and JMI=26). I visually inspected the overall features of the test score data through visual representation of the data, including tables, graphs and histograms. Descriptive statistics also allowed me to explore the distribution of the EMI and JMI students' individual test scores to identify the proportion of high and low performing students according to the different mediums of instruction.

Triangulation of ANCOVA and gain score analysis

Two common methods used for pre-post data are analysis of covariance (ANCOVA) and gain score analysis defined as post-test minus pre-test, both of which were used in the present study. In this section, I delineate the differences between the two approaches in the analysis of the test scores. Knapp and Schafer (2009) argue that the two approaches are mathematically inequivalent and yield different interpretations to research questions. Gain score analysis addresses "what is the effect of the treatment on the change from pretest to posttest?" while ANCOVA answers "what is the effect of the treatment on the posttest that is not predictable from the pretest?" (p. 2).

ANCOVA examined differences in covariate-adjusted post-test scores between the two groups and test a hypothesis (H_0) that the groups sharing the same baseline scores from the pre-test improved at the same rates. Adjusting for the pre-test scores helped to detect the size of the effect of the medium of instruction and ensure that the post-

test differences were not linked to the pre-existing score differences. ANCOVA was also statistically more powerful than gain score analysis and suitable for a small sample size in a quasi-experimental study (Stevens, 2002).

In contrast, gain score analysis, regardless of their prior test scores, essentially compared any gains between the two groups to test a hypothesis (H_0) that the groups improved the scores at the same rates. The analysis addresses the "question of whether the two groups differ in terms of their [group] mean change." (Fitzmaurice, Laird & Ware, 2004, p. 124). Hence, ANCOVA expected the groups to have the same level of the prior content knowledge, whereas gain score analysis did not presume the same assumption about the baseline content knowledge. Accordingly, as it was both the post-test and gain score means across the groups that were of interest, ANCOVA was first administered to test the potential differences in the adjusted mean post-test scores, which was then substantiated by gain score analysis (Appendix 21) to triangulate the data analysis to explore the differences in gains, growth, and changes.

Spearman's rank correlation coefficient and regression analyses

To address the third research question, Spearman's correlation was first adopted to explore the relationship between the EMI students' prior content knowledge (pre-test scores), their L2 proficiency (IELTS scores), their perceptions of challenge towards EMI learning (i.e. perceived challenges in English-mediated tasks), and their overall academic performance (post-test scores). Subsequently, I adopted three separate simple linear regression analyses to explore the extent to which pre-test scores, EMI challenges, and proficiency as independent variables explained the variance in the outcome variable (i.e. post-test scores). A multiple regression was also performed to explore the extent to which proficiency and prior content knowledge as independent variables predicted academic success in EMI. As Spearman's correlation revealed a strong correlation coefficient between EMI challenges and proficiency, the multiple regression model only included proficiency and prior content knowledge as a predictor variable. Adopting the explanatory design, I first investigate the association between the variables through quantitative analysis (RQ3.1) and then through qualitative analysis (interviews with the EMI students) (RQ3.2). The next section will explain the approach adopted to analyse the qualitative data.

3.12 Qualitative data analysis

To code the volume of qualitative data including the interviews and fieldnotes, my study applied qualitative content analysis (Mayring, 2019), which is a common method of qualitative data analysis and adopts the same analytical framework as thematic analysis (Braun & Clarke, 2006) and qualitative text analysis (Kuckartz & McWhertor, 2014). These methods have been employed in numerous qualitative EMI studies (e.g. Rose & McKinley, 2018; Aizawa & Rose, 2019; Xie & Curle, 2020) are described as a data-reductive analysis (Zhang & Wildemuth, 2009; Patton, 2002). Researchers can identify recurring patterns of qualitative data to uncover main and sub themes within a specific analytic method instead of merely describing the data (Ryan & Bernard, 2000). The analysis also involves an exploration of themes identified across multiple cases. In my study, a cross-case synthesis was an effective method to contrast the students' learning experience between multiple cases (e.g. EMI/JMI and IELTS levels). I now turn to explain the process of qualitative data analysis, which is divided into the following six steps (adopted from Mayring, 2019; Kuckartz & McWhertor, 2014).

3.12.1 Qualitative content analysis

Step 1: Initial work with the text

As the qualitative analysis requires data to be transformed into written text, the audio-recordings of the interviews were transcribed for detailed analysis. The qualitative data used in the study included the following:

- EMI and JMI student pre-post interview transcripts (RQ1, 2, 3)
- EMI and JMI teacher interview transcripts (RQ 2)
- EMI and JMI lecture observation fieldnotes (RQ 2)

Listening to the full recordings of the interviews numerous times during the transcribing process familiarised the researcher with the data (Braun & Clarke, 2006). To ensure the transfer of the same meaning, tone and nuance as that of the original language, the interview data were transcribed and analysed in Japanese through NVivo (a computer-assisted qualitative analysis software). Excerpts were later translated into English for data presentation. Regarding the fieldnotes, the first stage of the analysis was undertaken manually by reading and annotating the data with comments, identifying reoccurring themes in different colours, inserting memos in

margins, and cutting and sorting key quotes and statements based on the identified themes. This initial step was particularly important because I could highlight and underline the recurring themes and familiarise myself with the overview of the data.

Step 2: Developing a start list of themes (coding scheme)

The next analytical stage after the initial work with data was code development. The coding procedure employed in the present study entailed both deductive coding – directed by existing literature and inductive coding – stemming from the empirical data, which is an approach referred to as the abductive approach by McKinley (2020). Kuckartz and McWhertor (2014) also call for a hybrid approach, noting that "in most cases, a multi-stage process of categorizing and coding is used; completely inductively or completely deductively are rarely found in research practice" (p. 6).

To address research question one and three (challenges and success), a start list of themes was first constructed deductively based on previous EMI research on academic challenges (see for example Macaro, 2020; Bradford, 2016; Evans & Morison, 2011; Flowerdew & Miller, 1992; Ekoç, 2020) and success in EMI (see for example Galloway et al., 2017; Kym & Kym, 2014; Rose et al., 2019; Xie & Curle, 2019) and subsequently developed inductively from the pilot interview data (Zhang & Wildemuth, 2009) (see Appendix 13 for a start list of themes). Similarly, themes for research question two (implementation) will also be constructed abductively resembling the same code development process, through existing literature and the empirical data (i.e. the interviews and fieldnotes).

Step 3: First coding

The researcher then used qualitative data analysis software (NVivo 12) to begin coding the volume of the interview data collected for the main investigation, which helped to organise, code and retrieve qualitative data in an efficient manner. Sentences, phrases and paragraphs of the text data were coded according to the initial abductive themes identified through the previous step. The data were coded based on wording with significant meaning rather than frequency of emergence as suggested by Marshall and Rossman (2006) and Weber (1990). Therefore, the analysis entailed using individual themes as the unit of analysis, rather than the numerical units (e.g. counting individual words, sentences and paragraphs).

It is important to clarify that this initial list of themes was constantly checked, combined, collapsed and discarded within the collected data since it was possible for new concepts to constantly emerge through the initial and later cycles of coding (Huberman & Miles, 2002). During the first coding phase, it became clear that the following themes in Figure 7 that had been initially included in the start list of themes (Appendix 13) were removed due to a lack of coverage across the source data

Figure 7: A list of themes discarded during first coding

Challenge: Academic English speaking (presentation, asking questions)
Challenge: Academic English writing (laboratory reports, essays)
Challenge: Language use (reliance on code switching, simplified teacher language)
Challenge: Quality of learning (simplified content, reduced content coverage)
Challenge: Others (inadequate language support scheme, inadequate computer literacy)

As can be seen in this list, the themes related to academic English speaking and writing challenges, for example, only emerged in the pilot interviews involving second-year chemistry students. This is because laboratory reports and oral presentations (i.e. productive language skills) were only part of the evaluation criteria for the advanced level courses rather than the current introductory level program. Similarly, the themes related to the use of L1 in EMI (e.g. code switching) were also only present in the pilot interviews where the participants experienced L1 use in their EMI courses. In contrast, the EMI Chemistry course in the main study was conducted entirely in English. Hence, the start list of themes constructed through the literature and pilot interview data was constantly modified to create broader and more specific themes according to the further analysis of the source data.

Step 4: Developing sub-themes

After first coding, I reviewed all the text passages coded under the same main themes (text retrieval) to construct sub-themes, which were all compiled in NVivo. This process involved exploring the phrases and (to a lesser extent) entire paragraphs which highlighted regularities, patterns and sentences with significant meanings in the data. At this point, I also carefully checked the coded passages several times to avoid 'drifting into an idiosyncratic sense of what the codes mean' (Huberman & Miles,

2002, p. 64). Accordingly, the new sub-themes were all subsumed under each of the main themes, which were then used for the subsequent second coding stage.

Step 5: Second coding

The entire dataset was then carefully coded again for the second time but this time using the main and sub-themes revised during the previous stage. The list of themes still constantly changed and developed at this stage as the data were further analysed. Accordingly, a final list of main themes (see Figure 8 below) was generated at the end of this stage as follows (see Appendix 14 for a final list of main and sub themes):

Figure 8: A final list of main themes for RQ1 and 3

Challenge: Comprehension of the course textbook
Challenge: Comprehension of the course lectures
Challenge: Final Grade
Challenge: Adjustment to the new academic and cultural conventions of the Chemistry course
Success: Improved English knowledge
Success: Improved motivation
Success: Increased access to learning resource
Success: Long-term career advancement

Step 6: Cross-case analysis (Case overviews)

The last phase of data analysis was shifting towards integration and synthesis of the findings around the research questions. The current study draws from "case overview" approach to qualitative data analysis suggested by Kuckartz and McWhertor (2014). This critical step in the analysis process involved "compar[ing] a selection of cases with each other with respect to characteristics that are deemed especially relevant to the research questions" (Kuckartz & McWhertor, 2014 p. 22). While applying different terms, such as "a thematic matrix" (Ritchie & Lewis, 2003, p. 220), "cross-case analysis" (Huberman & Miles, 2002, p. 172) and "between-case analysis" (Mills, Durepos & Wiebe, 2010, p. 971), the other qualitative researchers also support creating case overviews. In cross case analysis, matrices were used in the comparative examination of the identified themes based on the selected cases, such as the medium of instruction and L2 proficiency level in my study. Table 12 summarises the data sources and analysis in accordance with each of the research questions.

Table 12: Overview of qualitative data analysis

Research questions	Data sources	Data analysis
1.2 Perceived content knowledge	・EMI student interviews ・JMI student interviews	・Coded according to challenges and success. Results then compared across cases (JMI and EMI) with references.
2. Implementation	・EMI and JMI student and teacher interviews ・Unstructured observation fieldnotes	・Coded according to implementation. Results then compared across cases (JMI and EMI) with references.
3.2 Perceived relationship	・EMI student interviews	・Coded according to challenges and success. Results then compared across cases (IELTS level and prior knowledge) with references.

3.12.2 Summary

This section has explicated the analytical methods. Quantitative data analysis was first described in terms of data preparation, normality and outlier inspection, followed by an explanation of descriptive and inferential statistics performed in relation to each of the research questions. I then illustrated the steps of qualitative analysis, describing the process of abductive theme development and data presentation. Based on cross-case analysis, my study uncovers patterns between the identified themes according to each case against the full range of the data.

3.13 Ethical considerations

Prior to my 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). I also followed the local ethical regulation arranged by the Japanese university where I was granted to conduct the research. Permission to conduct the research was obtained in accordance with the Chemistry professors before any research activities commenced.

After research clearance was approved by the university, I visited the first lecture of the semester to distribute the participant information sheet (Appendix 16), which stipulated the aim and procedure of the study, details of data protection, anonymity and confidentiality. Since the project measured student academic performance in their ordinary courses, I orally explained the participants how the data would be stored and

treated after my fieldwork to maintain the anonymity and confidentiality of the collected data. To ensure anonymity, I modified the details of the content teachers and students to disguise their identity (e.g. gender, titles, names). The participants were also informed that they could withdraw from the study at any time.

Prior to conducting interviews, all the participants were also asked to sign the consent forms (Appendix 17) and return them to the researcher in person. I explained the purpose of the interviews and informed them of the use of a voice recorder and asked if they were comfortable with the interviews being recorded. Therefore, this study was conducted in accordance with the necessary ethical guidelines abided by the institutional ethics policies in Britain and Japan.

3.14 Reflexivity of my positionality

My own lens, identity, worldview and prior experience shape and construct many aspects of the research process. In this study, I adopt McKinley's (2017) observation of positionality that "I would expose my own subjectivity and explore what influence it might give to my data collection and interpretation" (p. 42). Therefore, rather than pursuing an objective interpretation of my data, I embrace the impact of my subjectivity on my data collection, analysis and interpretation. Consequently, throughout the writing up of my research findings, I remained transparent and truthful about my own positionality.

My own background, prior knowledge and past experience influenced my analysis of the data. In particular, when analysing interview data regarding students' challenges, I had to be mindful of my status as a L2 English speaker as challenges faced by students would be perceived differently according to whether they were L1 or L2 English speakers. My past experience of learning content subjects through English as an undergraduate student also affected how I analysed the data regarding students' challenges. Having studied through an EMI program, I faced many of the challenges expressed by the students in the study. Thus, my subjectivity was pivotal in my research as it provided me with the background knowledge necessary to capture the full extent of how the medium of intrusion affected students' academic performance.

While the study embraced my own positionality as a lens through which I interpreted the data, the triangulation of the data in the analysis phase enabled me to check consistency between what was observed in the lectures and what was reported by students, as well as between what was reported by teachers and students in interviews. Comparing various data sources enabled me to capture a comprehensive picture of the impact of the medium of instruction on students' challenges regarding content learning.

It was also noticeable that my positionality influenced my data collection process. In the field, insider/outsider dynamics were shaped by my prior relationship with the research site university. On one hand, my perceived identity as a former graduate of the university with a wealth of insider knowledge (e.g. names of professors, buildings, student dorms, clubs and societies) created an insider dynamic, which greatly facilitated the data collection procedure. My familiarity with the research site enabled me to establish a rapport with the students and the content teachers during interviews. On the other hand, I also achieved outsider dynamics as I approached and interviewed the participants in the position of an external researcher through the formalities of introducing consent forms and information sheets. Thus, my unique positionality as a former student of the university was part of the data collection of my own research. Although it was my duty to reflect on my positionality as a researcher, I also embraced my own positionality. My experience of learning content subjects through EMI at the university greatly contributed to my data collection and interpretation of data.

3.15 Chapter summary

In summary, this chapter has discussed the methodological considerations in relation to the development of the research framework of the study. The framework of the study draws on mixed-methods approaches which have allowed a much richer exploration of this relatively unexplored are of EMI outcome research; test results nor student experience alone can give researchers a full picture of complexity of student content knowledge acquisition. Since most applied linguists conducting classroom-based research have no control over the operationalisation of teaching and assessment, this chapter has highlighted the importance of collaborations with content experts. I

have explained in detail how the detailed contextual information offered by the content teachers allowed a much richer understanding of EMI implementation. Overall, the chapter has illustrated the following methodological contributions this study aims to make to the previous literature in this field: longitudinal design over the time frame of three months (one semester); directly measured instruments (test scores); content tests developed specifically for the research purpose (pre-post-test); adjustment of the effects of the prior knowledge as a covariate (ANCOVA); triangulation of data sources (test scores with final course grades); a direct comparison group of an L1 instruction group (inclusion of baseline); and finally a mixed-methods design (interviews and observations as follow-up data of test scores).

The following chapter will present research findings in relation to the first research question regarding the effects of the medium of instruction on the students' directly measured (RQ1.1) and self-perceived (RQ1.2) content knowledge.

Chapter 4 - Results: Impact of the medium of instruction on chemistry content learning

4.1 Introduction

The chapter draws on quantitative and qualitative approaches to investigate the positive and negative effects of medium of instruction on students' learning of Chemistry. The quantitative evidence (based on direct measures of content test performance) detected measurable differences in the development of their understanding of Chemistry (research question 1.1). The qualitative evidence (involving student interviews) offered indirect measures of their perceived learning achievements (research question 1.2). This two-pronged, direct and indirect approach helped to unpack the complicated nature of student content learning and how their achievement was affected by their chosen language of instruction.

4.2 Research question 1.1

The first sub-research question presented below examined the extent to which the achievement of directly measured content learning differed between the students enrolled in an English and Japanese mediated introductory Chemistry course. 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.

- RQ1: Is there an impact of medium of instruction on undergraduate students' content learning?
    - RQ 1.1: To what extent does the achievement of content learning of the EMI and JMI students differ? (Directly measured content knowledge)

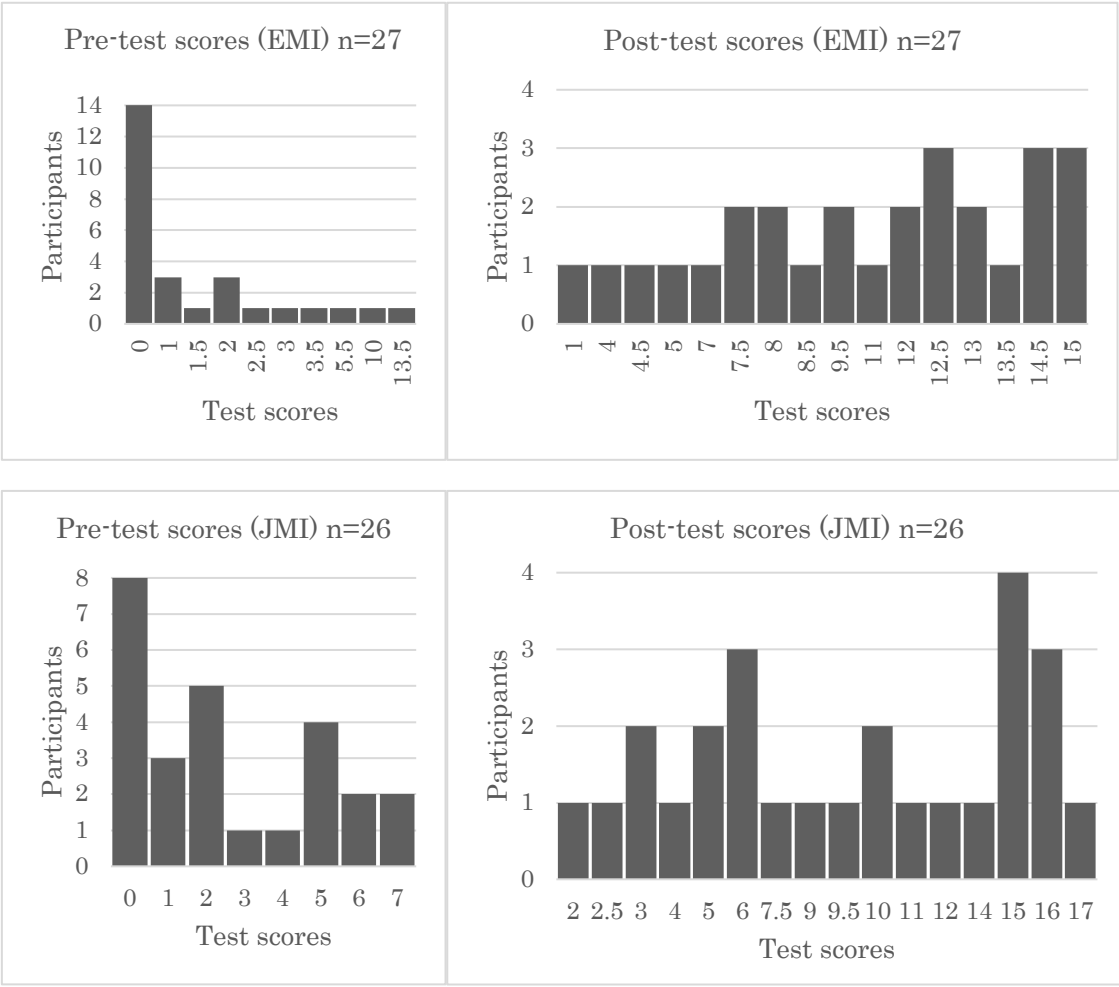
4.3 Within groups comparisons: whole sample

Table 13 below shows descriptive statistics for the pre-test and post-test scores for the whole sample and for the EMI and JMI samples separately. As can be seen in Figure 9 below, 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 statistically examine the students' score changes between Time 1 and 2, I compared their pre-test and post-test scores. Due to the violation of normality (see section 3.11.2), a Wilcoxon Signed Rank test was conducted to reveal a 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. Therefore, the results indicated that both JMI and EMI groups improved their knowledge of Chemistry between the start and end of their course.

Table 13: Descriptive statistics – test scores for the EMI (n=27), JMI (n=26) and 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

Figure 9: Pre-test and post-test scores by EMI (n=27) and JMI (n=26) students



4.4 Between groups comparison: whole sample post-test scores

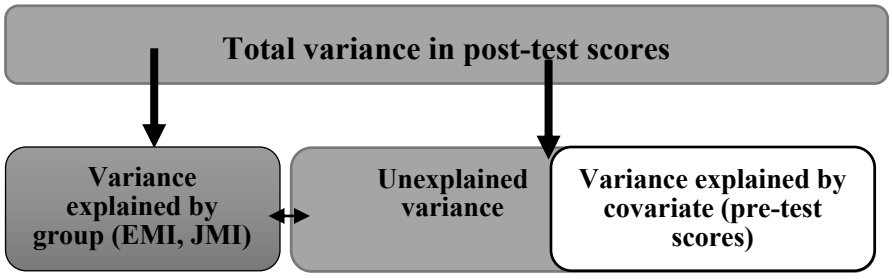
A one-way ANCOVA was conducted to examine whether there was a significant difference in the post-test scores for the EMI and JMI groups while controlling for the pre-existing differences in their content knowledge between the groups (i.e. their pre-rest scores). The independent variable was the medium of instruction, and the dependent variable consisted of scores on the post-test administered at the end of the semester. The pre-test scores were used as the covariate in this analysis.

Before running an ANCOVA, the pre-test scores were first compared to disambiguate any pre-existing differences in their content knowledge between the samples at the start

of the course. Due to the violation of normality (see section 3.11.2), a Mann-Whitney U Test was carried out to reveal 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 results indicated that the JMI and EMI groups were not significantly different in terms of their knowledge of Chemistry at the start of the term.

To conduct an ANCOVA, preliminary checks were carried out to ensure no violation of the assumptions of normality, homogeneity of variances and reliable measurement of the covariate. Additionally, the linear relationship between the dependent variable and covariate was confirmed according to the scatterplot (Appendix 20). I also confirmed that there was no statistically significant interaction between the dependent variable and covariate at an alpha level of .111, safely above the cut-off line, and thus no violation of the homogeneity of regression slopes assumption. Further, independence of the covariate and experimental effect (i.e. medium of instruction) was another prerequisite for an ANCOVA. It was revealed that the level of the students' prior knowledge was independent of the effect of their mediums of instruction (see Figure 10 below). Thus, I adopted the students' mediums of instruction as a covariate to reduce the confounding influence of the group differences.

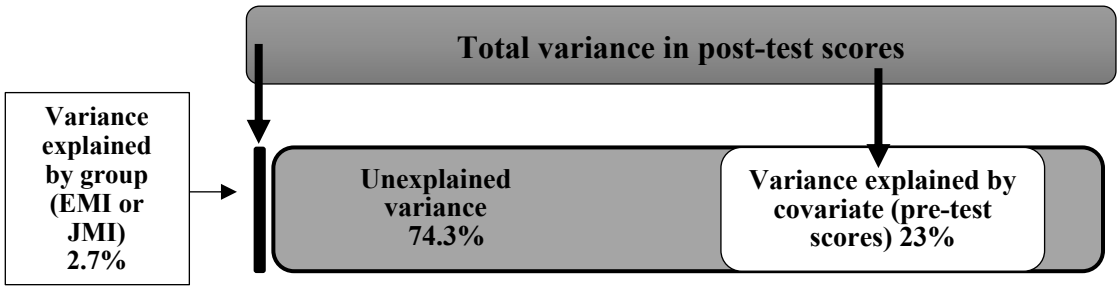
Figure 10: Independence of the covariate and experimental (group) effect



When 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 $\eta 2=.027$. As can be seen in Figure 11 below, only 2.7% of the variance in the dependent variable (post-test scores) was explained by the independent variable (medium of instruction). There was

also a relationship between the pre-test and post-test scores when controlling for the medium of instruction, as indicated by a partial eta squared value of .23. The pre-test scores thus explained 23% of the variance in the dependent variable. Although an ANCOVA reduced within-group unexplained variance by allowing the covariate to explain some of this unexplained variance, the current results showed that the total unexplained variance still amounted to 74.3%. To sum up, the results of the ANCOVA showed that the JMI and EMI groups were not significantly different in terms of their content knowledge at the end of the term when controlling for their prior knowledge.

Figure 11: Explained and unexplained variance



4.5 Between groups comparison: L2-English only sample

The analyses so far focused on how the achievement of content learning differed between the EMI and JMI students, with the condition that the EMI sample consisted of the L1 English (n=7) and L2 English (n=20) students. Seeing that previous studies have supported the role of L2 English proficiency in student academic success, this part of the subsequent analysis only included the L2 English students to investigate whether there would be any differences in content learning between the EMI and JMI cohorts, and if any, how much of an impact the medium of instruction had on their acquisition of content knowledge. After excluding both English L1 students (n=3) and bilingual English/Japanese students (n=4) from the EMI sample (a total of seven participants), the same analyses were carried out with 20 EMI and 26 JMI students, totalling 46 students.

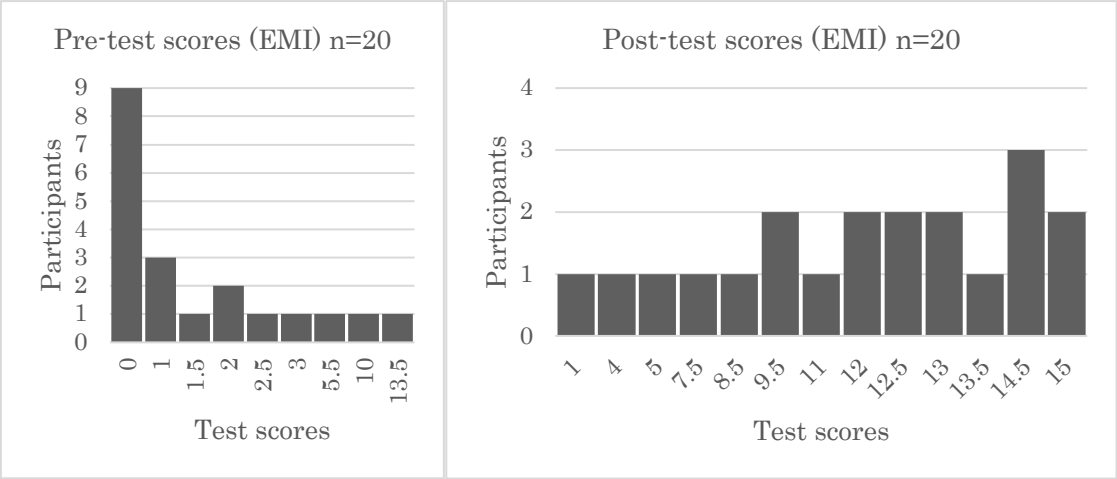
4.6 Within groups comparisons: L2-English only sample

Table 14 below shows descriptive statistics for the pre-test and post-test scores for the whole L2-English only sample and for the EMI and JMI samples separately. As can be seen in Figure 12 below, 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). I also statistically examined their score changes between Time 1 and 2. Due to the violation of normality (see Section 3.11.2), a Wilcoxon Signed Rank test was conducted to reveal 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). Therefore, the results indicated that the L2 English students improved their knowledge of Chemistry between the start and end of the term.

Table 14: Descriptive statistics – test scores for the EMI (n=20), JMI (n=26) and total L2-English only 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

Figure 12: Pre-test and post-test scores by EMI (n=20) students



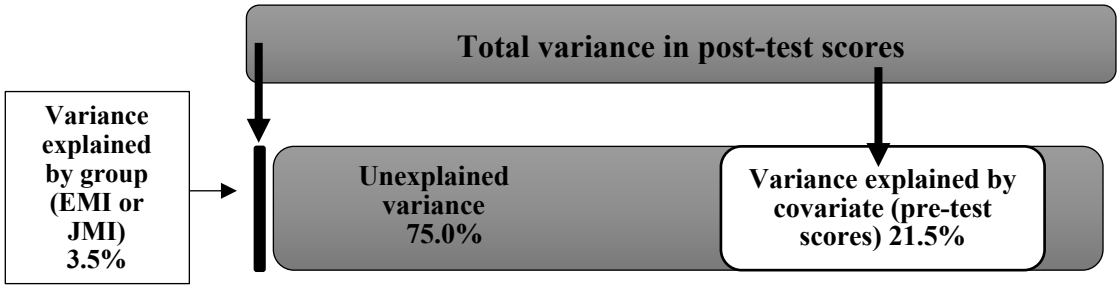
4.7 Between groups comparison: L2-English only sample post-test scores

Prior to conducting an ANCOVA, a Mann-Whitney U test first compared the pre-test scores to disambiguate any pre-existing differences in content knowledge between the groups, indicating no significant difference in the pre-test 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 results indicated that the JMI and L2 English EMI students were not significantly different in terms of their prior knowledge of Chemistry at the beginning of the course.

To employ an ANCOVA, I conducted preliminary checks to confirm that all the assumptions were met (see Appendix 20). A one-way ANCOVA was adopted to test whether there was a significant difference in the mean post-test scores between the EMI (excluding the L1 English students) 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 independent variable (medium of instruction) only explained 3.5% of the variance in the dependent variable (post-test scores). The relationship was also found between the pre-test and post-test scores while controlling for the medium of instruction, as indicated by a partial eta squared value

of .215. The pre-test scores explained 21.5% of the variance in the dependent variable (see Figure 13 below), highlighting that the total unexplained variance still amounted to 75%. Thus, the results showed that the JMI and L2 English EMI students were not significantly different in terms of their content knowledge at the end of the term when controlling for their prior content knowledge.

Figure 13: Explained and unexplained variance

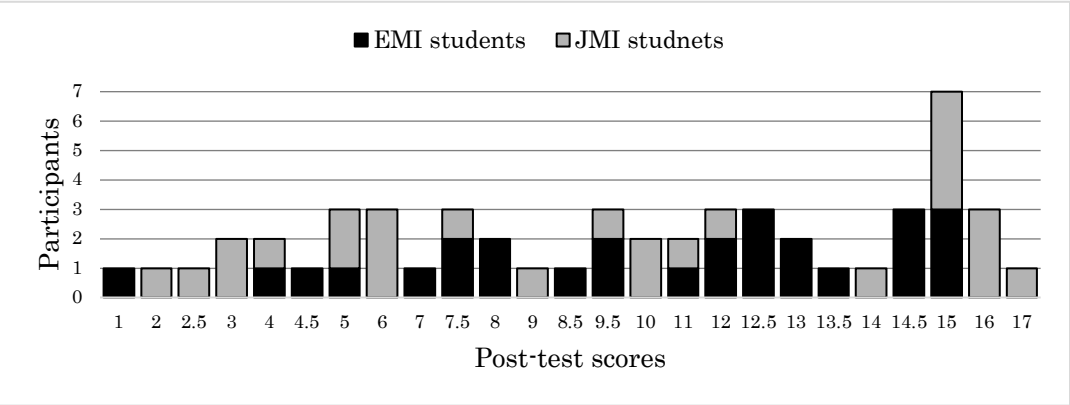


Based on these subsequent analyses, the finding was still upheld that there was no significant difference in Chemistry knowledge according to the medium of instruction, even when analysis was confined to the students who studied through L2 English. Thus, despite the previous literature suggesting that L2 English EMI students face considerable language-related challenges (Macaro et al., 2018), the current results indicated that learning through L2 did not compromise the EMI students' academic performance compared to that of their JMI counterparts.

4.8 Medium of instruction and achievement

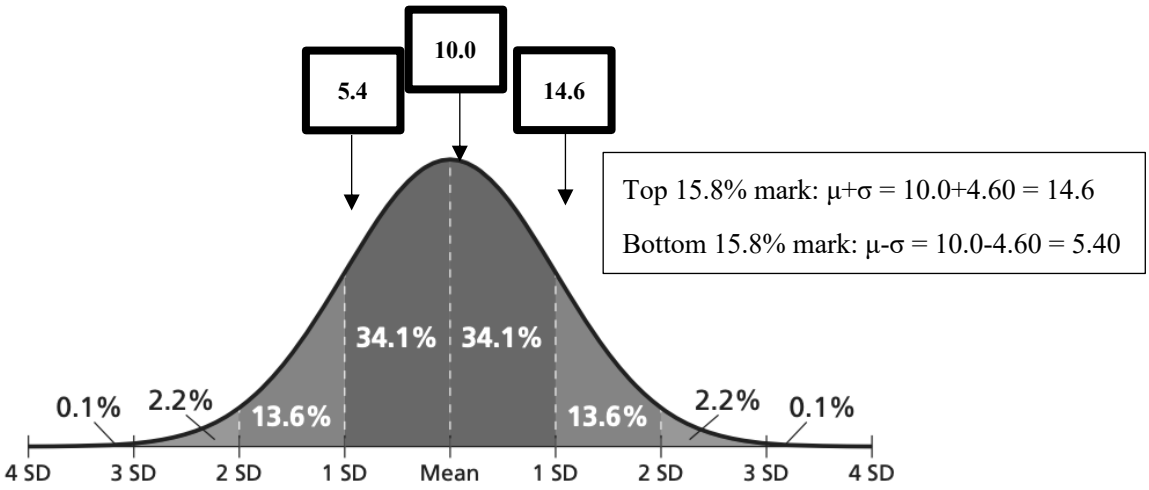
Descriptive statistics were calculated to explore the proportion of high and low performing students according to the mediums of instruction. When comparing the EMI (n=27) and JMI (n=26) students, the analyses have so far indicated that their mean test scores as groups were not significantly different from each other. Nevertheless, subsequent analyses based on the combined dataset of the EMI and JMI students below (n=53) revealed that there were clear descriptive differences in the proportion of the high and low performing students based on the medium of instruction, highlighting the need to explore the experiences and perceptions of individual students (see Figure 14).

Figure 14: Distribution of the participants' post-test scores (n=53)



To compute the cut-off lines of the higher and lower end of the test score distribution, I adopted the three-sigma rule because the data (n=53) for the post-tests were approximately normally distributed (see Appendix 18 for a z-score table) (Pukelsheim, 2012). Accordingly, the values (above 14.6 for the top end and below 5.40 for the bottom end mark out of 17) within two and three standard deviations of the mean accounted roughly for the top and bottom 15.8% of the distribution (see Figure 15).

Figure 15: Top and bottom 15.8% of the post-test score distribution



- μ is the mean of the distribution; σ is the standard deviation.
- Z-scores for 5.4 and 14.6 are .1587 and .8413, respectively, corresponding to the top and bottom 15.8% of the observations using the Z-score table.

When exploring the descriptive statistics of the observations above 84.2% and below 15.8% of the post-test score distribution (see Table 15 below), it became clear that a higher proportion of the JMI students dominated the top range of the post-test performance distribution. Table 15 below reveals that 30.8% of the overall JMI students (n=8) fell within the top 15.8% range of the test performance distribution, compared to just 11.1% of their EMI counterparts (n=3). Conversely, this proportion for the bottom 15.8% end of the distribution was more equal across the groups, entailing the 14.8% of the overall EMI students (n=4) and the 26.9% of their JMI (n=7) counterparts. Thus, although the results of an ANCOVA showed no significant overall mean differences between the EMI and JMI students, the observation of their individual scores suggested that the 19.7% of the higher proportion of the JMI students were represented in the top 15.8% of the post-test distribution than their EMI counterparts. To this end, I emphasise the potential issues regarding simply computing and contrasting group mean scores. In my study, the average values did not tell the whole story; a qualitative examination was hence required to tease out the experiences and perceptions of individual students and their learning experiences in JMI or EMI, as the subsequent section reveals.

Table 15: Top and bottom 15.8% of the post-test score distribution

		EMI students (n=27)		**JMI students (n=26)**	
Post-test score	**Post-test score %**	**Frequency**	**Cumulative percent**	**Frequency**	**Cumulative percent**
17.00	100.0%	0	0%	1	3.8%
16.00-16.99	94.1-99.9%	0	0%	3	11.5%
14.60-15.99	85.9-94.1%	3	11.1%	4	15.4%
Total		**3**	**11.1%**	**8**	**30.8%**
5.00-5.40	31.8-32.4%	1	3.7%	2	7.7%
4.00-4.99	23.5-29.4%	2	7.4%	1	3.8%
3.00-3.99	17.6-23.5%	0	0%	2	7.7%
2.00-2.99	11.8-17.6%	0	0%	2	7.7%
1.00-1.99	5.9-11.8%	1	3.7%	0	0%
0.00-0.99	0.0-5.9%	0	0%	0	0%
Total		**4**	**14.8%**	**7**	**26.9%**

4.9 Research question 1.2

In the quantitative analysis, the achievement of content learning was operationalised in terms of the pre-post test scores, whereas in the qualitative analysis of the interview data, the second sub-research question presented below explored the students' perceptions of their achievement of content learning in more nuanced terms. While the quantitative results indicated no significant overall measurable mean differences in the gain scores and adjusted post-test scores between EMI and JMI students, the qualitative data offered more detailed insight into the participants' perspectives of content learning according to the two mediums of instruction.

- RQ1: Is there an impact of medium of instruction on undergraduate students' content learning?
    - RQ1.2: How do the EMI and JMI students perceive the medium of instruction to affect their achievement of content learning? (perceived content knowledge)

4.10 Content learning-related challenges

The first part of the qualitative results is presented drawing on four particular themes developed deductively from previous EMI research on academic challenges and inductively by the analysis of qualitative data that emerged from the first round of interviews involving EMI (n=13) and JMI students (n=4) and the second round of interviews with EMI (n=6) and JMI (n=2) students (see Appendix 2 for a summary of the interview participants). The 'Time 2' interviews identified their challenges that emerged at the end of the semester and how their challenges manifested at 'Time 1' persisted and dissipated throughout the semester. Four problems facing the EMI and JMI participants (the four themes) are: (1) comprehension of the course textbook; (2) comprehension of the course lectures; (3) Final Grade; and (4) adjustment to the new academic and cultural conventions of the Chemistry course. These emergent themes are based on challenges distinctive to EMI (e.g. understanding English grammar structures)

as well as general academic challenges common to both EMI and JMI (e.g. acquiring technical terms). Below, the results will be presented and discussed in relation to these themes. First, four particular challenges facing the EMI and JMI participants will be discussed followed by the second part of the chapter which then examines four types of positive learning outcomes through EMI and JMI. Table 16 below presents a summary of the interview participants.

Table 16: Interview participants

Name	Language of instruction	Gender	IELTS	First Language	Pre-test	Post-test	Final Grade
1 Hikaru*	EMI	Male	6.5	Japanese	2	5	C
2 Rina*	EMI	Female	6.5	Japanese	0	7.5	C
3 Marii*	EMI	Female	7.0	Japanese	0	1	C
4 Ayaka*	EMI	Female	7.0	Japanese	0	8.5	C
5 Minori*	EMI	Female	7.0	Japanese	0	9.5	B
6 Lisa	EMI	Female	8.0	Japanese	3	12.5	A
7 Miyazaki*	EMI	Female	8.0	Japanese	13.5	15	A
8 Charlotta	EMI	Female	8.0	German	0	9.5	C
9 Tikamon	EMI	Female	7.5	Thai	5.5	12.5	A
10 Kiley	EMI	Female	7.0	Swazi	1	4	E
11 Maxwell	EMI	Male	-	English	0	8	D
12 Saika	EMI	Female	-	Bilingual	0	7.5	C
13 Emma	EMI	Female	-	Bilingual	0	12.5	A
14 Oda*	JMI	Male	5.0	Japanese	2	15	B
15 Kyoka*	JMI	Female	5.5	Japanese	5	16	A
16 Nao	JMI	Female	7.5	Japanese	0	3	C
17 Shizuka	JMI	Female	6.0	Japanese	2	6	C

Notes: Bilingual students self-reported to obtain both Japanese and English as their first languages; Grading scale for the pre-post-test ranges from 0 to 17, and Final Grade ranges from A to E (A, B, C, D, E) where A represents the highest possible grade, and E the lowest possible pass grade (fail); Student with an asterisk (*) participated in both the first and second round interviews.

4.10.1 Comprehension of the course textbook

A first theme related to the challenges facing the EMI and JMI students was their inadequate comprehension of the course textbook. For ease of discussion of results, this theme was divided into four sub-themes during data analysis: insufficient vocabulary knowledge; slow reading speed; reliance on translation; and inefficient reading.

A first sub-theme related to the students' low comprehension of the course textbook was their difficulty in understanding vocabulary used in the course textbook, reported as an issue by the EMI students (coded 1.9 times per participant) as well as the JMI students (coded 1.3 times per participant). The EMI students faced challenges with technical and general academic vocabulary, reporting their reliance on dictionaries to search for unfamiliar words. For instance, Rina (IELTS 6.5, post-test 7.5) 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, Minori (IELTS 7.0, post-test 9.5) stated 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. Further, Hikaru (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, Rina, Minori and Hikaru all had to look up the meanings of unfamiliar technical and academic words. Conversely, the JMI students, Oda (IELTS 5.0, post-test 15) and Shizuka (IELTS 6.0, post-test 6) only occasionally had to look up technical terms. When doing so, it was merely to consolidate their disciplinary knowledge.

> **Oda (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.
>
> **Shizuka (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 it was a common problem for all participants to learn unfamiliar vocabulary, they experienced this challenge differently according to the medium of instruction. In the instances given above, the EMI students, Rina, Minori and Hikaru struggled to comprehend both technical and general academic terms, whereas the JMI students Oda and Shizuka only lacked discipline specific terms. The most apparent difference was the type of vocabulary they struggled with and how frequently they relied on dictionaries.

A second sub-theme that emerged from the interviews was the students' frustration with their somewhat slow reading speed, reported as an issue by the EMI students (coded 2.7 times per participant) and to a lesser extent by the JMI students (coded 1.0 times per participant). Unlike the JMI students who learned in their L1, the EMI students' problem was magnified by their limited grammatical knowledge in L2 English. For example, although re-reading the same sections multiple times, Marii (IELTS 7.0, post-test 1) struggled to fully comprehend certain English sentence structures which impeded her reading speed. Consequently, the bigger consequence of this challenge was that she could not complete her assigned readings before attending her lectures.

> **Marii (EMI):** I tend to stop and read the same sentences many times. It takes me a while to understand the meanings of English sentences. I have to think about English grammar, like SVO, SVC. So I always run out of time to finish my assignments before going to my lectures.

On the contrary, the JMI students, Kyoka (IELTS 5.5, post-test 16) and Shizuka (IELTS 6.0, post-test 6) could maintain reading at a higher speed without any loss of comprehension.

> **Kyoka (JMI):** I usually skim through my textbook before my lectures. Reading quickly is enough to follow my lectures. I only read the same sections more than once when the concepts introduced there are difficult to understand.

> **Shizuka (JMI):** I often just read through the textbook quickly right before I go to my lectures. I often do it when I commute to the university on the train. I think I generally have no problem reading and understanding the textbook.

However, another JMI learner, Nao (IELTS 7.5, post-test 3), who had not studied Chemistry before university, also struggled to understand the textbook. Unlike Kyoka, she had to read the textbook carefully and slowly as she had to look up even the basic terms which Kyoka would have learned from her high school science lessons.

> **Nao (JMI):** It takes me a while to read the textbook because I often look up basic science words on the internet, such as “oxidation” and “unit cell”. I didn’t study Chemistry at high school so I don’t have enough content knowledge. The textbook doesn’t always explain these basic terms.

Thus, while Nao (JMI) and Marii (EMI) both struggled to read the textbook efficiently, Nao lacked technical vocabulary knowledge, whereas Marii had the additional difficulty of having to understand English grammar structures. This indicates that studying through L1 did not by osmosis result in reading at an increased speed. In contrast, Kyoka (JMI) with more advanced knowledge of technical vocabulary, could read the text at higher speeds without sacrificing her comprehension.

The third challenge facing the participants was their reliance on translation of English terms into their L1 equivalents, which was a theme that was present in the EMI student data (coded 1.6 times per participant), but none of their JMI counterparts mentioned this challenge. The EMI students who had previously studied Chemistry through L1 were

acutely aware that their educational background placed them at a considerable disadvantage for acquiring specialist terminology. For example, Ayaka (IELTS 7.0, post-test 8.5) and Minori (IELTS 7.0, post-test 9.5) often looked for the alternative L1 words to compensate for their inadequate L2 vocabulary knowledge. Adopting translation more often, Hikaru (IELTS 6.5, post-test 5) even had to translate a whole sentence or paragraph because he found it easier to understand concepts in Japanese.

> **Ayaka (EMI):** I studied Chemistry in Japanese at high school. When I come across English Chemistry terms in the textbook, I have to translate them to see if I know them in Japanese. It's an ineffective way to learn Chemistry.
>
> **Minori (EMI):** I translate English words into Japanese because I don't understand a lot of them. It takes more time but helps me understand the textbook more easily.
>
> **Hikaru (EMI):** It's not the best way to improve my English but I often translate a whole sentence into Japanese to understand difficult concepts in Japanese first.

In contrast, none of their JMI counterparts reported adopting such a translation strategy.

> **Shizuka (JMI):** I never translate Japanese terms into English. My lectures and textbooks are all in Japanese. **(IELTS 6.0, post-test 6)**

Therefore, Ayaka, Minori and Hikaru who had previously learned Chemistry in L1 prior to university all placed great reliance on translation. In particular, Hikaru had to translate a whole sentence and paragraph due to his limited English proficiency. While they all appeared to recognise the value of translation as a useful coping strategy, they also found it problematic to resort to Japanese constantly in the EMI program.

The final challenge, and one indeed that was more prevalent among the EMI students (coded 2.8 times per participant) than the JMI students (coded 1.8 times per participant) was their inefficient reading of the textbook. Although spending longer independent study hours than those in L1, the EMI students were unable to achieve the same comprehension as that of their L1. For example, Minori (IELTS 7.0, post-test 9.5),

Marii (IELTS 7.0, post-test 1) and Lisa (IELTS 8.0, post-test 12.5) perceived reading in English to be cumbersome and inefficient. Their challenge entailed sacrificing clarity, retaining less information, simplifying concepts, and lacking a nuanced understanding, as epitomised in the excerpts below.

> **Minori (EMI):** I usually spend longer hours preparing for my EMI lectures than my Japanese lectures, but I only understand 85% of my textbook. I also simplify difficult concepts, so my understanding is always shallow.
>
> **Lisa (EMI):** I read the textbook more carefully because it takes me a while to memorise information in English.
>
> **Marii (EMI):** I'm never confident in my ability to understand the textbook correctly in English. I have to ask my course mates and professor to check my understanding.

Similarly, the JMI student, Nao (IELTS 7.5, post-test 3) and Shizuka (IELTS 6.0, post-test 6) were also eager to improve their comprehension of the textbook. Consistent with the problems facing the EMI students above, their main problem also lay in sacrificing clarity, simplifying concepts, and relying on peers and teachers. This indicates that textbook comprehension was a general academic challenge common to both the EMI and JMI students. However, the JMI students mainly grappled with their limited prior disciplinary knowledge and ability to understand the subject matter.

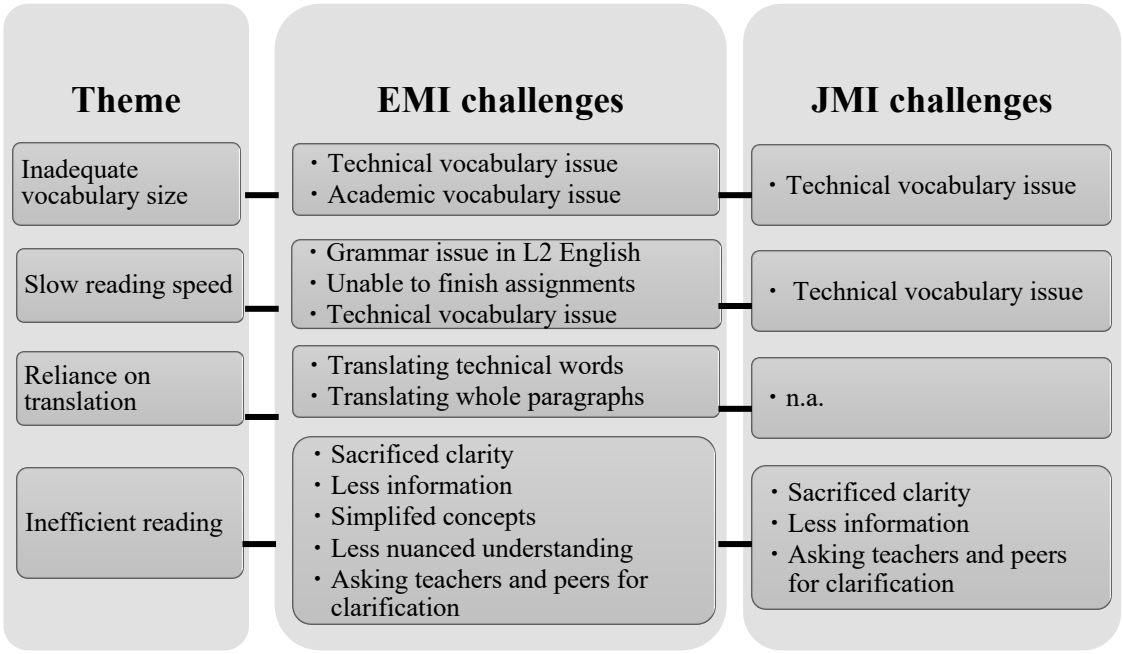
> **Nao (JMI):** I'm not good at studying Chemistry so I usually spend longer hours reading the textbook for this course than other courses. But I still don't understand many concepts in the textbook. I didn't take natural science subjects at high school.
>
> **Shizuka (JMI):** My understanding of the textbook is always shallow because I don't have a good understanding of the subject. I tend to ask my teachers and course mates for help.

Thus, while the EMI and JMI students both expressed the comprehension problems, their challenges were caused by different factors (language or content) depending on the medium of instruction. Unlike the JMI students, the EMI students could not achieve the same level of comprehension as that of their L1, indicating that EMI learning could be

cumbersome and inefficient. Nevertheless, seeing that causes of student challenges were merely perceived causes, the causes of their challenges could be caused by a number of factors, such as prior experience of learning through EMI, topical knowledge, and ability to learn the subject matter.

Overall, the first theme of reading comprehension was explored according to the medium of instruction. Figure 16 below illustrates the challenges specific to EMI in addition to the challenges common to both EMI and JMI. Although there were general academic challenges (e.g. technical term issues), their challenges also varied according to the medium of instruction. The EMI students had to adapt to a linguistically demanding learning environment while also wrestling with a new academic discipline.

Figure 16: Textbook comprehension challenges by MOI (EMI/JMI)



4.10.2 Comprehension of the course lectures

The second theme related to the challenges facing the EMI and JMI students was their inadequate comprehension of the course lectures. This main theme was divided into four sub-themes: low concentration; low comprehension of humour and jokes from the

lecturer; a ‘double whammy’ of content and language-related issues; and low comprehension of technical terms.

A first sub-theme related to students’ low comprehension of the lectures was their difficulty in sustaining concentration, which was a theme that was reported by both the EMI (coded 1.8 times per participant) and the JMI students (coded 1.3 times per participant). Transitioning from a Japanese-medium school, the EMI student, Hikaru (IELTS 6.5, post-test 5) was unaccustomed to listening to long stretches of English mediated-lectures, expressing constant frustration with being easily distracted. Below he shared an instance where he missed a deadline due to concentration fatigue in class.

> **Hikaru (EMI):** Listening in English is generally exhausting. I have missed an important announcement before because I was not paying attention to my professor’s announcement.

Similarly, Minori (IELTS 7.0, post-test 9.5) emphasised a negative impact of her shorter attention span on her understanding of EMI lectures. She perceived that listening through English was a large drain on her energy. In the instance given below, she was unable to fully understand a concept introduced towards the end of the lecture.

> **Minori (EMI):** I remember the professor decided to use the last five minutes of the lecture to explain ‘the effective nuclear charge’. It was a particularly difficult concept and I was tired of listening in English so I zoned out and stopped listening to him.

Despite learning through Japanese, Oda (IELTS 5.0, post-test 15) also struggled to engage in a consistent level of concentration during the entire duration of his lectures. This indicates that concentration fatigue was a general school-university transitional challenge common to all the students, irrespective of their mediums of instruction.

> **Oda (JMI):** I can usually stay focused on my lectures, but I sometimes stop listening when I’m tired. It’s hard to concentrate on the entire duration of the 70-minute lectures. My high school classes were only 50 minutes. I also have two more classes in the morning so I’m usually already tired when I arrive at the Chemistry class.

In short, although both the EMI and JMI students struggled to focus on long stretches of university-style lectures, it was also evident that they faced different types of challenges. Hikaru and Minori (EMI) whose lectures were delivered in L2 English, constantly suffered from a short attention span. Listening through L2 English required them a high level of concentration, which in turn undermined their comprehension. Conversely, Oda's (JMI) problem was not as acute as that of Hikaru and Minori. In most cases he was able to maintain a certain level of concentration.

The second sub-theme related to the students' low lecture comprehension was their difficulty in understanding humour and jokes used by the EMI lecturer, which was a theme that was only reported by the EMI students (coded 1.5 times per participant). Although disciplinary understanding itself was not directly affected, Minori (IELTS 7.0, post-test 9.5) asserted that understanding humour and jokes was essential for creating inclusive learning atmospheres and rendered lengthy stretches of lectures more engaging. However, Ayaka (IELTS 7.0, post-test 8.5) perceived that humour and jokes were less accessible than disciplinary concepts because they generally emerged more spontaneously and were not always closely linked to the content of the lecture.

> **Minori (EMI):** I like it when my professor tells a joke because it makes everyone laugh and we get to know his personality. My English is too bad to understand it properly, but it's an important part of EMI and being taught by a foreign teacher.

> **Ayaka (EMI):** Even if I listen carefully, I am not confident in understanding jokes because the teacher speaks more colloquially and quickly. I want to understand them better because that's what makes the lectures enjoyable.

Similarly, Miyazaki (IELTS 8.0, post-test 15) with higher English language proficiency also desired to understand humour and jokes more precisely.

> **Miyazaki (EMI):** The professor sometimes starts lectures by introducing recent Chemistry news. One of them recently was about a handmade rocket in a London park. He tells a joke to make us laugh. I can usually understand it well but I would like to understand it better.

On the contrary, none of the JMI interviewees reported this issue, indicating that this was a problem only specific to the EMI students, and worse for those of low English proficiency who attributed this to be the main reason for their poor comprehension. Irrespective of the L2 proficiency, all the EMI students however recognised humour and jokes as an essential part of their EMI learning experience.

The third challenge, and one indeed that was only reported by the EMI students (coded 2.7 times per participant), was a 'double whammy' of content and language-related issues. Unlike the JMI students, the EMI students had to combat additional language-related challenges while understanding complicated academic concepts. In contrast to reading textbooks at their own pace, students had to listen, interpret and understand lectures instantaneously through L2 English. Consequently, this extra linguistic hurdle coupled with a demanding task of learning new concepts impeded their lecture comprehension. For instance, Tikamon (IELTS 7.5, post-test 12.5) wrestled with English language challenges amid concerns over the quality of her listening comprehension.

> **Tikamon (EMI):** Learning chemistry is difficult, but it is more challenging because I learn in English.

Similarly, Marii (IELTS 7.0, post-test 1) and Ayaka (IELTS 7.0, post-test 8.5) noted that understanding English medium-lectures was already taxing but their problem became even more insurmountable when having to take notes during their lectures.

> **Marii (EMI):** It's difficult to understand concepts in English, but I struggle even more when taking notes. When I want to listen, I have to stop taking notes. It's hard to listen and write at the same time.

> **Ayaka (EMI):** I have to use both English and Japanese when taking notes because my English is too bad. I don't know how to spell many English words. It's also quicker to write in Japanese. But I want to be able to take notes entirely in English one day. Listening in English and writing down notes in Japanese is not natural.

Conversely, the JMI student, Kyoka (IELTS 5.5, post-test 16) below did not experience any of these problems mentioned by the EMI students, given that her lectures occurred through Japanese. She even mentioned that she would have struggled to achieve the same learning quality if she had studied through EMI.

> **Kyoka (JMI):** I learn in Japanese and I can listen to my professor and take notes at the same time. I practised note taking in high school. But it would be impossible to listen and take notes in English. I couldn't do both together.

Thus, the additional linguistic issues (e.g. inadequate vocabulary size) severely impeded the EMI students' lecture comprehension. Unlike Kyoka (JMI), both Marii and Ayaka (EMI) had to sacrifice the standard of their note taking in their EMI lectures.

The final sub-theme related to lecture comprehension was the students' frustration with their low understanding of technical terms used in their lectures, which was a challenge reported by the EMI students (coded 2.0 times per participant) and to a lesser extent by the JMI students (coded 1.3 times per participant). The professor's use of a plethora of unfamiliar technical terms hindered both the EMI and JMI students' comprehension of the subject matter.

For instance, Ayaka (IELTS 7.0, post-test 8.5) had to relearn the English version of the Chemical elements in the periodic table which she had already learned at high school, including "hydrogen", "helium" and "nitrogen". Similarly, Rina (IELTS 6.5, post-test 7.5) had previously learned "valence electrons" at high school but did not realise when the same concept was introduced in her lecture because she was not familiar with the English word "valence".

> **Ayaka (EMI):** I learned Chemistry in Japanese at high school, so I have to memorise so many words in English and all the Chemical elements in the periodic table all over again in English. It's time-consuming.

> **Rina (EMI):** I always read the textbook before my lectures but in one of the lectures 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.

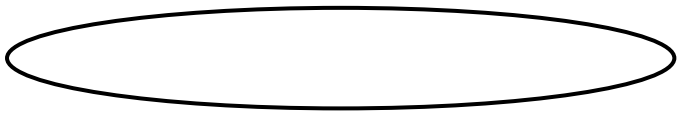
On the contrary, the JMI student Shizuka (IELTS 6.0, post-test 6) who had previously studied technical terms in Japanese did not have to learn these terms again, which she described as "lucky" in the excerpt below. Instead, she was overwhelmed with the amount of new technical terms she was exposed to during the course, which was a general academic challenge also reported by Ayaka (EMI). Students' as-yet limited knowledge of technical terms was a recurring theme in the interviews which prevented them from fully understanding lectures. Additionally, the JMI student Kyoka (IELTS 5.5, post-test 16) could retain new technical terms more easily than Rina (EMI).

> **Shizuka (JMI):** There are many technical terms I need to learn in the course. But I'm lucky that there are many words I remember from my high school class.

> **Kyoka (JMI):** I usually only read the textbook briefly before my lectures but I understand almost all the terms when I hear them in lectures.

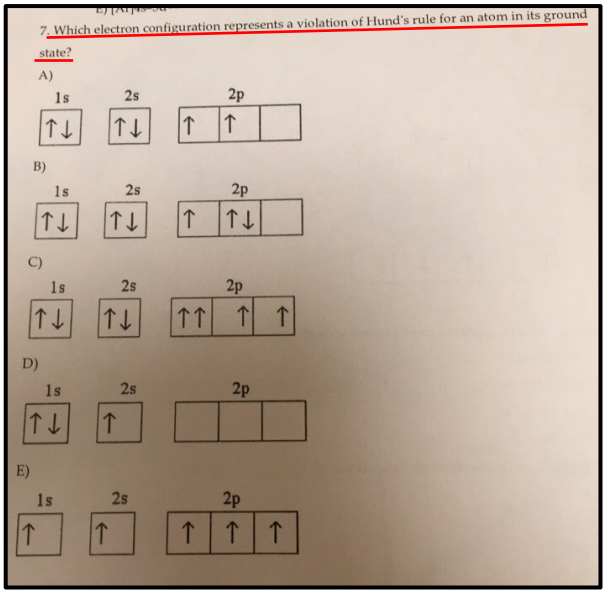
The EMI students also lacked a rich repertoire of vocabulary containing both scientific and everyday dual meanings used in the Chemistry domain (such as "shell", "element", "activity", "stability", "capacity"). For example, Marii (IELTS 7.0, post-test 1) shared the following example problem set from one of her lectures which she could not complete due to unfamiliar vocabulary, such as "violation" and "ground state".

> **Marii (EMI):** I was not confident in understanding the instruction because I didn't know the meaning of the words "violation" and "ground state".



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

Resource document shared during an interview by Marii: Problem sets of 'Concepts in Chemistry'.



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). Hikaru (IELTS 6.5, post-test 5) identified the relation between his vocabulary knowledge and lecture comprehension.

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

1. What is the electronic configuration of O?
2. Write a plausible Lewis structure for $SiCl_3^-$, showing all valence electron pairs. Note that the species are charged – be sure to write the charges next to the structures. If there is a resonance hybrid, draw it too. Give the VSEPR notation (AXE).

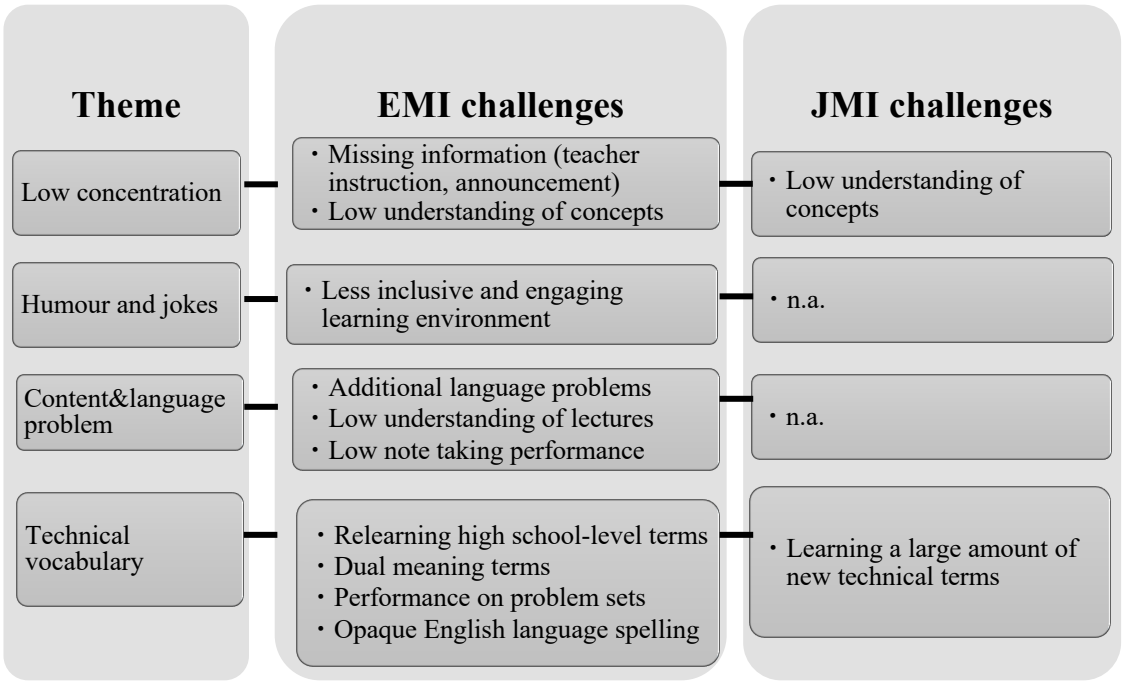
It also became apparent that their vocabulary-related issues stemmed from the opaque English language spelling system. Hikaru (IELTS 6.5, post-test 5) and Ayaka (IELTS 7.0, post-test 8.5) were unable to look up basic technical vocabulary (e.g. tetrahedral molecular) due to its unclear English spelling.

> **Hikaru (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.
>
> **Ayaka (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.

To sum up, although all the interviewees faced difficulty in comprehending lectures laden with specialised domain-specific vocabulary (e.g. molecular orbital theory), they experienced the challenges differently depending on their medium of instruction. The EMI students, in particular those who had previously learned Chemistry in Japanese faced the significant additional difficulty of having to acquire the English versions of the high school-level terms (e.g. "hydrogen", "helium") and dual meaning terms (e.g. "shell", "activity"), which impeded their performance of problem sets (e.g. calculating equations). The opaque English language spelling system also exacerbated their problems. In contrast, the JMI students were also overloaded with the large amount of new technical terms introduced in the course but were more advantaged in their prior vocabulary knowledge acquired from high school. As a result, the JMI students in most cases could easily overcome the problems through reading the textbook before lectures.

Overall, the second theme (i.e. lecture comprehension issues) was explored according to the medium of instruction. Figure 17 below illustrates the challenges distinctive to EMI and the challenges common to both EMI and JMI. Although there were general challenges (e.g. low concentration issues), L2 English students in the EMI course faced a considerably more taxing challenge than their JMI counterparts while adapting to a linguistically more demanding learning environment.

Figure 17: Lecture comprehension challenges by MOI (EMI/JMI)



4.10.3 Final Grade

The third theme related to the students' challenges was their difficulty in achieving a respectable Final Grade. For ease of discussion of results, this theme was divided into four sub-themes during data analysis: poor management of workload; low performance of the final course exam; low performance on the 'Comment Worksheet' (explained later); and motivation and effort as mediators of Final Grade.

A first challenge facing the EMI and JMI students related to their difficulty in obtaining a respectable Final Grade was to manage workload for all enrolled courses, which was a theme that was reported by the EMI students (coded 2.3 times per participant) and to a lesser extent by the JMI students (coded 1.3 times per participant). Seeing that participants enrolled in three or four modules on top of the Chemistry course, it was a common problem for students to have to manage their independent learning hours to ensure the overall performance in all of their courses irrespective of the medium of instruction.

For example, Hikaru (IELTS 7.0, Final Grade C) needed to invest longer learning hours when studying through EMI because it was more linguistically demanding than other courses taught in Japanese. Despite his efforts to strategically study through a study timetable, heavy workload prevented him from managing time effectively, which in turn affected his grades for not only the Chemistry course but also all the other courses.

> **Hikaru (EMI):** I wish I were just taking the Chemistry course because I spend most of my study time revising for Chemistry because it's in English. On top of this, I'm also taking four other courses. I chose to study them in Japanese, but I still need to prepare a lot for them. I created a study timetable and set aside specific times to study for each course, but I'm still worried that I might fail on my courses, especially the Chemistry course because it's in English.

As addressed in the excerpt by Oda below (IELTS 5.0, Final Grade B), managing time effectively was one of the general academic challenges also facing the JMI participants.

> **Oda (JMI):** I've registered for four courses alongside the Chemistry course. I generally have no issue managing my study time. But I spend far longer hours to prepare for my biology course because it's not my expertise and the professor gives us lots of assignments. I would have been able to spend more time on other courses if I were taking an easier course instead. It was a good decision to study biology in Japanese. I'd have to spend far longer hours if it were in English.

Thus, Hikaru (EMI) and Oda (JMI), on the face of it, both struggled to manage workload effectively. However, Hikaru felt obliged to spend longer hours studying for the EMI Chemistry course than the other courses due to its additional linguistic demand. Conversely, Oda (JMI) did not have to prioritise the Chemistry course, but instead perceived that the biology course was more demanding because of other factors, such as course difficulty, professor, assignments, and his own disciplinary knowledge.

The second sub-theme related to the students' Final Grade was their low performance on the final course examination, which was a challenge reported by the EMI students (coded 1.5 times per participant) and to a lesser extent by the JMI students (coded 1.0 times per participant). Rina (IELTS 6.5, Final Grade C) and Hikaru (IELTS 6.5, Final

Grade C) who participated in the second round of interviews carried out after the administration of the final course exam, noted that taking the exam in L2 English negatively affected their exam performance. They admitted that their linguistic challenges (e.g. slow reading speed) persisted throughout the term, leading to reduction in the understanding of disciplinary concepts and a slowing down of content coverage. Rina believed that her scores would have been higher if she had studied through Japanese. Hikaru also could not understand exam questions and instructions accurately and quickly.

> **Rina (EMI):** I could have done better if I had used Japanese. I understood some concepts inaccurately and memorised the incorrect version when preparing for my exam. My understanding of many concepts would have been better in Japanese.

> **Hikaru (EMI):** I wasted a lot of time by reading instructions for each exam question and ran out of the time. I had to read each sentence carefully and sometimes twice or three times to make sure I understood them correctly.

The JMI student Oda (IELTS 5.0, Final Grade B), who completed the exam in Japanese, also expressed some similar problems (e.g. poor time management, low comprehension of disciplinary concepts). Nevertheless, there were also issues exclusive to the EMI students (e.g. reduction in content coverage, misunderstanding of exam questions written in English). This suggests that unlike Oda (JMI), Rina and Hikaru's (EMI) exam performance suffered due to their additional linguistic challenges, whereas Oda's performance was affected by other non-linguistic factors.

> **Oda (JMI):** I think I did very well on the exam but I wasn't confident in one question which required the knowledge of a concept I didn't understand well. I ran out of time to go over my answers at the end because I spent too much time on some of the questions at the beginning.

The third challenge, and one indeed that was only apparent among the EMI students (coded 1.7 times per participant), was their low performance on the Comment

Worksheet[1], which determined 15% of the EMI and JMI students' Final Grade for the Chemistry course. The Comment Worksheet was distributed in every lecture through the semester for various purposes: (1) to administer problem sets (pop quiz); (2) to evaluate students' academic progress; (3) to assess their learning needs; (4) to gather their general questions; and (5) to record their attendance. Based on their responses, the professors assigned scores on a scale ranging from 1 (poor) to 5 (excellent) in increments of 0.5.

Unlike the JMI students, the EMI students struggled to solve problem sets and formulate comments in English due to their low lecture comprehension and insufficient English writing skills. For example, Rina (IELTS 6.5, Final Grade C) failed to understand lectures adequately to perform well on problem solving quizzes. Kiley (IELTS 7.0, Final Grade E) struggled to relay her ideas using precise wording in English.

> **Rina (EMI):** I don't like these pop quizzes because I don't necessarily understand every part of my lectures sufficiently. I sometimes just don't know how to solve these problem sets.
>
> **Kiley (EMI):** When I want to ask follow-up questions, I don't know how to ask them. I always end up leaving only one or two sentences because I need more time to revise what is covered in my lectures.

Conversely, the JMI students, Shizuka (IELTS 6.0, Final Grade C) and Oda (IELTS 5.0, Final Grade B) below were largely enthusiastic about the Comment Worksheet without highlighting any pitfalls. They were inclined to report greater ease than difficulty, assuming that JMI content learning would ensure trouble-free performance.

[1]Grading scale ranges from 0 to 25, combining a total score of five Comment Sheets, amounting to a maximum of five points. Although Comments Sheet was distributed in every lecture through the semester, the professor only included five Comment Sheets as part of the assessment (25 points) because Comment Sheet was used on these five occasions to administer problem sets. On other occasions, he distributed Comment Sheet just to gather the students' questions and assess their learning needs.

> **Shizuka (JMI):** I'm thankful for the opportunity to ask questions at the end of each lecture through Comment Sheet. I can write my comments when things are still fresh in memory.

> **Oda (JMI):** I like having quizzes at the end of the lectures because I can clarify my understanding of the lecture. I also like it when the professor explains the questions in the following lectures.

The final sub-theme related to the students' Final Grade was their motivation and efforts as mediators of their challenges, which was a theme reported by both the EMI (coded 3.7 times per participant) and the JMI students (coded 4.0 times per participant). Despite the observed negative influence of EMI on their Final Grade, all participants emphasised the importance of their increased motivation and efforts as a potential mediator of their performance. For instance, the EMI student, Minori (IELTS 7.0, Final Grade B), while acknowledging her limited L2 proficiency, rationalised her decision for studying through EMI by stressing the importance of diligence and determination. Similarly, the JMI student, Shizuka (IELTS 6.0, Final Grade C) argued that achieving high exam scores and Final Grade required a whole different skill set from just being able to understand lectures.

> **Minori (EMI):** My final grade will certainly be negatively affected by learning in English but it will be ultimately decided by how determined I am to make efforts outside my lectures. I can still receive a high final grade as long as I study hard for the final exam.

> **Shizuka (JMI):** Achieving a high final grade requires certain practical skills. I need to be strategic and only study what will be on the exam. I don't need to understand every part of my lectures, but must be able to solve all of the key practice problems. The same questions may be on the exam.

The second round of interviews showed that studying through L1 did not by osmosis result in a respectable Final Grade. The JMI student, Nao and the L1 English EMI student, Saika both expressed their challenges associated with L1 content learning.

> **Nao (JMI):** I realised after taking this course that I'm hopeless at Chemistry. I don't feel motivated to study for this course because it's not enjoyable to learn Chemistry.

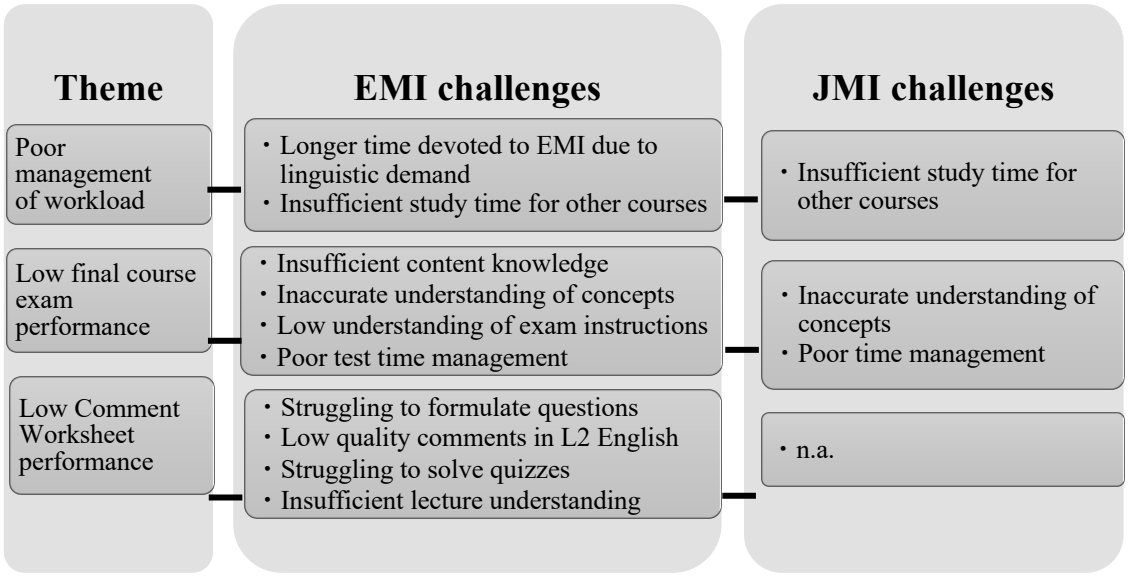
> **Saika (EMI):** I'm worried about my final grade for the Chemistry course because I've not been able to spend a lot of time revising for this course.

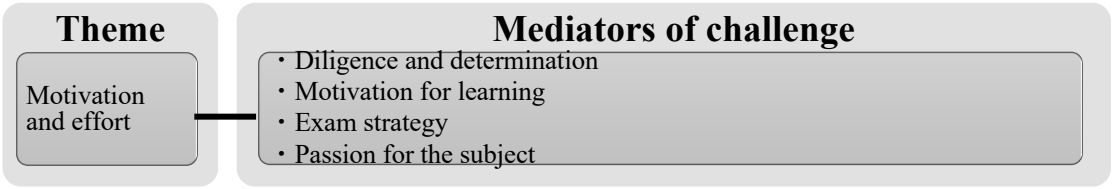
Similarly, Rina (IELTS 6.5, Final Grade C) reported below that it did not necessarily mean that she would have gained higher test results if she had studied in JMI.

> **Rina (EMI):** It would be easier for me to receive a higher final grade in the Japanese course but I still need to study for the course as much as I would for the English course. Having good English skills would help, but I should always be driven to work hard on my subject.

Thus, an exploration of the theme related to students' low Final Grade based on the medium of instruction suggested that the EMI students experienced the additional linguistic challenges (e.g. long learning hours, inaccurate understanding of instructions, low quality comments) which prevented them from achieving a high Final Grade. However, the students also perceived that the crucial factor for successful Final Grade was their diligence, determination, and passion for the subject, which acted as a catalyst for strong motivation, hard work, and effective learning strategies. The figure below summarises these themes based on students' mediums of instruction (see Figure 18).

Figure 18: Problems and mediators of Final Grade by MOI (EMI/JMI)





4.10.4 Adjustment to the new academic and cultural conventions of the Chemistry course

The fourth and final challenge facing the EMI and JMI students was to adjust to the new specialised disciplinary culture and academic conventions of the Chemistry course. This main theme was divided further into two sub-themes: adjustment to the new educational culture; and persistent problems.

A first sub-theme related to the students' adjustment issues was their difficulty in conforming to the new educational culture and academic conventions of the specialised discipline, which was a challenge reported by the EMI students (coded 1.8 times per participant) and to a lesser extent by the JMI students (coded 1.3 times per participant). After receiving schooling in L1, the EMI students encountered insurmountable problems adapting to the novel educational culture and ways of teaching and learning specific to EMI, which were more student-centred, bidirectional, interactive and participatory. Their problems were exacerbated by the new pedagogical methods, indicating that their school-university transition required not only a change in the medium of instruction but also in their ways of thinking about teaching and learning.

Below Hikaru (IELTS 6.5, post-test 5) and Rina (IELTS 6.5, post-test 7.5) highlighted how teaching was conducted differently between school and university, suggesting that their domestic educational background placed them at a considerable disadvantage in adapting to the new educational culture. Conversely, the JMI student Nao (IELTS 7.5, post-test 3) noted that her professor adopted a teaching method similar to how she was taught at her Japanese-medium high school, which she described as "teacher-centred" in

the excerpt below. Although neither Nao nor another JMI student Kyoka (IELTS 5.5, post-test 16) had to adapt to studying in English, they were overwhelmed with the demand of academic acculturation into the new disciplinary norms and practices that were central to their university learning experience (e.g. longer lecture duration).

> **Hikaru (EMI):** The professor likes nominating us for our opinions. I usually just need to say one or two words and his questions aren't linguistically difficult. But I'm still not comfortable with the idea of thinking on my feet and sharing my ideas with the rest of my class. I was always quiet and passive at high school.
>
> **Rina (EMI):** Some of my course mates especially the international students have good English and ask many questions to challenge the teacher. That's how they learn overseas? I'd like to do the same but I'm worried about my English.
>
> **Nao (JMI):** My lectures are mostly teacher-centred. I prefer this teaching style because I learn well if I listen first and then absorb what is being covered. One difference from school is that we're now expected to study more independently. Our lectures are also more difficult, fast-paced and longer.
>
> **Kyoka (JMI):** I wouldn't like to study in English because I'd need to be more active and answer teachers' questions in front of the whole class. It's more of an American style. In my class, I can just listen to the teacher the whole time. Because the class is far more difficult than my high school class, I want to first focus on learning Chemistry without worrying about the different learning culture.

Hence, an examination of the sub-theme related to the new educational culture based on the medium of instruction indicated that both the EMI and JMI learners had to conform to the novel academic and social conventions of learning and teaching required for the university courses (e.g. long lecture duration, independent learning, fast-paced lectures, and advanced disciplinary level). Despite these common transition challenges, the EMI students faced the additional difficulty of having to be more critical, analytical, challenging what is being taught, and to share opinions and provide answers on the spot. In the aforementioned excerpt, this unique EMI disciplinary culture was the very reason which led Kyoka (JMI) to decide against studying through EMI. She was not fully prepared for both the linguistic and cultural transition she would have experienced by learning through EMI.

The final problem facing the participants during the whole semester, and one indeed that was only reported by the EMI students (coded 1.4 times per participant) was that their content learning-related issues were long-lasting rather than short-term teething problems. Based on the second round of interviews involving the six EMI and two JMI participants, Minori (IELTS 7.0, post-test 9.5) and Rina (IELTS 6.5, post-test 7.5) who completed the EMI course, suggested that many of their problems that manifested during the first round of interviews (e.g. slow reading speed, insufficient lecture comprehension, and reduction in understanding) persisted through the entire term.

> **Minori (EMI):** I spent a long time revising for lectures, but I still couldn't understand many concepts. I had to give up trying to learn as well as how I'd learn in Japanese. I also memorised less information after reading the textbook and attending the lectures.

> **Rina (EMI):** Towards the end of the semester, I was getting more and more efficient at learning. My reading speed became faster and I learned many new English terms such as "Equilibrium constant". But I was never fully able to learn as efficiently as how I'd learn in Japanese. I was worried if this problem would drag on for longer.

On the contrary, the JMI student Oda (IELTS 5.0, post-test 15) and Kyoka (IELTS 5.5, post-test 16) compared their learning experience between the first and second round of interviews, describing many of their challenges dissipated through the semester, compared to the long-standing challenges described by Minori and Rina.

> **Oda (JMI):** I remember I initially couldn't keep up with the amount of homework because they were much harder than the one from high school but I eventually got used to working much faster. I just needed some time to overcome my problems because the course was basically similar to how I learned Chemistry in high school.

> **Kyoka (JMI):** I found it difficult to learn so many new technical terms every week but it eventually became easier. I also spent a long time revising for my lectures and textbook at home. I hope I'll get a good grade.

Therefore, Oda and Kyoka (JMI) revealed that their issues dissipated throughout the

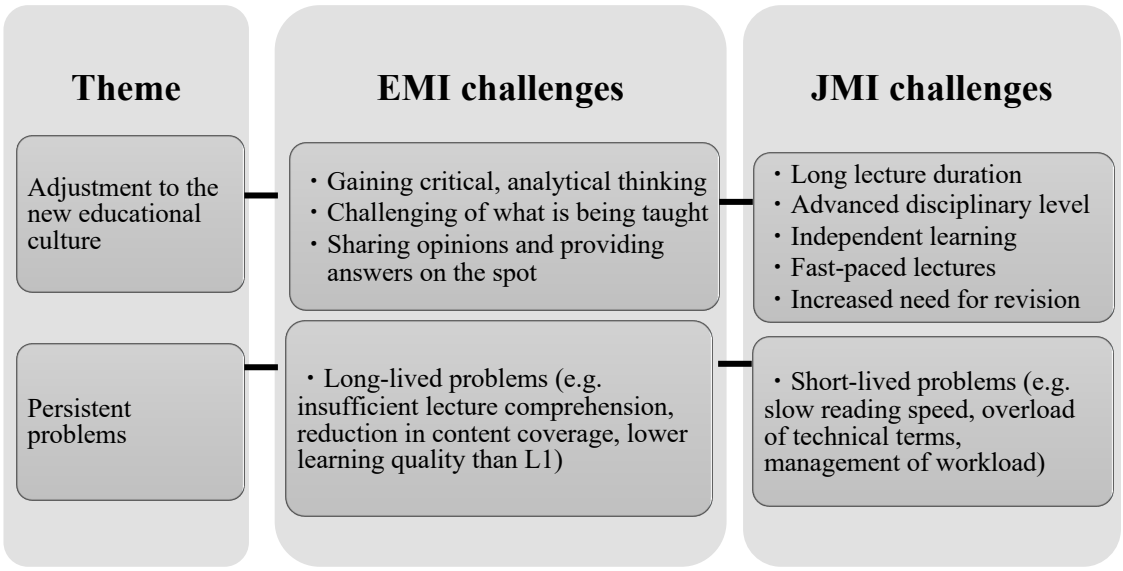
semester. This indicated that one semester was a sufficient length for them to fully mitigate their course specific challenges (e.g. particular course assessment requirements) and other general academic challenges (e.g. management of workload and homework). By the end of the term, they successfully adjusted to their new responsibilities, specific requirements of grading and evaluation, and the disciplinary norms and practices. Similarly, Rina and Minori (EMI) were also able to rectify some of their content learning-related issues identified during the first round of interviews (e.g. acquiring technical terms, improving reading speed, and studying more efficiently). However, they also came to terms with the consequence of learning content through L2 English, admitting that it was unfeasible to fully achieve the same level of disciplinary learning as that of JMI by the end of the term, irrespective of the duration of their invested learning hours. However, it also became clear that it was not all the EMI students whose challenges persisted through the entire term. The EMI student, Miyazaki (IELTS 8.0, post-test 15) who also participated in the second round of interviews was successful in adjusting to the new disciplinary norms and practices by the end of the term, indicating that the array of differential factors, including students' L2 proficiency, influenced the speed and ease of their transition into the EMI learning environment.

> **Miyazaki (EMI):** It used to take a long time to read the textbook at the beginning but I improved my Chemistry knowledge as the term went on. It gradually became easier to read and understand the textbook quickly too. I just had to get used to the professor's teaching style.

To sum up, Figure 19 below outlines the adjustment challenges distinctive to the EMI students as well as the challenges common to both the EMI and JMI students. Although students had to adapt to the demands of the specific educational culture of the new Chemistry course (e.g. advanced disciplinary level), irrespective of their medium of instruction, the EMI students had to make the additional effort to conform to the new academic culture specific to EMI (e.g. answering questions in English on the spot). Similarly, while the JMI students could fully overcome their adjustment challenges by the end of the term (e.g. acquiring the requisite technical terms), the EMI students encountered longer-lasting problems throughout the semester (e.g. reduction in content

coverage and comprehension), admitting that their disciplinary understanding did not reach the same level as that of their L1 instructed courses.

Figure 19: Adjustment challenges by MOI (EMI/JMI)



4.11 Success in content learning

This section of the qualitative results centres on data surrounding the positive impacts of learning content through EMI and JMI, going beyond educational outcomes of achieving a high grade on tests. Accordingly, patterns in the interview data indicated that positive outcomes of content learning through EMI and JMI primarily centred on (1) improved English proficiency; (2) improved motivation; (3) access to a wider range of learning resources; and (4) long-term career advancement. These themes involved success distinctive to EMI (e.g. accessing English medium resources) as well as general academic success common to both EMI and JMI (e.g. improved content knowledge). The EMI students compensated their costs to content learning for these subsidiary benefits exclusively linked to EMI content learning.

4.11.1 Improved English proficiency

A first theme related to the positive outcome of gaining content knowledge through EMI and JMI was their improved English proficiency, which was the benefit that was almost exclusively reported by the EMI students (coded 4.6 times per participant) rather than the JMI students (coded 0.5 times per participant). All the L2 English students in the EMI course interviewed regarded the development of their English language proficiency as one of the expected outcomes of their EMI education. Although the primary aim in the course was to achieve a sufficient level of understanding in Chemistry, Hikaru (IELTS 6.5, post-test 5) suggested the practical value of the EMI course as an additional platform to improve his English skills on top of the in-sessional English for academic purposes (EAP) courses which he was taking at the same time.

> **Hikaru (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, Tikamon (IELTS 7.5, post-test 12.5) valued the EMI course in relation to the development of her academic speaking skills.

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

Based on the second round of interviews, Rina (IELTS 6.5, post-test 7.5) observed an improvement in her English proficiency in the areas of listening and reading, compared to those at the beginning of the semester.

> **Rina (EMI):** Listening to my professor and reading the textbook in English through this semester was helpful to improve my listening and reading skills.

The EMI learners generally regarded linguistic improvement as a crucial achievement measure of EMI, with some drawing links between the EMI and EAP provision as additional mechanisms to support their English proficiency development goals.

Although this was a theme that emerged almost exclusively from the EMI data, one JMI student Oda (IELTS 6.0, post-test 6) reported an improvement in technical vocabulary knowledge in English. Below he explained that some of his optional reading assignments included English based sources and the lecture slides occasionally introduced English translations alongside the original Japanese terms. This indicates that while the increased proficiency was a particular benefit almost exclusively linked to EMI, Oda (JMI) also reported an improvement in his English vocabulary knowledge.

> **Oda (JMI):** I learnt some technical vocabulary in English because some optional reading I did was in English and the teacher sometimes introduced terms in English.

Such benefits may be the result of technical words in English being used in academic disciplines in Japanese due to its dominance as an academic lingua franca.

4.11.2 Improved motivation

The second theme related to the benefits of gaining disciplinary knowledge through EMI and JMI was the students' improved motivation, which was a theme that was present in the EMI data (coded 1.5 times per participant) and to a lesser extent in the JMI data (coded 1.0 times per participant).

Despite the observed language-related challenges, all the EMI students interviewed recognised the English medium nature of the course as a driving force of their diligence because they were intrinsically motivated to learn through English. The students were also highly motivated because of their desire to participate in the university's study abroad exchange program. This factor, coupled with the necessity to achieve the required IELTS and Grade Point Average (GPA) scores for the application criteria, served as a vital reason for their determination and perseverance.

> **Minori (EMI):** One of the reasons why I chose this university is because I can use 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 the exchange program. **(IELTS 7.0, post-test 9.5)**

While the participants expressed the additional hours spent on EMI learning were cumbersome and time-consuming, they also considered them as further opportunities to consolidate their disciplinary knowledge. For example, Tikamon (IELTS 7.5, post-test 12.5) was actively eager to maximise the repertoires of her learning materials, resources and dictionaries. Similarly, Ayaka (IELTS 7.0, post-test 8.5) developed new learning strategies by studying through EMI, such as carrying a separate notebook to regularly record unfamiliar words and technical concepts.

> **Tikamon (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.

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

Similarly, the JMI student, Nao (IELTS 7.5, post-test 3) who studied non-science subjects at high school was strongly determined to excel in Chemistry. As it was her long-awaited dream to study Chemistry at university, she was motivated to work hard to improve her content knowledge, recognising its centrality in her academic vision.

> **Nao (JMI):** I've always wanted to study Chemistry at university because I want to research on environmental problems. I didn't study natural science subjects, so I'm now spending a lot of time learning it from scratch.

Thus, although the JMI and EMI students were both highly motivated, learning through EMI in particular appeared to help the students maintain strong motivation, diligence and determination. The EMI cohort—as Tikamon and Ayaka reported—regarded the increased exposure to the content matter as an opportunity 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 comprising of those with the higher L2 proficiency and more positive attitudes towards EMI who volunteered to study through EMI.

4.11.3 Access to a wider range of learning resources

The third theme, and one indeed that was only present in the EMI data (coded 4.8 times per participant) was their access to a wide range of learning resources. Gaining information bilingually using both English and the students' L1 was a distinctive feature of EMI content learning. In contrast, the JMI students reported their excessive reliance on Japanese based materials, as illustrated in the excerpts below:

> **Minori (EMI):** When I don't understand concepts, I use all the available 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. **(IELTS 7.0, post-test 9.5)**

> **Kyoka (JMI):** I use additional materials to help my learning at home but they are always Japanese. Why would I use English sources when I have all the information in Japanese? **(IELTS 5.5, post-test 16)**

> **Nao (JMI):** I tried to read some materials in English because the teacher included several English sources on Moodle but I didn't really understand them so instead I just read the Japanese materials. I avoided the English ones. **(IELTS 7.5, post-test 3)**

Further to this, due to the wider range of online learning resources available in English than their L1s alone, the EMI participants preferred English as the language of instruction. Admitting this subsidiary benefit of using English, Ayaka (IELTS 7.0, post-test 8.5) and Minori (IELTS 7.0, post-test 9.5) supplemented their learning with online materials and practice exercises in English.

> **Ayaka (EMI):** There are more science information available written in English than Japanese on the Internet.

> **Minori (EMI):** I like revising Chemistry concepts using a Youtube channel called Khan Academy which offers short lessons in the form of videos.

Thus, the EMI students would not have been able to access the same types of information without using English, supporting the value of using English in the chosen academic discipline. Conversely, Kyoka and Nao (JMI), who relied exclusively on Japanese based sources, had limited access to a smaller range of resources.

4.11.4 Long-term career advancement

The final theme related to the benefits of gaining content knowledge through EMI and JMI was the students' long-term career advancement, which was reported by both the EMI students (coded 2.9 times per participant) and the JMI students (coded 1.3 times per participant). The EMI participants identified the advancement of their internationally oriented career as a crucial measure of their EMI success. Rina (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 future academic career overseas.

> **Rina (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, Lisa (IELTS 8.0, post-test 12.5) valued the international classroom landscape where her learning involved interactions with overseas students, serving as an additional advantage of EMI to accomplish her future employment goals.

> **Lisa (EMI):** I prefer studying in English because I like talking with my classmates from different countries. I want to work for an international organisation using English.

Similar to Rina (EMI), the JMI students, Kyoka (IELTS 5.5, post-test 16) and Shizuka (IELTS 6.0, post-test 6) were also motivated to improve her disciplinary knowledge through the JMI course because their main long-term aim was to study at a graduate school and become a scientist in Japan.

> **Kyoka (JMI):** I've always wanted to study Chemistry at university because I want to study at a graduate school and then work as a scientist at a pharmaceutical laboratory in Japan.

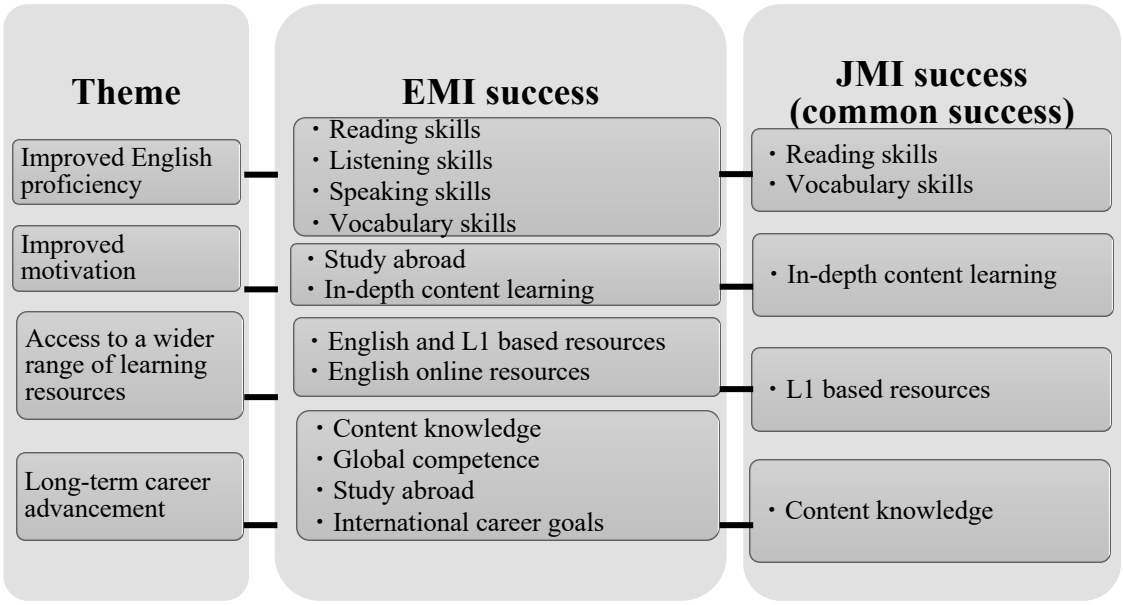
> **Shizuka (JMI):** When I entered the university, I majored archology but I realised that I would also need to learn chemistry and biology to specialise in my area of study. I will need to improve my chemistry knowledge to study at a graduate school.

Hence, the students regarded their EMI and JMI Chemistry courses as platforms to improve their career prospects. However, the EMI students saw the opportunities to interact with their peers and teachers in English as the additional benefits of learning through EMI where they could develop their intercultural competence and English language knowledge. Interestingly, Rina and Lisa (EMI) had more internationally oriented career goals (e.g. studying abroad and working overseas) compared to their JMI counterparts, Kyoka and Shizuka.

4.11.5 Summary

Thus, while the quantitative findings underpinned the positive evidence of EMI content learning—EMI had at least no detrimental impact on the students' content learning—the qualitative evidence revealed both the costs and benefits to content learning through EMI and JMI in more detail. Figure 20 below shows the overview of the students' perceived success depending on their mediums of instruction, highlighting the benefits specific to gaining content knowledge through EMI. The EMI students thus perceived these additional benefits to be weighed up against its deficits, which was a recurring theme in the interviews with the EMI participants.

Figure 20: Success in gaining content knowledge through EMI and JMI



Finally, although it became apparent from the qualitative analysis that the students faced different types of challenges and success depending on their medium of instruction, EMI was also noted to have led to unequal consequences on their content learning depending on individual factors, such as L2 English proficiency and prior-content knowledge. For example, Hikaru and Rina (IELTS 6.5) with lower linguistic and prior content knowledge experienced long-lasting, detrimental problems of content coverage reduction, whereas Miyazaki (IELTS 8.0) with higher linguistic and prior content knowledge encountered relatively few problems adjusting to the EMI learning conventions. Similarly, the magnitude of the challenges associated with understanding the professor's humour and jokes were also shown to be influenced by the students' L2 proficiency level. Recognising the influence of an array of the demographic, individual and contextual factors which all shaped the students' learning experiences, the second research question will explore the differences in implementation of the EMI and JMI class. Finally, the third research question will explore the relationships (if any) between the student demographic variables: L2 proficiency (IELTS scores); EMI challenges (perceived challenges in English-mediated tasks); prior content knowledge (pre-test scores); and post-course content knowledge (post-test scores).

4.12 Chapter Summary

In summary, I outline the key quantitative and qualitative findings in this chapter below in bullet point format for ease and clarity of discussion. The quantitative findings revealed no significant differences between the EMI and JMI students in terms of their content learning achievement (measured by the post-test). The qualitative findings revealed that there were challenges (e.g. poor time management) and success (e.g. improved content knowledge) which were common to both the EMI and JMI students. However, there were also challenges (e.g. inaccurate L2 listening) and success (e.g. accessing English medium resources) which were only distinctive to the EMI students. Thus, while the quantitative findings suggested no difference between the groups on content test scores, the qualitative findings revealed many differences, indicating that while the groups had similar post-test outcomes, the pathways to those outcomes (student challenges and success) differed substantively. The students saw a number of benefits of learning through EMI despite the challenges, suggesting that success was more than just exam scores.

Quantitative findings:

- No significant pre-test scores were found between the EMI and JMI students.
- No significant post-test scores were found between the EMI and JMI students.
- No significant gain scores were found between the EMI and JMI students.

Qualitative findings:

Perceived challenges

- Insufficient comprehension of the course textbook
- Insufficient comprehension of the course lectures
- Low performance on the final course exam and Final Grade
- Difficulty in adjustment to the new academic and cultural conventions

Perceived success

- Improved English proficiency
- Improved motivation
- Access to a wider range of learning resources through EMI
- Long-term career advancement

Chapter 5 – Results: Context-related differences between the two mediums of instruction

5.1 Introduction

The aim of this chapter is to explore a range of context-related differences between the two mediums of instruction which students ascribed to their achievement of content learning. The current study involved quasi-experimental research designs, thereby introducing a huge range of unpredictable contextual factors which may have influenced the extent to which the students developed content knowledge. While the first research question detected no statistically significant differences in the post-test scores between EMI and JMI students, many context-dependent factors may shed further light on explanations of this non-significant result. Given that the present study is classroom-based research, it was not the intent to control for the multitude of contextual differences between the JMI and EMI classes, but rather to bring them to the forefront in the interpretation of the results. All in all, this chapter draws on student and teacher interviews triangulated with field notes from observations to shed light on some of the plausible explanations for why the current research did not detect the post-test score differences between the two mediums of instruction.

5.2 Research question 2

The research question presented below considers the extent to which the students attributed a range of contextual factors to their content learning. The main data set for this research question was the student interviews, although the teacher interviews and the fieldnotes were also analysed to uncover recurring themes across the data sources .

- RQ2: In what ways do the two mediums of instruction differ? What context-related differences do the students ascribe to their achievement of content learning?

To answer this research question, the interview data were coded to capture recurring patterns in the data. In this initial stage of the analysis, several initial themes emerged from the data which represented an array of contextual differences between the two mediums of instruction that may have affected the processes of content learning.

These initial themes were then considered across all of the data sources, whereby some of them were collapsed into main themes and subthemes. Data from the fieldnotes were consulted to add a further dimension and ascertain any contextual similarities and differences between themes identified in the EMI and JMI data. The triangulation of the data sources was crucial to account for the highly contextualised and multifaceted nature of the current classroom-based research. To ease the discussion of findings, the main themes that captured the contextual differences between the EMI and JMI courses are presented in turn drawing on the following four sections: language use (section 5.3); peer learning (section 5.4); support scheme (section 5.5); and educational culture (section 5.6).

5.3 Language use

Analysis of the data revealed that the first contextual difference between the two mediums of instruction was whether students resorted to English or first language (L1) learning resources. Many of the EMI students relied on learning resources which were available not only in English but also in their L1, such as library books, journal articles, online materials and textbooks from high school. However, their JMI counterparts appeared to only refer to learning resources that were available in Japanese. Thus, this section explores the differences in how the students adopted English and L1 learning resources. Table 17 shows examples of how the data were analysed according to the main themes and subthemes.

Table 17: Examples of analytic themes

Data source	Extract from data	Subtheme	Main theme
Student interview	"I use the textbook and notebook from my high school Chemistry class." (**EMI student** Minori)	Use of Japanese in EMI	Language use
	"I do not see the point of reading materials in English. The class is conducted in Japanese." (**JMI student** Nao)	Use of Japanese in JMI	
Teacher interview	"I often encourage my students to use Japanese language materials." (**EMI content teacher**)	Use of Japanese in EMI	
	"Students can select Japanese or English to submit any of the problem-solving worksheets." (**JMI content teacher**)	Use of English in JMI	

Fieldnotes	Key terminologies on lecture slides were offered in Japanese and English. (**Fieldnotes, JMI**)	Use of English in JMI	
	There were no instances where the EMI content teacher translated terms from English into Japanese. (**Fieldnotes, EMI**)	Use of English in EMI	

5.3.1 Language use in EMI

The student interviews laid bare the strong presence of L1 medium learning resources among the EMI students. All L2 English EMI interviewees embraced their L1 as a useful tool to aid content learning and expand the repertoire of available learning materials. In the examples below, Rina reported using textbooks and notebooks from her high school Chemistry lessons to revise basic technical terms in Japanese. Marii also referred to online resources to gain a deeper understanding of academic concepts in Japanese when preparing for her lectures. Despite possessing advanced knowledge in English, Lisa still occasionally used a dictionary and supplemented her learning with Japanese medium resources, such as library books.

> **Rina (EMI)**: I use textbooks and notebooks from my high school Chemistry class. It is helpful to revise basic academic concepts in Japanese (IELTS 7.0, post-test 8.5).

> **Marii (EMI):** I like watching tutorial videos that my old cram school uploads on YouTube. They explain difficult concepts using simple Japanese. I usually watch their videos before attending lectures (IELTS 7.0, post-test 1).

> **Lisa (EMI):** I often visit the university library to borrow Japanese books when I do not understand certain difficult concepts. But I only sometimes use dictionaries to look up unfamiliar words in lectures (IELTS 8.0, post-test 12.5).

The EMI content teacher also encouraged the students to take advantage of their L1 as a useful pedagogical tool. In the excerpt below, he emphasised that it was important for the students to use their L1 to build a solid foundation of content knowledge. He observed that some of his students who had previously studied Chemistry in Japanese at high school used both the English and Japanese versions of the course textbook.

> **EMI teacher:** I do not give lectures in Japanese but I always encourage my students to use both English and Japanese language materials. The library has several copies of a Japanese version of the textbook that we use in the course. It is important that they are able to use their first language to build a strong foundation in the knowledge and understanding of Chemistry. I see some students from

> Japanese medium high schools use both the Japanese and English versions of the course textbook in lectures.

While the EMI content teacher was supportive of the use of L1 medium learning resources, the fieldnotes revealed that all the EMI lectures were conducted entirely in English, and the teacher did not make reference to Japanese medium resources. There were no instances where the EMI teacher translated technical terms from English into Japanese. According to the EMI teacher, the reasons for his exclusive use of English in lectures were twofold: firstly, the EMI cohort consisted of students with different first languages; secondly, those who wished to study the Chemistry course through English were expected to achieve the sufficient level of English proficiency.

In summary, data pointed to the strong presence of L1 medium resources among the EMI students, although use of these materials was relegated to outside of the classroom. While the EMI students were expected to reach a sufficient English language level to adequately understand the EMI lectures, their learning was hugely assisted by the use of their L1. While the EMI teacher did not necessarily use Japanese to conduct lectures, he still acknowledged the benefits of using L1 medium learning resources. Hence, the EMI students unanimously agreed that integrating all of their linguistic resources enabled them to achieve academic success in EMI. It appears that the current learning environment where the students could use their L1 was one of the conditions which may have contributed to successful content learning in EMI. That is, the non-statistically significant differences in academic performance of the EMI and JMI groups was under the condition that the EMI group utilised their L1 to support their EMI studies.

5.3.2 Language use in JMI

Compared to the EMI students who utilised both English and Japanese medium resources, none of their JMI counterparts reported using English medium resources. In the examples below, Nao and Oda showed little to no interest in using English medium resources. Nao described that her low English proficiency was the very reason that she had decided against enrolling in the EMI course. Oda also noted that his low motivation to improve his English skills may have influenced his decision to select the JMI course.

> **Nao (JMI):** I do not see the point of reading materials in English. The class is conducted in Japanese. I would do much worse if I had taken the English class because of my poor English knowledge (post-test 3).
>
> **Oda (JMI):** I want to improve my content knowledge in my first year so it is more efficient to study in Japanese. Right now I am not interested in improving English (post-test 15).

Despite Nao and Oda's negative stance towards the use of English, the JMI content teacher was supportive of using English in the JMI lectures, claiming that English had become an academic lingua franca in many of the STEM disciplines. He encouraged his students to memorise technical terms in both Japanese and English. He also allowed them to submit any of their problem-solving worksheets in Japanese or English. This discrepancy in attitudes towards the use of English between the JMI teacher and students was evident in the excerpts below.

> **JMI teacher:** I would like my students to learn technical terms in English and Japanese. Being able to read journal articles in English has become more and more important in many of the science disciplines. They can also submit any of their problem-solving worksheets in Japanese or English.
>
> **Shizuka (JMI):** My professor introduces technical words in English during lectures but I do not think I learn any of the English words because there are already too many difficult terms in Japanese (post-test 8).
>
> **Kyoka (JMI):** My reading list contains English books and materials but I have never actually read any of these English ones. I take a long time to read in English (post-test 16).

To support the JMI teacher, the fieldnotes revealed that many of the JMI lecture slides were presented bilingually, introducing technical terms in both English and Japanese. A list of suggested reading and resources was also available to the students in both English and Japanese. Given that many of the JMI students rarely used any of the English medium learning resources, there was a gap between how the language of instruction was envisaged by the teacher and how it was actually being practised by the students.

5.3.3 Summary - Language use

Overall, this section has discussed the strong presence of L1 medium resources among the EMI students as one of the contextual differences between the EMI and JMI courses. According to the EMI students and teacher, the use of an L1 was a useful pedagogical tool associated with better outcomes in content mastery. The EMI

students turned to both Japanese and English language materials to alleviate their linguistic challenges. In contrast, their JMI counterparts reported little to no interest in using English medium resources. Despite reference to English terminology in their classes by their JMI teacher, these students exclusively relied on Japanese medium resources for their learning.

5.4 Peer learning

Another theme which was largely absent in the JMI data, but reported by every EMI student was interactions that occurred between students to help each other learn content collaboratively. Students attributed this peer learning as a crucial factor supporting their academic success. This section therefore explores the differences in how the students resorted to peer learning according to the mediums of instruction. Table 18 below shows examples of how the data were analysed according to the main themes and subthemes.

Table 18: Examples of analytic themes

Data source	Extract from data	Subtheme	Main theme
Student interview	"I ask my course mates to clarify what I don't understand from lectures." (**EMI student** Hikaru)	Presence of peer learning in EMI	Peer learning
	"I rarely need help from my classmates." (**JMI student** Oda)	Absence of peer learning in JMI	
Teacher interview	"Chemistry major students should build good relationships with their colleagues in their first year." (**EMI content teacher**)	Presence of peer learning in EMI	
	"In the introductory level courses, it is often difficult for students to interact with each other." (**JMI content teacher**)	Absence of peer learning in JMI	

5.4.1 Peer learning in EMI

An analysis of the student interviews revealed that many of the EMI students valued peer learning as a useful pedagogical tool to ensure content mastery. Examples of how the students benefited from peer interactions with other class members are offered below.

> **Charlotta (EMI):** My friend helps me when we do homework problem sheets. She is very good at explaining difficult concepts clearly (IELTS 8.0, post-test 9.5).

> **Miyazaki (EMI):** I often sit next to my friend in lectures and we study together afterwards. I understand most of the concepts covered in my lectures but it is still helpful for me to clarify my understanding with her (IELTS 8.0, post-test 15).

In these examples, although Miyazaki had a higher level of content knowledge than Charlotta, they still found it mutually beneficial to help each other. While Miyazaki played a more dominant role in aiding in Charlotta's academic development, she still enriched and expanded her own understanding through explaining difficult concepts with clarity and ease. Their collaborative learning experience showed that peer learning was a useful pedagogical tool in the current mixed-ability learning environment in which the learners did not necessarily possess the similar level of content knowledge.

In the teacher interview data, the EMI content teacher also noted that the students should establish a rapport with their peers during the crucial first year at university. In the excerpt below, he went on to argue that there were clear long-term benefits associated with peer collaboration for Chemistry degree students.

> **EMI teacher:** Chemistry major students should build good relationships with their course mates in their first year. They need to work collaboratively when they advance to junior and senior year courses where they work on lab based projects.

In this assertion, the EMI content teacher emphasised the importance of peer learning. He claimed that Chemistry degree students should turn to their peers for support when they progress to more advanced level courses where laboratory based practicals would become part of the regular curriculum and assessment.

To mirror the EMI teacher's view, Ayaka and Rina below shared their positive experience of working closely with their colleagues who ultimately became their close friends. They attributed supportive peer networks to their long-term academic success.

> **Ayaka (EMI):** I always do my homework with my classmates after lectures. I have made good friends with some of them because we spend a lot of time together. I think having close friends from the course has helped me make it through some of my difficult challenges at university (IELTS 7.0, post-test 8.5).

> **Rina (EMI):** Our Chemistry department is very small so many of us will take the same courses in our third and fourth year. I am trying to be friends with as many people as possible this year (IELTS 6.5, post-test 7.5).

Hence, peer learning appeared to be an effective way to enable the students to improve their content understanding and achieve successful and satisfying university careers. This peer learning may have, therefore, contributed to the conditions that underpinned success in content learning for the EMI students.

5.4.2 Peer learning in JMI

With regard to the JMI interview data, it became clear that none of the JMI students reported the presence of peer learning. In the examples below, Nao and Oda admitted the sheer absence of peer learning. Oda in particular explained that he was still able to achieve a satisfactory understanding of academic concepts without support from his course members.

> **Nao (JMI):** I have not made any friends in this course because too many students are taking this course. I just go to the lecture hall, sit through one hour of my lectures and leave as soon as the lectures are finished (post-test 9.5).

> **Oda (JMI):** I usually sit alone in the lecture theatre because the room is too spacious. I also understand almost all concepts covered in the course so I rarely need help from other students (post-test 9.5).

When examining these responses more deeply, it became clear that one of the possible explanations for the perceived absence of peer learning in the JMI course was the large class size of 31 students as opposed to the EMI class of 18 students. The JMI teacher also highlighted that the set-up of the large class size was a key reason for the lack of peer learning in the JMI course. He noted below that a combination of the large class size and teacher-centred lecture style was not conducive to collaborative study.

> **JMI teacher:** In the introductory level courses, it is often difficult for students to interact with each other because we usually give formal lectures. This works well when we teach a large group of students in a spacious lecture hall.

Thus in conclusion, the large class size coupled with the JMI teacher's tightly constructed lecture style both seemed to have contributed to the absence of peer learning in the JMI course.

5.4.3 Summary – Peer learning

This section has discussed the strong presence of peer learning in the EMI course as one of the contextual differences between the two mediums of instruction. Unlike the JMI students, many of the EMI students enjoyed a range of benefits of peer learning, such as improving content knowledge, establishing close friendships, and creating long term support networks. These benefits significantly improved the speed and ease of the transition into EMI education. It also became clear that a mixed ability learning environment offered the EMI students opportunities to learn to communicate and breakdown complex concepts in a way which was easier for other students to understand. Overall, peer learning was highlighted by students as vital to their academic success in the EMI course.

5.5 Support schemes

One theme that was largely evident in both the EMI and JMI data concerns learning support schemes, which may be another key factor contributing to the students' academic success. Analysis of data revealed that a range of support provisions were put into place in the EMI and JMI courses, but there were also many contextual differences in how the support was integrated into these courses, some of which were highly relevant to content learning. This section therefore explores contextual differences in the way in which learning support schemes were implemented in the two mediums of instruction. The theme of "support schemes" was divided into three main themes for more detailed analysis: teaching assistant provision; the Comment Worksheet; and the academic English language program.

5.5.1 Support scheme: teaching assistant provision

Analysis of data revealed that the provision of a teaching assistant was a theme that was absent in the EMI data, but reported by many of the JMI students as a crucial factor influencing their academic success. Table 19 below shows examples of how the data were analysed according to the main themes and subthemes.

Table 19: Overview of analytic categories

Data source	Extract from data	Subtheme	Main theme
Student interview	"I would like to ask him for help but I do not know how I can contact him." (**EMI student**, Saika)	Absence of the teaching assistant in EMI	Teaching assistant provision
	"I have recently asked the teaching assistant for advice on how to revise for the final exam." (**JMI student**, Nao)	Presence of the teaching assistant in JMI	
Teacher interview	"He can't really help my students because he does not speak much English and he only attends the Japanese language lectures." (**EMI content teacher**)	Absence of the teaching assistant in EMI	
	"Students should utilise the teaching assistant because he can give them both academic and personal advice." (**JMI content teacher**)	Presence of the teaching assistant in JMI	
Fieldnotes	There were no instances where the teaching assistant was present in any of the EMI lectures. (**Fieldnotes, EMI**)	Absence of the teaching assistant in EMI	
	The teaching assistant was present at all times throughout the entire semester. (**Fieldnotes, JMI**)	Presence of the teaching assistant in JMI	

When exploring the JMI student interviews, it became clear that many of the JMI students valued the additional support given by the teaching assistant. In the examples below, Nao and Oda explained that a teaching assistant, a postgraduate research student, was available through the whole semester to offer them in-person support, assist them with their enquiries, and provide them with individually tailored advice. Examples of how Nao and Oda benefitted from this support provision are presented below.

> **Nao (JMI):** I have recently asked the teaching assistant for advice on how to revise for the final exam. It is helpful to turn to him for help because I sometimes feel too embarrassed to ask my professor basic questions (post-test 3).

> **Oda (JMI):** I made an appointment to speak to the teaching assistant to learn more about practical advice on applying for a graduate school. I would like to study Chemistry at a graduate school when I finish my undergraduate degree (post-test 15).

Here, Nao who struggled to prepare for the final exam received advice on revision techniques and exam preparation from the teaching assistance. Similarly, Oda

received pastoral care, mentoring and practical advice on careers guidance. The fieldnotes also revealed that the teaching assistant was present at all times throughout the entire semester. The JMI students could easily interact with the teaching assistant who was seated at the back of the lecture theatre, which was located next to its entrance.

In the teacher interview, the JMI content teacher stated he encouraged the students to seek support from the teaching assistant, explaining how this provision was useful for them. In the excerpt below, he argued that he was also able to benefit from this support scheme as a professor. He noted that this was a university-wide scheme designed to allow faculty members across all academic disciplines to recruit graduate teaching assistants who could support their teaching.

> **JMI teacher:** I have asked one of my former students to come and work for us as a graduate teaching assistant. Students should utilise this support system because he can offer academic and careers advice. The university allows us to hire teaching assistants to support with our teaching duties. It is helpful to have him because he can mark student homework and support me with technical aspects of the course.

Therefore, this support provision assisted not only the students but also the content teacher. The teaching assistant was able to respond to the students' questions and comments on behalf of the JMI content teacher. Furthermore, he was also involved in the assessment process and the revision of teaching materials to support the development of the course curriculum.

Conversely, none of the EMI students resorted to the teaching assistant for help. For example, neither Hikaru nor Saika had ever contacted the teaching assistant. In the comments below, Hikaru noted that he had never seen the teaching assistant in any of the lectures. Saika was not aware of how to make an initial contact with the teaching assistant. These comments reinforce the absence of this support provision in EMI.

> **Hikaru (EMI):** I remember that my professor mentioned about the teaching assistant but I have never actually seen him in any of the lectures (IELTS 6.5, post-test 5).

> **Saika (EMI):** I would like to ask the teaching assistant for help but I do not know how I can contact him in the first instance (L1 English, bilingual, post-test 7.5).

The EMI teacher also suggested that the EMI students faced practical difficulties in accessing this support provision. He emphasised that this scheme was, on the face of it, built into both the JMI and EMI courses; in reality it was only accessible to the JMI students. He went on to argue that the teaching assistant was not present in any of the EMI lectures and obtained little information about the progress of the EMI course. This was largely due to his limited command of the English language, which made it difficult for the students to turn to him for support in English.

> **EMI teacher:** The teacher in the Japanese course recruited the teaching assistant who was one of his former students. Unfortunately, he can't really help my students because he does not speak much English and only attends the Japanese lectures.

The fieldnotes also recorded no instances where the teaching assistant was present in any of the EMI lectures. This was in stark contrast to the JMI course in which the students could easily interact with the teaching assistant who was available throughout the semester, and physically present in their classes. Thus, this structural difference in an arrangement of the support scheme between the two mediums of instruction meant that the EMI students could not derive much support from the teaching assistant.

Overall, the teaching assistant supported both the students and content teacher in the JMI course. For example, the students turned to the teaching assistant for academic and careers advice. The teacher also received support from the teaching assistant to develop the course curriculum. Conversely, the EMI students were unable to access the teaching assistant provision due to various structural constraints. He had a limited command of the English language and little information about the progress of the EMI course. Thus, this support scheme was made only available to the JMI course. Thus, the EMI students success in EMI occurred despite structural disadvantages in the different conditions of their learning compared to the JMI learning condition.

5.5.2 Support scheme: The Comment Worksheet

The Comment Worksheet was another theme in the data associated with learning support schemes. Unlike the teaching assistant, mentions of the Comment Worksheet were evident in both the EMI and JMI data. Table 19 shows examples of how the data were analysed according to the main themes and subthemes.

Table 20: Overview of analytic categories

Data source	Extract from data	Subtheme	Main theme
Student interview	"I can use the Comment Worksheet to share my questions with my professor." (**EMI student**, Hikaru)	Usefulness of the Comment Worksheet in EMI	The Comment Worksheet
	"I like the Comment Worksheet to communicate with my professor more privately." (**JMI student**, Shizuka)	Usefulness of the Comment Worksheet in JMI	
Teacher interview	"The Comment Worksheet is an effective way to communicate with my students." (**EMI content teacher**)	Usefulness of the Comment Worksheet in EMI	
	"The Comment Worksheet is a good system that I can use to communicate with my students." (**JMI content teacher**)	Usefulness of the Comment Worksheet in JMI	
Fieldnotes	The EMI content teacher distributed the Comment Worksheet at the end of every lecture. (**Fieldnotes, EMI**)	Frequency of the Comment Worksheet in EMI	
	The JMI content teacher did not distribute or address the Comment Worksheet in every lecture. (**Fieldnotes, JMI**)	Frequency of the Comment Worksheet in JMI	

The EMI student interviews revealed that the Comment Worksheet was another source of learning support available to the EMI students. The EMI content teacher regularly used the Comment Worksheet to assess the students' academic progress, learning needs, and ask for general questions and comments. In the following examples, Hikaru and Saika noted that the Comment Worksheet offered them opportunities to express their concerns and issues in writing and discuss them with their content teacher. They also illustrated how this support scheme enhanced their overall learning experience.

> **Hikaru (EMI):** I am often too shy to ask questions in front of my classmates. But I can use the Comment Worksheet to share my questions with my professor. This helps me clarify difficult academic concepts (IELTS 6.5, post-test 5).

> **Saika (EMI):** I feel more comfortable when I can ask questions using the Comment Worksheet. I appreciate our teacher taking a lot of time to answer our questions at the start of every lectures. I like how he responds to our questions. It helps me understand academic concepts better (L1 English, bilingual, post-test 7.5).

In the excerpt below, the EMI content teacher emphasised the importance of student-teacher communication, explaining that the Comment Worksheet was an effective way to establish a rapport with his students, especially with those who were shy and quiet.

> **EMI teacher:** I distribute the Comment Worksheet in every lecture. There are some students who have good answers but do not have the courage to speak up in front of the whole class. So it is an effective way to communicate with my students.

The fieldnotes also recorded that the EMI content teacher distributed the Comment Worksheet at the end of every lecture and addressed the students' comments and questions at the start of the following lectures. Thus, he regularly used the Comment Worksheet to support their learning, and to build a communicative relationship with his students.

Similarly, the JMI student interviews suggested that Nao and Shizuka also relied on the Comment Worksheet to seek advice from their content teacher in the JMI course.

> **Nao (JMI):** I use the Comment Worksheet to share my questions with my professor. He does not necessarily address my comments in every lecture but I still like it when I receive his feedback. It is a useful way to understand difficult concepts (post-test 3).

> **Shizuka (JMI):** I often feel too shy to speak to my professor directly so I like the Comment Worksheet to communicate with him more privately (post-test 8).

However, the JMI content teacher noted that he did not necessarily adopt the Comment Worksheet in every lecture. While he still admitted that the Comment Worksheet was a useful way to build a good relationship with his students, he claimed that addressing their comments and questions during every lecture would be too time consuming.

> **JMI teacher:** The Comment Worksheet is a good system that I can use to communicate with my students. But I do not address their comments in every lecture because it takes too much time. There are a lot of new topics that we need to cover in the introductory level lecture.

The fieldnotes also revealed that the JMI content teacher did not distribute or address the Comment Worksheet in every lecture, which stood in contrast to the EMI content teacher. Given that the Comment Worksheet offered both the EMI students and teacher many benefits (e.g. improving content understanding, and building a good

relationship between the content teacher and students), the differences in how this support scheme was integrated into the two courses could be another key factor that contributed to the successful development of content learning.

5.5.3 Support scheme: The academic English language program

An analysis of data suggested that the academic English language program was a theme which was absent in the JMI data, but reported by many of the EMI student as a crucial factor determining their academic success. Table 20 below shows examples of how the data were analysed according to the main themes and subthemes.

Table 21: Overview of analytic categories

Data source	Extract from data	Subtheme	Main theme
Student interview	"I find it helpful to learn some useful tips on note-taking in the academic English language classes." (**EMI student**, Marii)	Importance of the academic English language program in EMI	The academic English language program
Teacher interview	"It is important for the students to attend the academic English language program to improve their academic English skills." (**EMI content teacher**)	Importance of the academic English language program in EMI	

The EMI student interviews revealed that many of the EMI students attributed the academic English language program with their improvement of academic English language skills required to undertake their regular content courses through English. The university offered its first year students across all academic disciplines a range of in-sessional academic English courses organised concurrently alongside the mainstream content courses. In the examples below, Ayaka and Minori noted that these academic English courses were crucial for enhancing their preparedness for EMI content learning.

> **Ayaka (EMI):** I take a long time to read the Chemistry textbook in English so it has been helpful to take the English language program this semester to learn how to read English language textbooks more efficiently (IELTS 7.0, post-test 8.5).

> **Marii (EMI):** I struggle to take notes when I attend lectures in English. I always seem to easily doze off and miss an important point. It is difficult to pin down the crucial points to write down during lectures. So I find it helpful to learn some useful tips on note-taking in the academic English language classes (IELTS 7.0, post-test 1).

Exploring their responses in greater depth revealed that these courses emphasised the improvement of general academic English skills, rather than disciplinary specific language skills catering for the learning of Chemistry. Nevertheless, Ayaka and Minori recognised that the skillsets they acquired from these courses were still transferrable to the learning of Chemistry. Ayaka learned to read English medium textbooks efficiently and Minori learned to take notes in English effectively, which was key to EMI content learning.

Similarly, Hikaru and Rina who had lower English proficiency than Ayaka and Minori appeared to have derived more benefits from the academic English language program. They asserted that it instilled greater confidence in using English in academic contexts and facilitated a smooth transition between their English language courses and regular content courses. They could mitigate their English language-related challenges to better prepare for their EMI content learning.

> **Hikaru (EMI):** I like reading authentic textbooks in the academic English language program because it usually takes a long time to read in English. It helps me feel more confident in taking more English medium content courses (IELTS 6.5, post-test 5).
>
> **Rina (EMI):** I have only studied English as a school subject before entering university so it is important for me to improve my academic English speaking skills. I am currently taking an academic English speaking class to practice speaking English as much as possible (IELTS 6.5, post-test 7.5).

In the teacher interview, the EMI content teacher also recognised the usefulness of the academic English language program, suggesting that the students were able to improve their productive English language skills. He stated that those who received schooling in Japan were proficient in their reading and listening skills but less so in their speaking and writing skills, which were still part and parcel of EMI study.

> **EMI teacher**: I usually advise my students that they should not take the English medium course if they do not have the sufficient English proficiency. I do not simplify my classroom language or translate technical terms from English into Japanese. So it is important for them to attend the academic English language program to improve their English skills. I think that they mainly need to improve their productive skills when they attend the program. Many of my students attended Japanese high schools so they can read and listen in English but they often lack confidence in speaking and writing in English.

Here, the EMI content teacher also asserted that he did not simplify his classroom language or translate terms into Japanese to accommodate those with the inadequate

English proficiency. The fieldnotes recorded that the EMI content teacher advised students at the beginning of the semester that those without the sufficient English proficiency were encouraged to select the JMI course. Thus, the academic English program was essential especially for those who needed to improve their academic English language skills to comprehend their EMI Chemistry lectures where English language support from the content teacher was not necessarily available.

Overall, while the academic English language program was not specifically catered for the preparation of the Chemistry course, it still contributed to the overall readiness for EMI study. For example, Ayaka could learn to read English medium textbooks more quickly and Marii could learn to take notes in English more effectively, which were still important transferable skillsets required to study their regular EMI content courses across all academic disciplines. Furthermore, given that the EMI students received little linguistic support from the content teacher, the academic English language program was vital to those who needed to improve their academic English language skills to perform satisfactorily in the EMI course.

5.5.4 Summary – Support schemes

In conclusion, this section has discussed the similarities and differences in terms of how the following three types of support schemes were put into place for the EMI and JMI courses. Firstly, the teaching assistant provision was made only available to the JMI students due to a range of practical constraints. For example, the teaching assistant had limited English knowledge and was completely absent in the EMI course. Secondly, the Comment Worksheet was introduced in both the EMI and JMI courses as an effective means to support the students. However, the frequency of its use was left open for each teacher and was not consistent across the two courses. Finally, the academic English language program was a support scheme for the EMI students and was particularly useful for students with lower English proficiency. They advocated the program for not only the direct improvement of their academic English skills but also their overall preparedness for EMI education, such as boosting their confidence in using English in an academic context. Thus, the EMI and JMI courses shared differences and similarities in how learning support provision was integrated into their curriculum, some of which were only available to one of the courses (e.g.

the teaching assistant provision) while others were equally available to both of them (e.g. the Comment Worksheet). Therefore, these contextual differences between the two courses may have influenced how successfully the students developed content knowledge.

5.6 Educational culture

A further distinction between the two mediums of instruction related to educational culture, which was another significant factor contributing to students' academic success. An analysis of data suggested that many of the EMI students perceived that their educational culture was interactive and participatory whereas their JMI counterparts retained the teacher-centred educational culture, which resembled that of their Japanese medium high schools. This section explores how the differences in educational culture between the two mediums of instruction affected their achievement of content learning. Table 21 shows examples of how the data were analysed according to the main themes and subthemes below.

Table 22: Examples of analytic themes

Data source	Extract from data	Subtheme	Main theme
Student interview	"The professor often poses questions and asks us to come up to the front of the class to share ideas with the rest of the class." (**EMI student**, Minori)	Interactive educational culture in EMI	Educational culture
	"I like how the professor organises lectures because they are teacher-centred." (**JMI student**, Shizuka)	Teacher-centred educational culture in JMI	
Teacher interview	"I nominate my students to share their answers with the rest of the class." (**EMI content teacher**)	Interactive educational culture in EMI	
	"I usually deliver lectures in a traditional way, using a blackboard and chalk." (**JMI content teacher**)	Teacher-centred educational culture in JMI	
Fieldnotes	Students came up to the front of the lecture theatre and explained key concepts to the rest of their course members using a blackboard and chalk. (**Fieldnotes, EMI**)	Interactive educational culture in EMI	
	There were no instances where interaction occurred between the content teacher and students. (**Fieldnotes, JMI**)	Teacher-centred educational culture in JMI	

5.6.1 Interactive educational culture in EMI

An analysis of the student interviews revealed that EMI lectures typically involved a range of student-teacher interaction to promote an exchange of ideas and discussion of key academic concepts. The teacher delivered a pre-planned, tightly formatted presentation, but also frequently paused between different class activities to encourage students to share their ideas and engage in discussions. Many of the EMI students valued this interactive educational culture as a useful way to enliven and clarify difficult academic concepts. Examples of how the EMI students benefited from this interactive style of learning are offered below.

> **Ayaka (EMI):** I have more opportunities to interact with my professor at university compared to high school. Our professor remembers our names and often nominates us to share our ideas. Because I can interact with my professor during lectures, I feel more confident in visiting his office and asking him questions (IELTS 7.0, post-test 8.5).

> **Minori (EMI):** Our professor often poses questions and asks us to come up to the front of the class to share our ideas with the rest of the class. It is not what I am used to from my high school education but I find it useful to explain difficult concepts using my own words (IELTS 7.0, post-test 9.5).

In these examples, Ayaka and Minori who had not experienced EMI education prior to university, compared their current educational culture with that of their high school. Ayaka suggested that she gained more confidence in reaching out to her professor for help. As a result, she often scheduled appointments to discuss assignments and materials which she was particularly struggling with. Similarly, Minori could organise her thoughts by explaining her ideas with her course members and teacher during her lectures. To support Ayaka's and Minori's views, the fieldnotes also revealed that the EMI content teacher nominated the students to come up to the front of the lecture hall and asked them to explain key concepts to the rest of their course members using a blackboard and chalk in every lecture. Other students could listen to their explanation and then pose follow-up questions, which generated further student-centred discussions.

In the teacher interview data, the EMI content teacher also emphasised the importance of student participation. In the example below, he suggested a range of techniques that he would typically employ to promote greater student participation, such as learning about his students, memorising their names, recognising those with low English

proficiency, learning where they sat in the lecture hall, offering them individually tailored feedback, and encouraging them to share ideas with the other course members. Here, the EMI teacher spared no effort to encourage greater student-teacher interaction.

> **EMI teacher:** I nominate my students to share their answers with the rest of the class rather than wait for volunteers. I try and remember their names and where they usually sit in the classroom. I also try and work out who is struggling and who is doing well, and who needs extra language support. It is important to give them individually tailored feedback and do my best to support them to succeed.

Although the EMI teacher created an interactive educational culture, some students without sufficient English proficiency or prior EMI learning experience viewed this interactive educational culture as a double-edged sword: on the one hand, it improved their motivation to actively participate in lectures; on the other hand, it caused anxiety if they were not confident in speaking English in front of other students. Examples of these two contrasting views are highlighted by two EMI students below.

> **Hikaru (EMI):** I really like how our teacher nominates us to share our ideas with the rest of the class. I feel like I am contributing to the class. But I also find it difficult to express my opinions well in English. I feel nervous about speaking English in front of my classmates (IELTS 6.5, post-test 5).

> **Rina (EMI):** Some of my course mates have very good English and ask many questions to challenge the teacher. I would like to do the same but I am worried about my English (IELTS 6.5, post-test 7.5).

Thus, the interactive educational culture generated plenty of opportunities for many of the EMI students to share their opinions with their teacher and course members and receive individually tailored feedback. However, those without sufficient English proficiency or prior EMI learning experience could find it difficult to adapt to this educational culture because of their fear of speaking English in front of their course members.

5.6.2 Teacher-centred educational culture in JMI

Based on the JMI student interviews, all students perceived that their educational culture closely resembled how they had been taught at Japanese medium schools. In the excerpts below, Shizuka and Oda described their learning as “teacher-centred” and “passive”, respectively. While they were accustomed to the teacher-centred educational culture through the Japanese school system, they also expressed issues

inherent in this educational culture. For example, Oda noted that he usually sat through lectures passively listening to his professor. He was often not motivated to do his pre-lecture reading homework as he knew that all key concepts would be covered in lectures. Shizuka noted that many of her course members did not usually answer questions posed by the teacher, which often lapsed into uncomfortable and awkward silences.

> **Oda (JMI):** Our professor does not usually check if we have read the textbook. I only read the most important part of the textbook usually after attending lectures. This passive learning works well for me because it is similar to how I was taught at high school (post-test 15).

> **Shizuka (JMI):** I like how the professor organises teacher-centred lectures because I can first listen and then digest what is being covered. I learn well this way. It is similar to how I learned at high school. The professor sometimes asks us questions but no one gives answers because everyone is too shy so it is a bit awkward (post-test 8).

Here, Oda and Shizuka suggested the absence of student-teacher interaction, pointing to the strong reliance on the more traditional and formal teaching method in the JMI course. The fieldnotes also revealed that there were no instances where interaction occurred between the content teacher and students in any of the JMI lectures.

Similarly, the JMI content teacher supported the teacher-centred educational culture, offering two reasons: firstly, the course involved delivering lectures to a group of up to 31 students in a large lecture hall; secondly, the introductory level course aimed at the mass transmission of a large volume of disciplinary knowledge from teachers to students through the medium of tightly constructed, well-organised lectures. Below the JMI teacher explained that he typically offered "chalk and talk" lecture in which he mainly lectured using a blackboard and chalk while students generated notes.

> **JMI teacher:** I usually deliver lectures in a traditional way, using a blackboard and chalk because I have to teach a large group of students. There is also a lot of content to cover in these foundation level courses.

Thus, Oda and Shizuka largely subscribed to the teacher-centred educational culture which mirrored how they had been educated at Japanese medium high schools. From the practical point of view, the JMI teacher also embraced the teacher-centred educational culture in the introductory course.

5.6.3 Summary – Educational culture

To summarise, this section has discussed educational culture as one of the contextual differences between the two mediums of instruction, which may have influenced how the students learned Chemistry. Many of the EMI students attributed the interactive educational culture to their greater academic achievement. This educational culture offered a range of benefits to aid learning of the academic concepts. For example, the students were given additional opportunities to interact with their content teacher, clarify issues with their course members, gain more confidence in seeking help, and organise their thoughts by explaining concepts using their own words. While those who had received Japanese medium high school education faced a considerably more testing transition into this interactive educational culture, they still enjoyed these benefits. Conversely, the JMI classes retained a more traditional, teacher-centred learning culture which followed that of the Japanese medium high schools. They had little or no chance to engage with their content teacher, which some of them described as "passive" learning. Overall, given the myriad of additional benefits unique to EMI, the differences in educational culture between the two mediums of instruction could be another crucial factor that contributed to the students' academic performance in EMI.

5.7 Chapter summary

The current classroom-based study involved research with real-world learning environments, with in-tact JMI and EMI classes. As a result, there were several differences in the two learning conditions, which may have influenced the students' content learning, and need to be brought to bear in interpretation of the study's findings in the previous two chapters. Accordingly, this chapter probed a range of contextual differences between the two mediums of instruction which were key to the development of content knowledge based on the students' and teachers' perceptions.

To create a detailed picture of the students' learning experience, student and teacher interviews were corroborated with fieldnotes. As shown in the chapter, the triangulation of the data sources played a vital role in achieving nuanced contextual data to improve the transparency of the inevitably messy research context in which a range of contextual differences existed between the two courses.

Thus, the findings of the previous chapters that EMI students did not significantly underperform in their Chemistry classes compared to the JMI students is underpinned by the following contextual conditions:

- The EMI students had access to supplementary learning support in the form of an academic English program which helped them to develop their academic English and study skills.
- The EMI class utilised more frequent learning support from a learning scheme called the Comment Sheet, where they could ask questions to their lecturer.
- The EMI class utilised L1 learning resources outside of the EMI class to support their learning.
- The EMI class engaged in peer learning, in stark contrast to the JMI class
- The EMI class did not have access to a teaching assistant, whereas the JMI class had weekly access to one.
- The EMI classes were interactive in style, whereas the JMI classes were more teacher-centred.

Hence, due to an array of complex interactions between contextual, individual and external factors, I emphasise the difficulties in teasing out a single powerful factor which singularly accounted for the development of content learning. Nevertheless, while acknowledging the messiness of research involving intact classrooms, this study successfully shed light on some of the important contextual factors which the students ascribed to their achievement of content learning in the EMI and JMI courses.

Chapter 6 - Results: Prior content knowledge, academic English challenges, proficiency and chemistry content learning in EMI

6.1 Introduction

Chapters four and five explored the potential differences between the two mediums of instruction in terms of 1) student content learning achievement and 2) course implementation, respectively. Chapter six focuses on the EMI participants, reporting on the quantitative and qualitative findings concerning the relationship between their prior content knowledge (i.e. pre-test scores), English language proficiency (i.e. IELTS scores), perceptions of ease towards academic study through English (i.e. the EMI Challenges scale) and content knowledge achievement (i.e. post-test scores). The quantitative evidence will be first presented drawing on descriptive and inferential statistics of the relationship between the quantitative variables, whereby proficiency, challenges, and prior content knowledge may affect the achievement of EMI content knowledge. Subsequently, the qualitative evidence involving 13 semi-structured student interviews will be presented and discussed to elicit a more nuanced understanding of how these factors may interact and influence student perceptions of difficulties and ease towards academic study in EMI contexts.

6.2 Research question 3.1

The first sub-research question presented below examines the extent to which the EMI students' achievement in content learning (as measured by chemistry test scores) was associated with English language proficiency, prior content knowledge, and perceived EMI challenges associated with different tasks oriented around academic English skill areas. Findings are first based on descriptive statistics pertaining to the differences in the proportion of high and low performing students according to proficiency and challenges. Subsequently, correlation analyses will be performed to explore the relationship between the variables, followed by regression analyses to probe the extent to which the predictor variables (i.e. proficiency, prior content knowledge, and challenges) may explain the variance in the student' content learning achievement (i.e. post-test scores) as an outcome variable.

- RQ3: Are there relationships between the EMI students' English language proficiency, academic English challenges, prior content knowledge, and content knowledge achievement?
    - RQ 3.1: To what extent are factors such as the EMI students' English language proficiency and perceived EMI challenges associated with content knowledge achievement? (Directly measured relationship)

6.3 Descriptive statistics: Proficiency, challenges, pre-test scores and post-test scores

6.3.1 Presenting descriptive statistics of key variables

Before exploring the relationships between different variables, this section introduces descriptive statistics of the key variables. Descriptive statistics (n=27) are shown in Table 22 for the EMI students' post-test scores, EMI Challenges Scale, English language proficiency, and pre-test scores (see Appendix 26 for histograms). The mean post-test score was 10.40 (SD=3.15), with a median score of 12.0. The lowest score was 1.0 and the highest was 15.0 (out of 17.0) achieving a range of 14.0. The skewness was -0.48 (falling within -0.5 to +0.5) and kurtosis was -0.44 (falling within -3 to +3), an acceptable range (Briggs Baffoe-Djan & Smith, 2020) with no outliers. Accordingly, the data was approximately normally distributed (see Appendix 25 for a z-score table). Secondly, the mean score of EMI challenges score was 4.10 (SD=0.54, range=2.2), with a median of 4.2, and a range from 2.8 to 5.0. Skewness (-0.41) and kurtosis (-0.77) fell within the acceptable range of +1 to -1. This variable therefore was neither significantly skewed or kurtotic. Thirdly, IELTS band score as an indication of English proficiency (Mean=7.83, SD=0.77) had the highest value of 9.0 and the lowest value of 6.5, with a range of 2.5, and a median of 8.0. The data was also approximately normally distributed (skewness=0.2, kurtosis=-0.73) with no outliers. Finally, in contrast to these other variables, the distribution of pre-test scores (Mean=1.80, SD=3.21) was positively skewed (skewness=2.63, kurtosis=5.17) with a large proportion of students being clustered around the lower end of the distribution (closer to 0 score). This means that the majority of the students in the sample had little to no prior content knowledge at the start of the semester. The lowest score was 0.0 and the highest score was 13.5, with a median value of 0.5. Consequently, the violation of the normality assumption restricted the selection of the statistical tests

planned for the later analyses (Field, 2018). Subsequently, I adopted a Spearman's rank-order correlation as a nonparametric alternative to a Pearson's correlation. The section below will draw on bivariate descriptive analysis to explore how the post-test scores varied according to the levels of English proficiency and academic challenges.

Table 23: Descriptive statistics of proficiency, challenges, pre-test, and post-test scores

Variable	N	Mean	SD	Median	Min	Max	Range	Skew	Kurtosis
Proficiency	27	7.83	0.77	8.0	6.5	9.0	2.5	0.20	-0.73
EMI Challenges	27	4.10	0.54	4.2	2.8	5.0	2.2	-0.41	-0.77
Pre-test scores	27	1.80	3.21	0.5	0.0	13.5	13.5	2.63	5.17
Post-test scores	27	10.40	3.15	12.0	1.0	15.0	14.0	-0.48	-0.44

6.3.2 Bivariate descriptive analysis: content knowledge achievement and proficiency

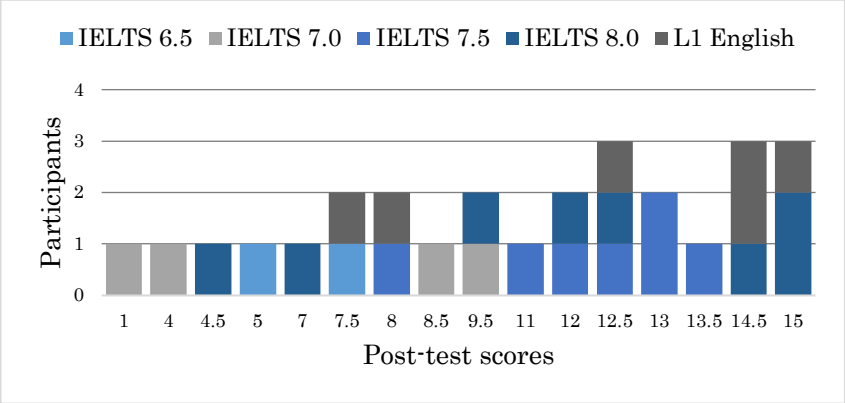
This section explores the associations between the key variables in relation to content knowledge achievement. Firstly, descriptive statistics in Table 23 below were calculated to explore how the EMI students' post-test scores (n=27) varied according to each level of English language proficiency (i.e. IELTS 6.5; 7.0; 7.5; 8.0; and L1 English). Due to the small sample size consisting of each proficiency group, the median and range were computed as a measure of central tendency in addition to the mean and standard deviation for each group. The descriptive statistics based on each proficiency group show that the mean and median scores for IELTS 6.5 were both 6.25 (SD=1.25, range=2.5); the scores for IELTS 7.0 were 5.75 and 6.25 (SD=3.44, range=8.5); the scores for IELTS 7.5 were 11.86 and 12.5 (SD=1.75, range=5.5); the scores for IELTS 8.0 were 11.25 and 12.5 (SD=3.91, range=10.5); and the scores for L1 English were 12.04 and 13.5 (SD=3.11, range=7.0). The potential differences in the mean and median scores according to each group therefore seem to indicate that the students with higher proficiency, on average, appeared to achieve the similar or higher test scores than their counterparts with lower proficiency. In particular, the L1 English students achieved the highest post-test scores (Mean=12.04, Median=13.5) among all the proficiency groups, suggesting that the L1 English students who were more linguistically advantaged could possibly achieve higher scores than those who learned through L2 English.

Table 24: Descriptive statistics of post-test scores and IELTS level (n=27)

Variable	Proficiency	N	Mean	Median	SD	Min	Max	Range
Post-test scores	IELTS 6.5	2	6.25	6.25	1.25	5	7.5	2.5
	IELTS 7.0	4	5.75	6.25	3.44	1	9.5	8.5
	IELTS 7.5	7	11.86	12.5	1.75	8	13.5	5.5
	IELTS 8.0	8	11.25	12.25	3.91	4.5	15	10.5
	L1 English	6	12.04	13.5	3.11	7.5	14.5	7
	Total	27	10.40	12.0	3.15	1	15	14

To further examine the differences in the proportion of higher and lower performing students according to proficiency, a distribution of the post-test scores was computed below (see Figure 21). Based on the three-sigma rule (Pukelsheim, 2012), the cut-off lines between the top end ($\mu+\sigma = 10.40+3.15 = 13.55$) and the bottom end ($\mu-\sigma = 10.40-3.15 = 7.25$) of this distribution was computed. Accordingly, the values above 13.55 for the top end mark and below 7.25 for the bottom end mark out of 17 fell within the two and three standard deviations of the mean and accounted approximately for the top and bottom 15.8% of the distribution, respectively. When exploring the observations above 84.2% of the distribution as the threshold of the higher test scores, it was clear that the top end of the distribution was exclusive to those who reached IELTS 8.0 or above.

Figure 21: Distribution of post-test scores by proficiency (n=27)



In order to explore the exact proportion of participants who fell within the top and bottom end of this distribution, Table 24 below shows a breakdown of the frequency and cumulative percentage of the higher and lower performing participants according to proficiency. In the top range, none of the participants with IELTS 7.5 or below were observed while 37.5% of the IELTS 8.0 students (n=3) and 50% of the L1 English students (n=3) fell within this range. In the bottom range, 100% of the IELTS 6.5 students (n=2) and 50% of the IELTS 7.0 students (n=2) were represented while

only 25% of the IELTS 8.0 students (n=2) and 16.7% of the L1 English students (n=1) fell within this region. This descriptive result reveals that there was a clear difference in the proportion of students represented at the top and bottom end of this distribution according to proficiency, and that the higher and lower proficiency students tended to be clustered around the top and bottom end of this distribution, respectively.

Table 25: Top and bottom 15.8% of post-test distribution by proficiency

	IELTS 6.5 (n=2)		IELTS 7.0 (n=4)		IELTS 7.5 (n=7)		IELTS 8.0 (n=8)		L1 English (n=6)	
Post-test score	**Frequency**	**Cumulative percent**	**Frequency**	**Cumulative percent**	**Frequency**	**Cumulative percent**	**Frequency**	**Cumulative percent**	**Frequency**	**Cumulative percent**
15.0	0	0%	0	0%	0	0%	2	25.0%	1	16.7%
14.5	0	0%	0	0%	0	0%	1	12.5%	2	33.3%
Total	**0**	**0%**	**0**	**0%**	**0**	**0%**	**3**	**37.5%**	**3**	**50.0%**
7.5	1	50%	0	0%	0	0%	0	0%	1	16.7%
7.0	0	0%	0	0%	0	0%	1	12.5%	0	0%
5.0	1	50%	0	0%	0	0%	0	0%	0	0%
4.5	0	0%	0	0%	0	0%	1	12.5%	0	0%
4.0	0	0%	1	25.0%	0	0%	0	0%	0	0%
1.0	0	0%	1	25.0%	0	0%	0	0%	0	0%
Total	**2**	**100%**	**2**	**50.0%**	**0**	**0%**	**2**	**25.0%**	**1**	**16.7%**

6.3.3 Bivariate descriptive analysis: EMI challenges and proficiency

Next, descriptive statistics in Table 25 below were calculated to explore how the EMI Challenges Scale scores (n=27) varied according to proficiency, with lower average values indicating that the students encountered greater difficulty. The mean and median scores for each proficiency group suggest that the students with higher proficiency, on average, appeared to achieve the similar or higher scores than their counterparts with lower proficiency. This indicates that the higher the IELTS score, the greater the ease with which the students experienced EMI; the lower the IELTS score, the greater the magnitudes of the challenges were expressed. The descriptive statistics based on each proficiency level in Table 25 show that the mean and median scores for IELTS 6.5 were both 2.8 (SD=0, range=0); the scores for IELTS 7.0 were 3.63 and 3.75 (SD=0.45, range=1.0); the scores for IELTS 7.5 were 3.87 and 3.75 (SD=0.38, range=1.1); the scores for IELTS 8.0 were 4.33 and 4.45 (SD=0.67, range=2.0); and the scores for L1 English were 4.64 and 4.7 (SD=0.24, range=0.7). This suggests that the L1 English students seemed to have achieved the highest EMI

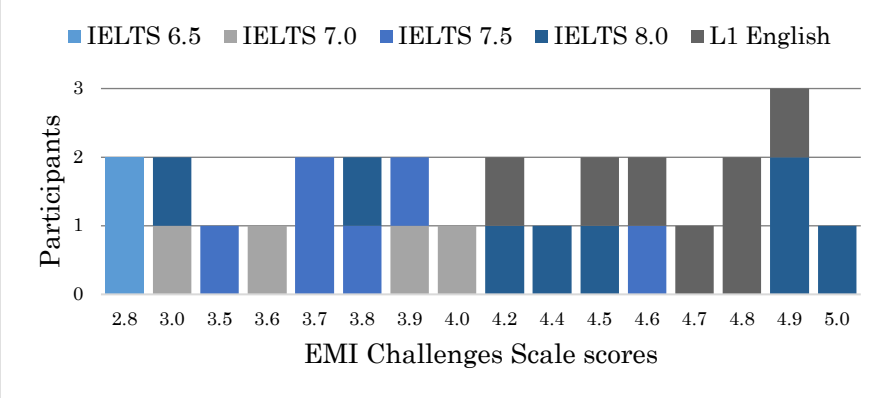
Challenges Scale scores (Mean=4.64, Median=4.7) among all the proficiency groups, indicating that those who possessed a sound foundation in the language expressed greater ease for EMI than their lower proficiency counterparts. Thus, the descriptive result suggests that students at different levels of proficiency expressed clear differences in the perceptions of academic challenges associated with EMI study.

Table 26: Descriptive statistics of challenges and IELTS level (n=27)

Variable	Proficiency	N	Mean	Median	SD	Min	Max	Range
Challenges	IELTS 6.5	2	2.8	2.8	0	2.8	2.8	0
	IELTS 7.0	4	3.63	3.75	0.45	3	4	1
	IELTS 7.5	7	3.87	3.75	0.38	3.5	4.6	1.1
	IELTS 8.0	8	4.33	4.45	0.67	3	5	2
	L1 English	6	4.64	4.7	0.24	4.2	4.9	0.7
	Total	27	4.10	4.2	0.54	2.8	5	2.2

To further examine the proportion of students who experienced the higher and lower levels of challenges according to proficiency, a distribution of the EMI Challenges Scale scores was plotted below (see Figure 22). When exploring the observations above 84.2% ($\mu+\sigma = 4.10+0.54 = 4.64$) and below 15.8% of the distribution ($\mu-\sigma = 4.10-0.54 = 3.56$) as the threshold of the higher and lower levels of challenges, it became clear that the top end of the distribution was exclusive to those who reached the IELTS 8.0 and above whereas the bottom end of the distribution represented more of the lower proficiency students who reached IELTS 6.5 and 7.0.

Figure 22: Distribution of challenges by proficiency (n=27)



In order to explore the exact proportion of participants who fell within the top and bottom end of this distribution, Table 26 below shows a breakdown of the frequency and cumulative percentage of the EMI participants according to proficiency. In the top range, none of the participants with IELTS 7.5 or below were observed, while 37.5%

of the IELTS 8.0 students (n=3) and 66.6% of the L1 English students (n=4) fell within this range. In the bottom range, while 100% of the IELTS 6.5 students (n=2) and 50% of the IELTS 7.0 students (n=2) were observed, only 14.3% of the IELTS 7.5 students (n=1) and 12.5% of the IELTS 8.0 students (n=1) fell within this region. Thus, the descriptive result indicates that the higher proficiency students who achieved IELTS 8.0 or above were more likely to report greater ease towards study through EMI than their lower proficiency counterparts who only reached IELTS 7.0 or below.

Table 27: Top and bottom 15.8% of challenges distribution by proficiency

	IELTS 6.5 (n=2)		IELTS 7.0 (n=4)		IELTS 7.5 (n=7)		IELTS 8.0 (n=8)		L1 English (n=6)	
Challenges	**Frequency**	**Cumulative percent**	**Frequency**	**Cumulative percent**	**Frequency**	**Cumulative percent**	**Frequency**	**Cumulative percent**	**Frequency**	**Cumulative percent**
5.0	0	0%	0	0%	0	0%	1	12.5%	0	0%
4.9	0	0%	0	0%	0	0%	2	25.0%	1	16.7%
4.8	0	0%	0	0%	0	0%	0	0%	2	33.3%
4.7	0	0%	0	0%	0	0%	0	0%	1	16.7%
Total	**0**	**0%**	**0**	**0%**	**0**	**0%**	**3**	**37.5%**	**4**	**66.6%**
3.6	0	0%	1	25.0%	0	0%	0	0%	0	0%
3.5	0	0%	0	0%	1	14.3%	0	0%	0	0%
3.0	0	0%	1	25.0%	0	0%	1	12.5%	0	0%
2.8	2	100%	0	0%	0	0%	0	0%	0	0%
Total	**2**	**100%**	**2**	**50.0%**	**1**	**14.3%**	**1**	**12.5%**	**0**	**0%**

6.3.4 Bivariate descriptive analysis: Content knowledge achievement and EMI challenges

Descriptive statistics in Table 27 below were calculated to explore how the EMI students' post-test scores (n=27) varied according to the levels of challenges measured on the EMI Challenges Scale. The mean and median scores for each of the challenges levels suggest that the students who reported greater ease for EMI study (i.e. higher challenges scores), on average, seemed to achieve higher test scores than their counterparts who experienced greater difficulty (i.e. lower challenges scores). The descriptive statistics based on the levels of challenges in Table 27 show that the mean and median test scores for the 3-point mark (neutral) were 9.0 and 8.0 (SD=4.26, range=10); the scores for the 4-point mark (easy) were 9.41 and 10.5 (SD=4.27, range=13.5); and the scores for the 5-point mark (very easy) were 11.95 and 12.75 (SD=3.05, range=8). The potential differences in the mean and median

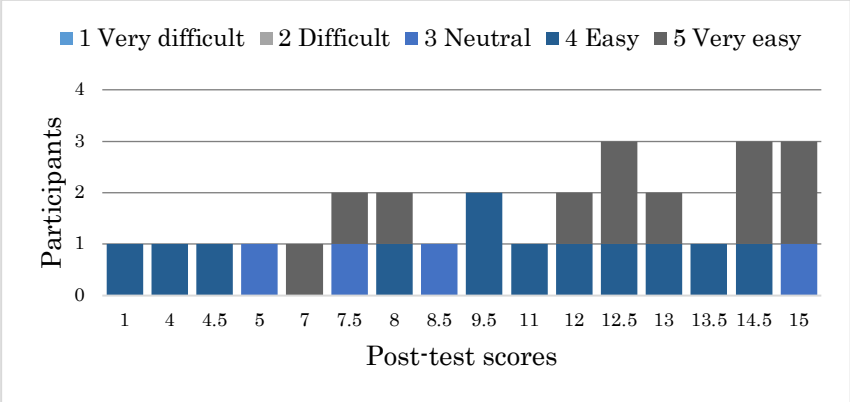
scores according to each group therefore seem to indicate that the higher the score on the EMI Challenges Scale, the greater the students performed on the post-test; the lower the challenges score, the lower they scored on the post-test. The descriptive result also highlights that all participants were over the 3-point mid-mark of the EMI challenges, indicating that none of them appeared to face significant challenges with EMI study.

Table 28: Descriptive statistics of post-test scores and challenges (n=27)

Variable	Challenges	N	Mean	Median	SD	Min	Max	Range
Post-test scores	1 Very difficult	0	0	0	0	0	0	0
	2 Difficult	0	0	0	0	0	0	0
	3 Neutral	4	9.0	8.0	4.26	5	15	10
	4 Easy	12	9.41	10.5	4.27	1	14.5	13.5
	5 Very Easy	11	11.95	12.75	3.05	7	15	8
	Total	27	10.40	12.0	3.15	1	15	14

To examine the difference in the proportion of the higher and lower performing students according to the levels of challenges, a distribution of the post-test scores was plotted (see Figure 23). When exploring the observations above 84.2% ($\mu+\sigma$ = 10.40+3.15 = 13.55) and below 15.8% ($\mu-\sigma$ = 10.40-3.15 = 7.25) as the threshold of the successful test scores, it was clear that the students who were clustered around the top and bottom end of this distribution represented a range of perceived academic challenges from 3 (neutral) to 5 (very easy). An exploration of the top and bottom ranges of this distribution indicated that there was necessarily no clear relationship between the students' post-test scores and the level of their academic challenges.

Figure 23: Distribution of post-test scores by challenges (n=27)



To ascertain the precise number and percentage of the students represented in the top and bottom end of the distribution according to the levels of challenges, the frequency

and cumulative percentage of the higher and lower performing participants was computed below (see Table 28). The top 84.2% end of the distribution represented 36.4% of the students on the 5-point mark (very easy), 16.6% of the students on the 4-point mark (easy), and 25% of the students on the 3-point mark (neutral). Consistently, the bottom end of the distribution was also broadly equal across the challenges levels, entailing 50% of the students on the 3-point mark (neutral), 25.0% of the students on the 4-point mark (easy) and 18.2% of the students on the 5-point mark (very easy). Thus, an exploration of this top 84.2% end and bottom 15.8% end of the distribution clarified that these specific ranges were not necessarily exclusive to those who reported greater ease or difficulty for EMI study, respectively. Hence, an inspection of the individual post-test scores within these top and bottom ranges suggests that the extent to which academic challenges were perceived to be more or less difficult may not be necessarily related to the level of academic success in EMI.

Table 29: Top and bottom 15.8% of post-test distribution by challenges

	1 Very difficult (n=0)		2 Difficult (n=0)		3 Neutral (n=4)		4 Easy (n=12)		5 Very easy (n=11)	
Post-test score	**Frequency**	**Cumulative percent**	**Frequency**	**Cumulative percent**	**Frequency**	**Cumulative percent**	**Frequency**	**Cumulative percent**	**Frequency**	**Cumulative percent**
15.0	0	0%	0	0%	1	25.0%	1	8.3%	2	18.2%
14.5	0	0%	0	0%	0	0%	1	8.3%	2	18.2%
Total	**0**	**0%**	**0**	**0%**	**1**	**25.0%**	**2**	**16.6%**	**4**	**36.4%**
7.5	0	0%	0	0%	1	25.0%	0	0%	1	9.1%
7.0	0	0%	0	0%	0	0%	0	0%	1	9.1%
5.0	0	0%	0	0%	1	25.0%	0	0%	0	0%
4.5	0	0%	0	0%	0	0%	1	8.3%	0	0%
4.0	0	0%	0	0%	0	0%	1	8.3%	0	0%
1.0	0	0%	0	0%	0	0%	1	8.3%	0	0%
Total	**0**	**0%**	**0**	**0%**	**2**	**50.0%**	**3**	**25.0%**	**2**	**18.2%**

6.3.5 Summary of descriptive statistics

In summary, the descriptive statistics broadly highlighted the relationship between the EMI students' English language proficiency, academic challenges and post-test scores. The students with higher English proficiency were more likely to achieve higher post-test scores and encounter fewer EMI challenges than their counterparts with lower proficiency. Consistently, those who reported fewer EMI challenges were more likely to achieve higher post-test scores than their counterparts who faced greater difficulty.

Nevertheless, based on Table 26 above which showed how the post-test scores differed according to the IELTS levels, it is important to note that the descriptive statistics also indicated that test scores for some of the proficiency groups ranged widely. This means that the values observed in the groups were widespread, some of which were outliers that affected the average values. For example, the highest value for the IELTS 8.0 group was 15 and the lowest 4.5, with a range of 10.5, a mean of 11.25, and a standard deviation of 3.91. The IELTS 8.0 group therefore ranged widely in its test scores, with a minimum value that fell outside of three standard deviations of the mean. The influence of this outlier score suggested that while the IELTS 8.0 group, on average, achieved greater test performance than their lower proficiency counterparts, some of the students in this group underperformed in their test scores compared to their peers. A inspection of the raw data (see Figure 21 above) subsequently revealed that two IELTS 7.0 students scored 1 and 4, respectively and two IELTS 8.0 students scored 4.5 and 7, respectively. This finding reveals that these students were not as successful as their peers in the same proficiency groups, highlighting the importance of inspecting the individual test scores.

For this reason, I emphasise the importance of inspecting raw data to examine the test score distribution instead of simply computing and comparing group mean scores. As such, I reviewed the individual test scores according to proficiency to explore whether any of them deviated widely from the rest of their peers (i.e. outliers) in the same proficiency groups. This in particular enabled me to identify struggling learners in the entire EMI student sample. To this end, I argue that quantitative data per se did not tell the whole story; a qualitative examination of student interviews would supplement this data to tease out the experiences and perceptions of individual students and their learning experiences, as research question 3.2 reveals.

6.4 Correlations: Proficiency, pre-test scores, challenges and post-test scores

This section reports on the results of correlation analyses which further explored the relationship between a range of factors, including the EMI students' English language proficiency, pre-test scores, post-test scores and academic challenges. A Spearman's rank-order correlation was carried out (see Table 29) to reveal that there was a moderate to strong correlation between English language proficiency and post-test

scores (r=.452, p<.05). There were also very strong correlations between proficiency and an overall composite mean for academic challenges (r=.735, p<.001) as well as each of the skill areas, such as academic speaking (r=.757, p<.001), academic reading (r=.684, p<.001) and academic listening skills (r=.706, p<.001). This indicates that students with higher proficiency encountered fewer academic challenges in all of the academic English skill areas and achieved higher post-test scores than their counterparts with lower proficiency. In contrast, there was no correlation between IELTS scores and pre-test scores (r=.049, p=.808), meaning that possessing advanced language knowledge at the beginning of the term was not necessarily indicative of the level of prior content knowledge.

Pre-test scores showed a strong and positive correlation with post-test scores (r=.424, p<.05), but were not correlated with academic English challenges in any of the skill areas. This suggests that possessing substantial prior content knowledge was highly facilitative of increased content test performance, but was not crucial to alleviating academic English challenges in any of the skill areas. Finally, there was a weak to moderate correlation between post-test scores and academic English challenges in the areas of reading (r=.340, p<.05). However, post-test scores were not correlated with any other skill areas including the composite mean score (r=.356, p=.068), academic speaking (r=.333, p=.089), and academic listening skills (r=.308, p=.118). This suggests that while those who reported fewer academic challenges in the area of academic reading could potentially achieve greater content test performance, generally no statistically significant correlations were observed between the perceived ease of academic study and content learning achievement in EMI.

Table 30: A Spearman's rank-order correlation between proficiency, pre-test score, post-test score and challenges (n=27)

	Proficiency	**Pre-test**	**Speaking challenges**	**Reading challenges**	**Listening challenges**	**Composite mean of challenges**
Proficiency						
r	-	.049	.757**	.684**	.706**	.735**
Sig. (2-tailed)	-	.808	.000	.000	.000	.000
Pre-test						
r	.049	-	.101	.231	.174	.175

Sig. (2-tailed)	.808	-	.617	.246	.387	.382
Post-test						
r	.452*	.424*	.333	.340*	.308	.356
Sig. (2-tailed)	.018	.028	.089	.047	.118	.068

Note: N=27, *. P<0.05 level, **. P<0.01 level (2-tailed), r=correlation coefficient. The correlation coefficient is a measure of an effect size: (r=.1) is small, (r=.3) is medium and (r=.5) is large effect (Field, 2018).

6.4.1 Summary of correlations

In summary, the results of the correlation analyses suggest that the students with higher English proficiency were more likely to achieve greater content knowledge acquisition and face fewer challenges in all of the academic English skill areas. Furthermore, more successful students were more likely to be equipped with higher prior content knowledge at the start of the semester and were possibly able to mitigate academic English challenges, in the areas of reading. However, it was also clear that possessing considerable prior content knowledge was not necessarily associated with the achievement of advanced English proficiency or greater ease for any of the academic English challenges associated with EMI. These associations between the variables are further explored through regression analyses in the subsequent section.

6.5 Regression analysis: Proficiency, pre-test scores, challenges and post-test scores

Finally, three separate simple linear regression analyses and one multiple regression analysis were performed to explore the extent to which the following variables predicted academic success in EMI. The variables entered into the simple linear regression models as potential predictor variables included the EMI students' English language proficiency (IELTS scores); prior content knowledge (pre-test scores); and perceptions of the academic challenges associated with EMI (EMI Challenges Scale scores). Academic success in EMI (post-test scores) was put into all of these models as an outcome variable. Subsequently, the multiple regression analysis explored the extent to which proficiency and prior content knowledge predicted academic success in EMI when these two predictor variables were put into one regression model. Due to the violation of the assumption of multicollinearity in which proficiency and challenges were highly correlated (r=.735, p<.001), as confirmed in Table 29 above,

the predictor variables included in the multiple regression model were only limited to prior content knowledge and proficiency to answer the research question.

In order to assess whether the data satisfied all the assumptions of simple and multiple linear regression, firstly, linearity was assessed by plotting three separate scatterplots for proficiency, challenges and pre-test score against post-test score with superimposed regression lines (see Appendix 27 for scatterplots). Visual inspection of these scatterplots confirmed linear relationships between these three variables and post-test scores. Secondly, there was independence of residuals, as examined by a Durbin-Watson statistic of 1.853 for proficiency; 2.092 for challenges; and 2.204 for pre-test score. Thirdly, the normal probability plots of standardised predicted values against standardised residuals demonstrated that the data met the assumptions of homoscedasticity of variance (see Appendix 28 for scatterplots), and the residuals were approximately normally distributed (see Appendix 29 for normal probability plots). Finally, no outliers were detected where the standardised residuals were not greater than the cut-off value of the three standard deviations. Accordingly, it was concluded that the data all met the assumptions of regression analyses.

6.5.1 To what extent does proficiency level predict post-test score in EMI?

Firstly, this section explores the association between the EMI students' English language proficiency and content knowledge achievement. Simple linear regression revealed a statistically significant positive linear relationship between IELTS scores and post-test scores ($\beta = 2.090$. $F(1, 25) = 6.497$, $p < 0.05$), indicating that the magnitude of this relationship was a moderate to strong effect, which was confirmed with a Pearson correlation coefficient of 0.45 ($p < 0.05$). The slope coefficient for IELTS scores was 2.090, 95% CI [0.440, 4.141], which suggests that post-test score would increase by 2.090 for every 0.5 point increase in IELTS score. The value of R^2 was 0.206, which means that 20.6% of the variance in post-test score was explained by this model containing proficiency score, with adjusted $R^2 = 17.5\%$. Additionally, Table 30 shows that proficiency score achieved a standardised beta coefficient (β) with a value of 0.454. This indicates that with every increase of one standard deviation in IELTS level (SD=0.77), post-test score rose by 0.454. To summarise, IELTS score statistically significantly predicted post-test score; and the higher the

proficiency, the higher the content test score.

Table 31: Linear regression coefficients table: Proficiency and post-test scores

	ΔR^2	*B*	*SE B*	*Standardised β*	*t value*	*p value*
Constant	0.206	-7.552	7.071		-1.068	0.296
Proficiency		2.090	0.899	0.454	2.549	< 0.017*

6.5.2 To what extent does challenges score predict post-test score in EMI?

Secondly, a simple linear regression was performed again to investigate the variance in post-test score in relation to a composite mean score of academic challenges in EMI. A linear regression revealed that the R^2 value was 0.128, indicating that challenges score could account for 12.8% of the variation in post-test score, with adjusted $R^2 = 0.093$ (9.3% of the variability) which was a smaller proportion of variance explained than that of IELTS score. However, when the statistical significance of the overall regression model was assessed, it became clear that challenges score did not achieve a significant positive linear relationship with post-test score ($\beta = 2.053$. $F(1, 25) = 3.678$, $p = 0.67$), which was also confirmed with a Pearson correlation coefficient of 0.356 ($p > 0.05$). Consistently, the non-statistically significant slope coefficient ($p > 0.05$) did not predict the increment of post-test score by the change in values of challenges score. Table 31 underpins these non-statistically significant finding as illustrated by the standardised Beta value of 0.358; post-test scores only increased by 0.358 for every one standard deviation increase in challenges scores (SD=0.54). In summary, challenges, as measured on the EMI Challenges Scale, was not a statistically significant predictor of content knowledge achievement in EMI, as measured by post-test scores.

Table 32: Linear regression coefficients table: Challenges and post-test scores

	ΔR^2	*B*	*SE B*	*Standardised β*	*t value*	*p value*
Constant	0.128	1.989	4.438		0.448	0.658
Challenges		2.053	1.070	0.358	1.918	0.67

6.5.3 To what extent does pre-test score predict post-test score in EMI?

Thirdly, in order to explore the effect of the EMI students' prior content knowledge measured by pre-test score on their post-test score performance, another simple linear regression was conducted. When evaluating the statistical significance of the

regression model, it became clear that there was a statistically significant linear relationship between the two variables (β = 0.500. F(1, 25) = 5.176, p < 0.05), which was also confirmed with a correlation coefficient ($r = 0.42$, $p < 0.05$), as illustrated in Table 29 above. The slope coefficient for pre-test score was 0.500, 95% CI [0.047, .953], suggesting that post-test score increased by 0.5 for every point increase in pre-test score. When exploring the proportion of variation explained by the model, the R^2 value was 0.172 in Table 32 below, which indicated that 17.2% of the variation in post-test scores was accounted for by the model containing pre-test score, with adjusted R^2 = 13.8%. Table 32 also presents the standardised Beta value (β = 0.414), which reports that for every one standard deviation rise in pre-test score (SD=3.21), post-test scores rose by 0.414. In summary, pre-test scores statistically significantly predicted post-test scores.

Table 33: Linear regression coefficients table: Pre-test scores and post-test scores

	ΔR²	***B***	***SE B***	***Standardised β***	***t value***	***p value***
Constant	0.172	9.490	0.800		11.864	< 0.001**
Pre-test scores		0.500	0.220	0.414	2.275	< 0.032*

6.5.4 To what extent do proficiency and pre-test score predict post-test score in EMI?

Finally, to explore the extent to which proficiency and pre-test scores explained the variance in the outcome variable (post-test scores), both of these variables were put into one multiple linear regression model as predictor variables. A significant linear regression equation was found (F(2, 24) = 6.545, p < 0.005**), with an R^2 of 0.353 (see Table 33). These two predictors therefore explained 35.3% of the variance in post-test scores, with adjusted R^2 = 30.0%. The strength of this linear association between the variables was also confirmed with the multiple correlation coefficient ($r = 0.59$, $p < 0.005$**). This model therefore explained a larger proportion of variance in post-test scores than any of the separate single linear models performed above. Table 33 also illustrates that both proficiency and pre-test scores were significant predictors of post-test scores. The slope coefficient for IELTS score was 2.053, 95% CI [0.440, 3.866], suggesting that post-test score increased by 2.053 for every point increase in IELTS level when other independent variables (pre-test scores) were held constant. Similarly, an increase in pre-test scores of one point was associated with an

increase in post-test scores of 0.464, 95% CI [0.053, 0.874] (i.e. the coefficient of pre-test scores is .464) when proficiency was kept constant. To conclude, proficiency and pre-test scores were both significant predictors of EMI success and explained a larger proportion of variance in post-test scores when these two variables were considered in one regression model.

Table 34. Multiple regression coefficients table: Proficiency, pre-test scores and post-test scores

	ΔR^2	*B*	*SE B*	*Standardized β*	*t value*	*p value*
Constant	0.353	-7.311	6.517		-1.12	0.273
Pre-test scores		0.464	0.199	0.384	2.332	0.028**
Proficiency		2.053	0.830	0.427	2.592	0.016**

6.5.5 Summary of regression analyses

To conclude, this section has first assessed the statistical significance of the three linear regression models, revealing that IELTS scores and pre-test scores were both significant predictors of academic success in EMI while the EMI Challenges Scale score did not significantly predict post-test score achievement. Thus, the slope coefficient ($p < 0.05$) showed that post-test score increased by 2.090 for every level increase in IELTS score, and increased by 0.5 for every point increase in pre-test score. The non-statistically significant result of the challenge score ($p = 0.067$) suggests that the students in the current sample might have expressed greater language-related challenges, but they still seemed to be able to achieve successful content test performance. A multiple regression model included proficiency and pre-test scores, which explained 35.3% of the variance in post-test scores. Seeing that proficiency and pre-test scores in simple linear regression models only explained 20.6% and 17.2%, respectively, this multiple regression model explained a larger proportion of variance in post-test scores. Overall, the regression analyses revealed that proficiency and pre-test score reflected a positive increase in the academic success in EMI as measured by post-test score.

6.6 Summary of research question 3.1

To summarise, research question 3.1 has explored the extent to which factors, such as the EMI students' English language proficiency, prior content knowledge, and perceptions of academic challenges are associated with content knowledge

achievement. Simple linear regression analyses revealed that IELTS and pre-test score statistically significantly predicted post-test score, accounting for 20.6% and 17.2% of the variation in the post-test score, respectively. A multiple regression was also performed to support this finding, accounting for 35.3% of the variation in the post-test score. This indicates that higher the proficiency and prior content knowledge, the more successful the students were in gaining content knowledge through English. Conversely, challenges score generated a non-significant regression result, which was confirmed by the correlation coefficient (r=.356, p=.068). The descriptive statistics also showed that the proportion of those who succeeded in EMI study was not determined by the level of academic challenges experienced in EMI.

The current quantitative analyses discerned a descriptive difference in the post-test scores between the groups. That is, the higher post-test scores were more apparent among the learners with IELTS 8.0 than IELTS 7.0. However, it is important to highlight that the sample size in the current study would not allow the use of t-tests or the one-way ANOVA in combination with post hoc tests. This means that it was not feasible to statistically test whether an overall significant difference in the post-test scores occurred between groups with different levels of English proficiency, pre-test score or challenges score. It was also not possible to detect which specific groups differed and confirm where exactly the differences lay. Without the multivariate analyses, it was not possible to establish that the students reaching IELTS 7.0 and 8.0, for instance, were statistically significantly different in terms of their perceived academic challenges or post-test performance. In the absence of more fined-grained statistical procedures, the current study utilised data triangulation of descriptive statistics, correlation, and regression analyses to illustrate the extent to which these key variables were associated with academic success in EMI.

In the quantitative analyses, student academic success and challenges have been operationalised in terms of numerical variables, while the interview findings presented in the subsequent section will focus on individual students' learning experiences, which offers a full picture of differences in student performance across different levels of English proficiency. The qualitative data in the following section explores various factors contributing to student content learning and examines why some of them were not as successful learners as their peers in the same proficiency group.

6.7 Research question 3.2

The quantitative analysis in the previous section explored the relationships between the EMI students' prior content knowledge (i.e. pre-test scores), English proficiency (i.e. IELTS scores), academic challenges (i.e. EMI Challenges Scale) and content knowledge achievement (i.e. post-test scores). The results seem to suggest that challenges were more noticeable particularly among the less proficient English learners than their counterparts with higher English proficiency. Similarly, those with higher English proficiency were also more successful in gaining content knowledge through EMI than their counterparts with lower proficiency. In the qualitative analysis, an exploration of the interview data investigates the EMI participants' perceptions of the relationship between English proficiency and challenges. It also explores the array of other factors to which the participants attributed their perceptions of challenges in EMI (e.g. motivation, prior content knowledge, previous experience of learning through EMI). In the quantitative analysis, a reductionist approach in the operationalisation of the EMI Challenges Scale was taken, which captured difficulties in pre-defined categories. Conversely, the qualitative analysis offers detailed insight into the students' perceptions of EMI challenges in more nuanced terms. Accordingly, the qualitative analyses organises interviewees according to two key dimensions: English proficiency levels, and levels of content learning achievement in the introductory Chemistry course. Specifically students from five proficiency levels (i.e. IELTS 6.5; 7.0; 7.5; 8.0; and L1 English) were represented in the interview sample. Thus, these participants offer cross sections of the lower and higher levels of English language proficiency from which to draw in-depth interpretations about perceptions of their challenges experienced in EMI. These data are used to answer the following research questions :

- RQ3: What are the relationships between the EMI students' English language proficiency, academic English challenges, prior content knowledge, and content knowledge achievement?
    - RQ 3.2: How do the EMI students perceive English language proficiency to affect their perceptions of challenges towards EMI study? (Perceived relationship)

6.8 Content learning-related challenges

The interview data of 13 EMI participants representing a range of English proficiency levels and content knowledge achievement were analysed using thematic analysis, which involved identifying emerging patterns across the data. Table 34 below shows a cross-tabulation of English proficiency and post-test scores. Some cells are blank due to the absence of certain combinations in the sample (e.g. students of low proficiency who achieved high content scores, and vice versa). Interview coding was conducted abductively, with initial deductive coding conducted using themes from the literature pertaining to EMI challenges and success. These codes were subsequently grouped into themes to capture recurring patterns. Accordingly, the following four main themes emerged (see Appendix 14 for a final list of main and subthemes): (1) comprehension of the course textbook; (2) comprehension of the course lectures; (3) Final Grade; and (4) adjustment to the new content learning environment. Below, the results will be presented according to these four themes.

Table 35: Matrices of post-test scores and IELTS level (n=13 EMI students)***

English Proficiency (IELTS Scores) / Content learning (Post-test scores)	Post-test low (1-6)** (n=3)	Post-test middle (6-12)** (n=6)	Post-test high (12-17)** (n=4)
IELTS 6.5 (n=2)	Hikaru (Post-test 5)*	Rina (Post-test 7.5)*	
IELTS 7.0 (n=4)	Kiley (Post-test 4) Marii (Post-test 1)*	Ayaka (Post-test 8.5)* Minori (Post-test 9.5)*	
IELTS 7.5 (n=1)			Tikamon (Post-test 12.5)
IELTS 8.0 (n=3)		Charlotta (Post-test 9.5)	Miyazaki (Post-test 15)* Lisa (Post-test 12.5)
L1 English (n=3)		Maxwell (Post-test 8) Saika (Post-test 7.5)	Emma (Post-test 12.5)

* Those who participated in both the first and second round of interviews ** Post-test scores: Low (0-40%), Medium (40-70%) and High (70-100%) adopted from the university's grading criteria *** See Appendix 24 for matrices of pre-test score, challenges and post-test scores

6.8.1 Comprehension of the course textbook

A first theme concerning EMI students' challenges was their inadequate comprehension of the course textbook. For ease of discussion, this theme was divided into four sub-themes: insufficient vocabulary knowledge; slow reading speed; reliance on translation; and inefficient reading.

A first sub-theme was the students' difficulty in understanding the vocabulary used in the textbook. When this theme was explored according to IELTS level, it became apparent that those with lower L2 proficiency were more likely to experience this challenge than those who had higher L2 proficiency. For instance, the most striking differences between the lower proficiency student, Hikaru (IELTS 6.5, post-test 5), and the higher proficiency student, Lisa (IELTS 8.0, post-test 12.5), lay in the type of vocabulary they struggled with and how frequently they relied on the use of dictionaries. While Hikaru frequently searched for both general academic and technical vocabulary, Lisa only occasionally had difficulty with technical terms and reported less reliance on using dictionaries.

> **Hikaru (EMI):** I look up words in a dictionary all the time. In the chapter 'Chemical Bonding', I didn't know words, like 'involve', 'polarised', 'charged' and 'law'.

> **Lisa (EMI):** I only sometimes look up words in a dictionary. I can usually correctly guess the meanings of unfamiliar words from the contexts. I sometimes don't understand certain technical terms but they are usually explained in detail in the textbook. So I usually don't have issues with English vocabulary.

When exploring this theme in light of the cut-off level between the point at which students felt at ease and the point at which they started to perceive difficulties, it was evident that those reaching IELTS 8.0 generally started reporting greater ease than difficulty in the area of general academic vocabulary. Their challenge revolved only around understanding unfamiliar but key technical terms, which was, needless to say, a general academic challenge for all participants. In comparison, the students with IELTS 7.5 or less were not even able to manage general academic terms (and thus were in stark contrast with the IELTS 8.0 participants who had at least mastered the academic terms). Below Tikamon (IELTS 7.5, post-test 12.5) and Emma (L1 English, post-test 12.5) achieved the same post-test scores but perceived their challenges differently depending on English proficiency.

> **Tikamon:** When I read the textbook, I have to look up not only technical words but also so many other basic English words. I wish I knew more English words.

> **Emma:** I usually understand the textbook without any language problem, like not understanding English words. My problem is more about not understanding technical terminologies in the textbook.

Thus, this evidence broadly suggests that English proficiency played a pivotal role in determining the magnitude of students' vocabulary-related challenges. For this reason, Lisa and Emma found it easier to understand both academic and technical words and used dictionaries less frequently than Tikamon and Hikaru of lower proficiency.

A second sub-theme was the participants' slow reading speed. An exploration of this theme according to IELTS level indicated that the lower proficiency students mentioned this challenge more frequently than the higher proficiency students. For example, Marii (IELTS 7.0, post-test 1) grappled with L2 English grammar structures and expressed that she had a slower reading speed than that of her L1, whereas Miyazaki (IELTS 8.0, post-test 15) could understand the textbook accurately without slowing down the speed of reading.

> **Marii:** I tend to stop and read the same sentences many times. It takes me a while to understand the meanings of English sentences. I have to think about English grammar, like SVO, SVC.

> **Miyazaki:** I may change how quickly I read depending on English or Japanese but it depends more on what I read. If it's a difficult concept, I'd read more carefully and slowly. Language doesn't matter.

While there was a clear relationship between proficiency and reading speed, having advanced proficiency did not necessarily guarantee a certain level of reading speed. For example, Saika (pre-test 0, post-test 7.5), still wrestled with a somewhat slow reading speed, despite being an L1 English speaker. In this case, her low prior content knowledge was the critical factor in her slow reading, not linguistic issues per se. Tikamon's case (pre-test 5.5, post-test 12.5) below also highlighted the significance of prior content knowledge in the learning of Chemistry. While expressing a bewildering proportion of unfamiliar vocabulary, she could maintain a fast reading speed owing to her advanced prior content knowledge.

> **Saika:** There are too many basic Chemistry terms and concepts I need to learn because I didn't study Chemistry before university. I constantly need to look up key words online and it takes a long time to read the textbook.

> **Tikamon:** When I read the textbook, I have to use online dictionaries a lot to search for unknown English words but having learnt Chemistry at high school has helped me process information faster.

For this reason, although L1 English speaker Saika was linguistically more favourably disposed towards learning through English than Miyazaki or Tikamon, the latter's advanced prior content knowledge allowed faster levels of reading speed. Thus, while low proficiency had a generally negative relationship to reading speed (indicated by the different experiences between Marii and Miyazaki), the level of prior content knowledge appeared to be facilitative of reading speed due to familiarity with content-related terminology.

The third theme was the students' reliance on translation. An analysis of this theme according to English proficiency revealed that the lower proficiency students had to rely on translation more frequently than the higher proficiency students. For instance, Hikaru (IELTS 6.5, post-test 5) regularly translated a whole paragraph into Japanese due to his limited English knowledge. In contrast, Lisa (IELTS 8.0, post-test 12.5) only occasionally translated unfamiliar technical terms into Japanese.

> **Hikaru:** I often translate a whole sentence to understand difficult concepts in Japanese first.

> **Lisa:** I usually search for the meaning of unknown technical terms in English but I also sometimes translate them into Japanese to see if I understand them more easily.

When exploring the data in terms of how frequently students had to use translation, it became apparent that the magnitude of this challenge was affected not just by proficiency, or whether students were L1 or L2 English speakers, but also by their high school teaching medium. For instance, those who came from a Japanese-medium high school system benefitted largely from understanding concepts in Japanese. Thus, Hikaru and Lisa who had previously learned Chemistry in Japanese, placed great value on adopting translation to compensate for their insufficient English vocabulary knowledge. Conversely, no students from an English-medium high school system reported adopting the same strategy. For example, Miyazaki (IELTS 8.0, post-test 15) noted that "I rarely use translation because my high school Chemistry class was taught in English".

Similarly, the use of translation was also likely to depend on whether students were L1 or L2 English speakers. For example, Lisa (post-test 12.5) despite achieving IELTS 8.0 still relied on translation while none of the L1 English counterparts reported this challenge. For example, Emma (L1 English, post-test 12.5) noted that "I am learning in English so it is unnecessary to translate English words". Thus, how frequently students had to rely on translation was influenced by an array of factors, including proficiency, whether students were L1 or L2 English speakers, and high school teaching medium.

The final theme that emerged from the interviews was students' frustration with their somewhat inefficient reading, defined as the inability to achieve the same comprehension as was possible in their L1 even when spending the same or even longer hours reading than in their L1. An exploration of this theme according to IELTS level revealed that the L2 English students all reported this challenge, regardless of English proficiency. For instance, Minori (IELTS 7.0, post-test 9.5) and Rina (IELTS 6.5, post-test 7.5) noted that L2 English reading was more cumbersome and time-consuming than that of their L1. Interestingly, even a student with higher proficiency, such as Lisa (IELTS 8.0, post-test 12.5) still perceived her reading comprehension to be somehow inferior than that of her L1.

> **Minori (EMI):** I usually spend longer hours preparing for my EMI lectures than my Japanese lectures, but I only understand 85% of my textbook.
>
> **Rina (EMI):** I read so much more slowly and less accurately in English. I have to spend longer time reading the textbook.
>
> **Lisa (EMI):** I read the textbook more carefully because it takes me a while to memorise information in English.

When comparing these students' experience, it was evident that all of them, regardless of proficiency, perceived their L1 learning to be more efficient than learning conducted in their L2. However, it is no surprise that Emma (L1 English, post-test 12.5) below who studied through her L1 did not face this issue. Thus, this challenge was more dependent on whether learning occurred through L1 or L2 English than proficiency.

> **Emma (EMI):** I'm fairly confident in my ability to learn chemistry. The textbook is not much more difficult than the one I used at high school.

Thus, although Emma (L1 English) and Lisa (IELTS 8.0) obtained the same post-test scores (post-test 12.5), it was only Lisa who found English-medium content learning to be inefficient. The post-test scores per se did not offer a full picture of their learning experience, highlighting that students experienced markedly different paths to the same outcome.

Overall, this section has explored the students' comprehension of the course textbook based on English proficiency, pointing to a relationship between challenges and proficiency. In general, the higher L2 proficiency students reported fewer challenges than the lower proficiency students. Table 35 below shows the number of times each sub-theme was coded (references) per participant in the data according to IELTS level. The only exception to this was that all participants across IELTS levels perceived reading through L2 to be less efficient than that of L1. While proficiency was crucial to alleviating challenges, no clear-cut linguistic threshold, according to IELTS level, was detected where students stopped reporting challenges. The data instead suggested a myriad of other factors which contributed to their perception of challenges (e.g. prior-content knowledge, educational background, L1 or L2 English learners).

Table 36: Matrices of challenges according to IELTS level

Participants Main & Sub-themes	EMI IELTS 6.5 (n=2)	EMI IELTS 7.0 (n=4)	EMI IELTS 7.5 (n=1)	EMI IELTS 8.0 (n=3)	EMI English L1 (n=3)
1. Comprehension of the course textbook	**13.5**	**10.2**	**9**	**5.9**	**2.6**
Inadequate vocabulary knowledge	3	2.3	2	1	1.3
Slow reading speed	4	3	2	0.5	1.3
Reliance on translation	3	1.3	2	1.7	0
Inefficient reading	3.5	3.6	3	2.7	0

6.8.2 Comprehension of the course lectures

The second theme emerging from the interview data was the students' comprehension of the course lectures. This main theme was divided into four sub-themes: low

concentration; low comprehension of humour and jokes from the lecturer; a 'double whammy' of content and language-related issues; and low comprehension of technical terms.

The first sub-theme was the students' concentration fatigue in lectures.When exploring this theme according to IELTS level, it became apparent that those with lower proficiency were more likely to encounter this challenge than those who had higher proficiency. For instance, Rina (IELTS 6.5, post-test 7.5) and Ayaka (IELTS 7.0, post-test 8.5) attributed the English medium nature of the course to be the primary reason for their difficulties, emphasising that low proficiency (e.g. inadequate vocabulary knowledge) prevented them from keeping focused on lectures.

> **Rina:** Listening in English is exhausting compared to Japanese because I have to listen to every single word carefully. When I hear unknown words, I stop processing information. I'm always drained of energy at the end of lectures.
>
> **Ayaka:** I always lose energy and become very tired when attending EMI lectures. My attention easily starts to wander. It is exhausting to try to understand every single word in English. It's easier to focus in Japanese.

Compared to Rina and Ayaka, some students with higher English proficiency reported greater ease for this challenge. For instance, Emma (L1 English, post-test 12.5) could selectively attend to key information. Similarly, Miyazaki (IELTS 8.0, post-test 15) also noted that EMI hardly ever affected her lecture comprehension negatively.

> **Emma:** It is very tiring to concentrate on the entire duration of lectures but I usually select information and pay attention only to the key parts of my lectures.
>
> **Miyazaki:** My concentration skills are very good when I study in English. Using English or Japanese doesn't affect my attention span.

Therefore, the data underpinned the relationship between proficiency and concentration fatigue in EMI lectures. Rina and Ayaka with lower proficiency had shorter attention spans than Emma and Miyazaki with higher proficiency who were more accustomed to EMI study.

The second theme was the students' frustration with their inadequate understanding of humour and jokes used by their L1 English lecturer. An exploration of this theme

based on IELTS level revealed that all participants reaching IELTS 8.0 or over stopped reporting this challenge. For example, Miyazaki (IELTS 8.0, post-test 15) was fully accustomed to listening to humour and jokes in English because of her previous exposure to EMI education. Her experience of learning from L1 English teachers at high school ensured that she was confident in understanding humour and jokes in English.

> **Miyazaki:** I remember many of my high school teachers were from America or England and often told us jokes to make us laugh in class. It would be difficult to understand jokes if I hadn't attended English-mediated classes at high school.

However, the challenge was more prominent among the lower proficiency students. For example, Rina (IELTS 6.5, post-test 7.5) who graduated from a Japanese-medium high school attributed a lack of prior experience of learning through EMI to be the major reason that she could not understand humour and jokes in English. Similarly, Ayaka (IELTS 7.0, post-test 8.5) struggled with the teacher's "spontaneous and quick" lecture delivery style. This indicates that the challenges were primarily a major concern for the lower proficiency participants who had little to no previous exposure to EMI education.

> **Rina:** It is difficult to understand jokes in English because I have never really studied through English at high school.

> **Ayaka:** I am not confident in understanding the nuances of jokes because the teacher speaks too spontaneous and quickly.

Thus, it was clear that the lower proficiency students mentioned this challenge more frequently than their counterparts with higher proficiency. Specifically, none of the students reaching IELTS 8.0 or above reported this issue. Prior experience of learning through EMI was also another crucial factor which helped to mitigate this challenge.

The third theme was a 'double whammy' of content and language issues. This was a challenge distinctive to the L2 English students who struggled not only with the linguistic issues but also with difficulties associated with learning new academic disciplines. An exploration of this theme in light of proficiency suggested that those with lower proficiency reported this challenge more frequently than their counterparts with higher proficiency. For instance, Hikaru (IELTS 6.5, post-test 5) struggled to

comprehend his EMI lectures. The reason for his challenge was not only because the subject matter was new but also the professor used many unfamiliar grammar and vocabulary. He attributed this comprehension issue to his low English proficiency.

> **Hikaru**: Learning natural science subjects is challenging for me because I only studied humanity subjects at high school. I want to try and learn a new subject at university. I am enjoying challenging myself but I sometimes wonder if I should have taken the Japanese course instead. Learning the subject I have never studied before in English has been really difficult. My professor uses a lot of grammar and vocabulary I do not know. I should have a better understanding in Japanese.

In contrast, Miyazaki (IELTS 8.0, post-test 15) did not experience any apparent language-related challenges. She occasionally faced difficulty in understanding disciplinary concepts but she attributed many of her challenges to content problems rather than linguistic problems.

> **Miyazaki**: I usually understand most of my English lectures very well. I do not face any language challenges. When I do not understand concepts or terms in lectures, it is not because I have problem with English but the content is simply difficult.

Nevertheless, it was not only proficiency but also prior content knowledge that mattered. For example, while Lisa (IELTS 8.0) had advanced English knowledge, her little prior content knowledge influenced her lecture comprehension. She argued that her learning could have been more successful if she had studied in the JMI course where she would be free from any form of linguistic issues.

> **Lisa:** Learning Chemistry is really enjoyable but also difficult because it is a completely new subject. I find it difficult to understand lectures when a complicated concept is introduced. I always wonder if I could understand better in Japanese.

Thus, this theme was one of the few challenges that was only prevalent among the L2 English students. Given that Miyazaki with higher proficiency could learn the subject without facing any linguistic problems, proficiency seemed to have played a pivotal role in mitigating this challenge. However, the level of prior content knowledge was also a crucial factor influencing the severity of this challenge. Despite having advanced English knowledge, Lisa could have performed more successfully if she had

studied in the JMI course where she could have compensated for the lack of her prior knowledge.

The final challenge facing the participants was to understand the professor's use of a plethora of unfamiliar technical terms used in lectures. An exploration of this theme based on proficiency revealed that even those with advanced English knowledge still faced this problem when attending lectures. For instance, Charlotta (IELTS 8.0, post-test 9.5) still grappled with the sheer number of unknown technical words introduced in lectures despite possessing a sound foundation in English.

> **Charlotte:** It is difficult to fully understand the lectures because I still do not know many technical words used by my professor.

Similar to Charlotta, Kiley (IELTS 7.0, post-test 4) and Ayaka (IELTS 7.0, post-test 8.5) also had difficulty in understanding vocabulary introduced in lectures. However, their challenges appeared to be affected by their previous experience of learning Chemistry through English rather than their proficiency. Ayaka from a Japanese-medium school had to relearn the Chemical elements in the periodic table in English. She struggled to follow her lectures without understanding these basic terms. Kiley also faced difficulty with technical terms but she had already learned many of these basic terms in English at high school.

> **Ayaka:** I learnt Chemistry in Japanese at high school, so I have to learn so many basic words used in lectures and all the Chemical elements in the periodic table all over again in English. It is time-consuming.

> **Kiley:** I do not understand many technical words when I attend lectures but it is helpful that I remember some of them from my high school Chemistry class. It would have been more difficult to study all the new words if I had not learned Chemistry in English.

Similarly, Miyazaki (IELTS 8.0, post-test 15) had previously learned Chemistry terms in English, which placed her at a considerable advantage in acquiring new terms covered in lectures.

> **Miyazaki:** I still need to learn new technical words introduced in lectures. But there are also many words that I already learned in my high school Chemistry class.

Thus, all participants irrespective of proficiency were confronted with a range of technical vocabulary in EMI lectures, and those with higher proficiency were no exception. In other words, understanding technical vocabulary used in lectures was a general academic challenge. However, the magnitude of their challenges was likely to depend on factors, such as previous experience of learning Chemistry through EMI. Those who had previously learned Chemistry through Japanese faced the significant additional difficulty of having to acquire technical terms in English. In contrast, those with previous EMI learning experience were well versed at learning through English. As their difficulties concerning technical vocabulary were stemming from the lack of their understanding of both vocabulary used to explain the lexically complex definition and the academic concepts attached to them, this relationship between lexical and disciplinary knowledge will be further discussed in Section 7.4.4 "Language-related challenges in EMI" in Discussion.

To summarise, I explored the second theme related to the lecture comprehension issues according to proficiency. The higher proficiency students generally reported fewer challenges than their lower proficiency counterparts. Table 36 below also highlights this pattern and shows the number of times each sub-theme was coded (references) per participant in the data according to IELTS level. However, the data has also identified some challenges facing all participants irrespective of proficiency. For example, Lisa and Emma (IELTS 8.0) despite their higher proficiency, still faced difficulty with acquiring technical vocabulary. This shows that while proficiency was crucial to alleviating the magnitude of some challenges, other factors might have accounted for a large proportion of their challenges (e.g. previous experience of learning via EMI).

Table 37: Matrices of challenges according to IELTS level

Participants Main & Sub-themes	EMI IELTS 6.5 (n=2)	EMI IELTS 7.0 (n=4)	EMI IELTS 7.5 (n=1)	EMI IELTS 8.0 (n=3)	EMI English L1 (n=3)
2. Comprehension of lectures	**12.5**	**12.2**	**8**	**4.0**	**1.3**
Low concentration	2.5	4	1	0.3	0.3
Humour and jokes	3.5	1.8	2	0	0
Content and language	3.5	3.6	3	2.7	0

Technical words	3	2.8	2	1	1

6.8.3 Final Grade

A third challenge facing the EMI students was to achieve a respectable Final Grade. This theme was further divided into three sub-themes: poor management of workload; low performance on the final course exam; low performance on the Comment Worksheet.

A first sub-theme was the students' difficulty in managing their workload. An exploration of this theme based on IELTS level revealed that this challenge was more prevalent among the less proficient English learners than their counterparts with higher proficiency. As an EMI course could be linguistically more demanding than its L1 counterpart, those with lower proficiency had to spend a longer study time and prepare more thoroughly for lectures when studying through English. In contrast, those with higher proficiency did not report this problem. For example, Hikaru (IELTS 6.5, Final Grade C) invested longer learning hours exclusively on the Chemistry course. Consequently, he was left with less time to revise for other courses, which resulted in lowering his Final Grade. In contrast, Emma (L1 English, Final Grade A) could focus on other courses which were "more demanding" than the Chemistry course.

> **Hikaru:** I spend most of my study time revising for Chemistry because it's in English. I'm also taking four other courses. I chose to study them in Japanese, but I still need to prepare a lot. I'm worried I might fail on my courses.
>
> **Emma:** I registered for five courses this semester. I usually spend most of my time reviewing my Biology course because I have a lot of assignments and tests. The content is more demanding than the Chemistry course because I haven't studied Biology before.

Nevertheless, possessing advanced English proficiency did not necessarily ensure that the students could manage their workload without any difficulty. Given that all participants enrolled in three or four courses on top of the Chemistry course, time management was a common problem for all students. For instance, a struggling learner, Maxwell (L1 English, Final Grade D) lacked adequate disciplinary knowledge and expressed unease at planning out his time effectively throughout the semester. He was confronted with a vicious circle where his insufficient content

knowledge resulted in undermining his passion for learning, which in turn led to a further loss of progress.

> **Maxwell:** This semester I'm taking four courses on top of the Chemistry course. I have problems with time management. I am taking far too many courses. I ended up not spending much time on Chemistry because I don't like the subject very much.

Therefore, the data broadly suggested a possible relationship between proficiency and the ability to manage workload. This is illustrated by the markedly different experiences expressed by Hikaru (IELTS 6.5) and Emma (L1 English). Hikaru attributed his workload problem to his low L2 proficiency. In contrast, Emma handled her heavy workload and allocated time to each element of courses effectively.

Table 37 below showcases the relationship between proficiency and the Chemistry course Final Grade, pointing to a stark contrast between those with higher proficiency (e.g. Miyazaki, Lisa) and those with lower proficiency (e.g. Hikaru and Ayaka). Most of the participants with IELTS 6.5 and 7.0 only achieved a "C" or below whereas most of their counterparts with IELTS 7.5/8.0 earned an "A". Nevertheless, it also became apparent that the challenges were associated with the constellation of other factors, such as course difficulty, professors, assignments, and prior disciplinary knowledge. As a result, Maxwell (L1 English), a demotivated and lower-ability learner, lacked the requisite disciplinary knowledge and struggled to manage his workload.

Table 38: EMI participants' Chemistry course Final Grade (FG) by IELTS level

English Proficiency (IELTS Scores) / Final Grade**	Low (C, D, E)	Medium (B)	High (A)
IELTS 6.5	Hikaru (FG: C)* Rina (FG: C)*		
IELTS 7.0	Kiley (FG: E) Marii (FG: C)* Ayaka (FG: C)*	Minori (FG: B)*	
IELTS 7.5			Tikamon (FG: A)
IELTS 8.0	Charlotta (FG: C)		Miyazaki (FG: A)* Lisa (FG: A)
L1 English	Maxwell (FG: D) Saika (FG: C)		Emma (FG: A)

* Those who participated in the first and second round interviews.

** A course Final Grade was awarded at the end of the semester on the basis of Course Work, Final Exam, Attendance, and the Comment Worksheet, ranging from A to E, A being the highest mark (distinction), E being the lowest mark (fail).

The second sub-theme was the students' low performance on the final course exam, which was one of the common challenges reported by all participants regardless of English language proficiency. For example, Marii (IELTS 7.0, Final Grade C) and Saika (L1 English, binlingual, Final Grade C) despite obtaining different levels of language proficiency, both expressed general challenges pertaining to the final exams (e.g. poor time management during the exam, low comprehension of key disciplinary concepts, lack of confidence, test anxiety, insufficient revision time, and critical reflection of exam questions).

> **Marii:** I did badly on the final exam. I did not know how to answer many questions. I panicked and ran out of time to finish the exam. I should have managed my time better to prepare for the exam more effectively.

> **Saika:** I did not understand some parts of the lectures and textbooks. I should have asked the teacher for help when I didn't understand them. There were quite a lot of questions about these concepts. I was also nervous before the test so it was difficult to concentrate on the test.

However, an exploration of this challenge based on IELTS level also supported the relationship between exam performance and English proficiency, which was further observed in Table 38 below. This table summarises the students' final exam scores based on IELTS level – that is, Miyazaki and Lisa (IELTS 8.0) earned 70% or more overall in their exams whereas none of their lower proficiency counterparts reached the same range of scores. The excerpts below also indicate that Rina (IELTS 6.5, Final Grade C) required longer time to write down answers in English. Hikaru also (IELTS 6.5, Final Grade C) struggled to read and comprehend each exam question accurately and quickly. As a result, he did not realise that an answer bank was provided at the end of the exam paper. Their linguistic challenges (e.g. low academic English vocabulary size) had lasted throughout the semester, leading to a slowing down of content coverage. It might well be that their exam performance was impeded by low linguistic level.

> **Rina (EMI):** It took me a long time to process what I wanted to write down in English. I would have liked to take the same test in Japanese.

> **Hikaru (EMI):** I had to spend a long time on reading instructions for each exam question and ran out of the time. I also didn't notice that I was given an answer bank for one of the questions. I could have saved a lot of time.

Thus, all students faced great difficulty with the final exam regardless of proficiency; yet there were certain challenges (e.g. misunderstanding of exam questions written in English) only reported by the lower proficiency students (i.e. Marii, Rina and Hikaru). Their challenges were linked to their linguistic problems, whereas Saika (L1 English) attributed exam challenges to non-linguistic factors (e.g. revision and preparation time, comprehension of concepts, motivation, prior content knowledge, learning strategy, and readiness for the academic subject).

Table 39: EMI participants' final exam scores by English proficiency

English Proficiency (IELTS Scores) / Exam Scores (ES)	Low (0%-60%)	Medium (61%-70%)	High (71%-100%)
IELTS 6.5	Hikaru (FE: 38.8%)* Rina (FE: 40.7%)*		
IELTS 7.0	Kiley (FE: 27.7%) Marii (FE: 34.0%)* Ayaka (FE: 36.7%)*	Minori (FE: 60.8%)*	
IELTS 7.5		Tikamon (FE: 64.8%)	
IELTS 8.0	Charlotta (FE: 40%)		Miyazaki (FE: 83.3%)* Lisa (FE: 70.7%)
L1 English	Maxwell (FE: 50.8%) Saika (FE: 36.5%)	Emma (FE: 66.2%)	

* Those who participated in the first and second round interviews

The third and final challenge facing the participants was their low performance on the Comment Worksheet, used by the professor to administer pop quizzes and gather student general comments and questions, which accounted for 15% of Final Grade of the Chemistry course. An exploration of this theme based on IELTS level revealed that the students' proficiency influenced their performance on the Comment Worksheet, which was also a pattern observed in Table 39 below. This table highlights that Miyazaki and Lisa (IELTS 8.0) were the only learners who could reach 18.0 or above overall in the Comment Worksheet. Those with lower proficiency were less likely to understand lectures sufficiently, solve comprehension quizzes, and write down comments and questions accurately and quickly in English. For example, Rina (IELTS 6.5, Final Grade C) attributed this problem to her lower lecture comprehension. Kiley (IELTS 7.0, Final Grade E) had difficulty in writing down

sentences quickly in English due to the short amount of time allocated (i.e. typically less than five minutes).

> **Rina (EMI):** I don't like these pop quizzes because I don't necessarily understand every part of my lectures sufficiently.
>
> **Kiley (EMI):** I always end up leaving only one or two badly phrased sentences because I need more time than five minutes to revise what is covered in my lectures.

On the contrary, their counterparts with higher English proficiency were generally more enthusiastic about the Comment Worksheet. For instance, Saika (L1 English, bilingual, Final Grade C) had ample academic English writing practice at her EMI high school. In the same vein, a higher-ability learner, Lisa (IELTS 8.0, Final Grade A) insisted on the importance of preparation and revision to achieve good lecture comprehension to perform well on the Comment Worksheet.

> **Saika:** I am happy with the Comment Worksheet system because it allows us to ask further questions. At high school we also had a similar system. My problem is to get good scores on the pop quizzes. I often don't understand lectures.
>
> **Lisa:** I like being given the opportunity to ask questions through the Comment Worksheet. But I also need read the textbook carefully being going to lectures. I would otherwise do badly on the pop quizzes.

Unlike Rina and Kiley, the higher proficiency students, Saika and Lisa were largely supportive of the Comment Worksheet. However, they also encountered some similar challenges as those expressed by their lower proficiency counterparts, such as difficulty in producing comments and questions, comprehending disciplinary concepts covered in lectures, and providing correct answers to comprehension check quizzes. Thus, while low proficiency impeded their performance on the Comment Worksheet, the evidence has also pointed to an array of other factors, such as previous exposure to EMI at high school, motivation, lecture comprehension, the amount of revision and preparation time.

Table 40: EMI participants' Comment Sheet scores by English proficiency

English Proficiency (IELTS Scores) / Comment Sheet (CS)	Low (0-10)**	Medium (11-17)**	High (18-25)**

IELTS 6.5	Hikaru (CS: 10.5)*	Rina (CS: 15.5)*	
IELTS 7.0	Kiley (CS: 6.5)	Marii (CS: 13.0)* Ayaka (CS: 13.5)* Minori (CS: 13.5)*	
IELTS 7.5		Tikamon (CS: 13.5)	
IELTS 8.0	Charlotta (CS: 6.5)		Miyazaki (CS: 18.0)* Lisa (CS: 19.5)
L1 English	Maxwell (CS: 7.5) Saika (CS: 2.0)	Emma (CS: 12.0)	

* Students who participated in the first and second round interviews
** Grading scale ranges from 0 to 25, combining a total score of five Comment Sheets, amounting to a maximum of five points.

Hence, the third theme regarding the students' Final Grade was explored according to proficiency in a cross-tabulation of coding frequency (Table 40). The data ascertained an intricate relationship between proficiency and challenges, whereby the students with higher proficiency faced fewer challenges than their counterparts with lower proficiency. The table also underpins this association by highlighting the number of times each of the sub-themes was coded per participant (references) based on IELTS level. While proficiency was key to higher levels of success, it also became apparent that the magnitude of their challenges were also affected by a myriad of other factors, such as motivation, learning strategy, course difficulty, previous exposure to EMI in high school, professors, assignments, prior content knowledge, and relationship with other course members. Thus, possessing advanced proficiency did not, by itself, lead to more successful performance.

Table 41: Matrices of challenges according to IELTS level

Participants	EMI IELTS 6.5 (n=2)	EMI IELTS 7.0 (n=4)	EMI IELTS 7.5 (n=1)	EMI IELTS 8.0 (n=3)	EMI English L1 (n=3)
Main & Sub-themes					
3. Final Grade	**7.5**	**6.9**	**5**	**4.4**	**3.3**
Managing workload for all the enrolled courses	3	3.6	2	1.7	1
Final course exam	2	1.5	1	2	1
Comment Worksheet	2.5	1.8	2	0.7	1.3

6.8.4 Adjustment to the new academic and cultural conventions of the Chemistry course

The fourth and final challenge facing the EMI students was to adjust to the new specialised disciplinary culture and academic conventions of the Chemistry course.

This theme was further divided into two sub-themes: adjustment to the new educational culture; and persistent problems.

A first challenge facing the EMI students was to conform to the new educational culture and disciplinary norms and practices of the Chemistry course. When this theme was explored according to proficiency, it became evident that this challenge was prevalent among those with low English proficiency. For example, Hikaru (IELTS 6.5, post-test 5) and Rina (IELTS 6.5, post-test 7.5) attributed low proficiency to issues associated with acculturation into the new academic conventions. Hikaru was generally not very confident in speaking in English, and this issue was further magnified by having to adapt to the new academic culture distinctive to EMI. The teacher often nominated students to share their opinions with other class members. Rina also had issues with spontaneous speech production in English. As a result, she often felt inferior to her peers who were more fluent English speakers.

> **Hikaru:** I am still not comfortable with the idea of thinking on my feet and sharing my ideas with the rest of my class. I was always quiet and passive at high school.

> **Rina:** Some of my course mates especially the international students have good English and ask many questions to challenge the teacher. That is how they learn overseas? I would like to do the same but I am worried about my English.

Although Tikamon (IELTS 7.5, post-test 12.5) had higher English proficiency than Hikaru and Rina, her limited prior EMI experience inhibited her motivation and willingness to participate in class discussion. This indicates that her readiness for EMI study was more likely to depend on factors, such as prior EMI learning experience.

> **Tikamon:** It is my first time to study in English. I do not like speaking in front of my course mates in English. I of course want to join class discussions but I feel very shy.

Exploring these excerpts suggests that high school educational culture and teaching medium influenced their preparedness for EMI education. The students faced new educational norms specific to EMI education, which were inherently student-centred, bidirectional, interactive and participatory. Consequently, those from Japanese medium schools faced a considerably more taxing induction compared to their

counterparts from EMI-based school systems, which provided more interactive educational cultures.

To illustrate, Miyazaki (IELTS 8.0, post-test 15) and Emma (L1 English, post-test 12.5) faced fewer problems adjusting to the demands of the EMI course. Having attended EMI high schools, they had already been versed in and favourably disposed towards EMI education. Nevertheless, the excerpts below highlight that they still experienced difficulties associated with learning a new academic discipline at university (e.g. longer class time, stricter evaluation criteria and more advanced disciplinary level).

> **Miyazaki:** My high school teacher always nominated us to speak in front of the class so I do not mind sharing ideas with my course mates. I feel sorry for my "Jun-Japanese" class mates (students from Japanese medium high schools) because they are not used to this teaching style. My problem is to keep focused on the entire 70 minute lecture because it is so much longer than my high school classes.
>
> **Emma:** I often ask questions in my chemistry class. At high school, it was common to ask questions and discuss ideas with my class mates and teachers. One big difference between high school and university is that university lectures are more difficult, fast-paced and longer. I also find it more difficult to receive good grades at university.

Therefore, while this challenge was prevalent among those with lower English proficiency (e.g. Hikaru and Rina), the school educational culture and teaching medium appeared to be the major factor influencing adjustment to EMI education. Compared to Miyazaki and Emma from English-medium schools, Tikamon, Hikaru and Rina without prior EMI experience faced a more testing transition into EMI. They had to learn to be critical, analytical, challenge what was being taught, and provide answers on the spot. Nevertheless, all participants, regardless of the school background or proficiency, still faced general school-university transition challenges, such as long lecture duration, independent learning, fast-paced lectures, and advanced disciplinary level.

The final sub-theme was that the students' adjustment issues were long-lasting rather than short-term teething problems. An exploration of this theme based on IELTS level indicated that this challenge was more prevalent among the lower proficiency students

than their higher proficiency counterparts, and was exclusively reported by the L2 English participants. For example, the lower proficiency students, such as Rina (IELTS 6.5, post-test 7.5) and Minori (IELTS 7.0, post-test 9.5) revealed that many of their linguistic problems reported at the start of the term persisted throughout the whole semester (e.g. small vocabulary size, slow reading speed, and grammar issue).

> **Rina:** I was never fully able to learn as efficiently as how I'd learn in Japanese. There are still so many English vocabulary and grammar I don't know.

> **Minori:** I spent a long time revising for lectures, but I couldn't understand many concepts. I gave up trying to learn as well as how I learn in Japanese.

In contrast, Emma (L1 English, post-test 12.5) and Miyazaki (IELTS 8.0, post-test 15) were confronted with a less onerous and more successful transition into the EMI course. Despite some common challenges posed by the school–university transition (e.g. longer lecture duration, advanced technical vocabulary), they rose to these challenges and ultimately fulfilled their desired goals of adapting to the demands of the EMI course by the end of the term. They attributed this progress to their pre- and post-lecture study rituals, coupled with effective learning strategies, and strong passion for Chemistry.

> **Miyazaki:** It used to take a long time to read the textbook at the beginning but I improved my Chemistry knowledge as the term went on. It gradually became easier to read and understand the textbook quickly too.

> **Emma:** I was overwhelmed with the amount of new technical terms I had to learn at the beginning of the term but I spent a lot of time revising thoroughly for lectures. By the end of the term, I successfully learned enough technical terms.

Nevertheless, obtaining advanced English proficiency did not by itself result in a trouble-free learning experience. For example, Charlotte (post-test 9.5, Final Grade C) was highly proficient (IELTS 8.0) but was never confident in reaching the same disciplinary comprehension as that of the L1 instruction, which stood in stark contrast to Miyazaki and Emma whose challenges completely dissipated by the end of the term. Charlotte was evidently neither as proactive nor motivated as Miyazaki or Emma in studying Chemistry which might have in part explained her relatively low content learning achievement. Thus, the data pointed to the array of factors that may influence the speed and ease of the transition into the EMI course.

> **Charlotta:** Chemistry is not the subject I have studied before so I still find it easier to follow lectures in my first language. I can understand individual science terms more efficiently.

Seeing that most of the lower L2 English interviewees (e.g. Rina, Marii) felt that studying through EMI was not as effective as through their first language, achieving a certain level of proficiency appeared to be a prerequisite for being successful for EMI study. However, those with higher proficiency (e.g. Miyazaki, Emma and Charlotte) revealed that a combination of EMI learning experience, motivation and rigorous study rituals ultimately led to their successful content knowledge development through EMI.

To sum up, table 41 below illustrates the relationship between challenges and proficiency and shows the number of times each sub-theme was coded (references) per participant based on IELTS level. With regard to the adjustment to the EMI course, several common challenges became apparent, some of which were only exclusive to the lower L2 proficiency students. Although all students irrespective of proficiency faced challenges associated with learning new academic subjects at university (e.g. advanced disciplinary level), the lower proficiency learners particularly struggled with the new educational culture distinctive to EMI (e.g. answering questions on the spot). Similarly, while the higher proficiency students felt ease at combating their school-university transitional challenges by the end of the term (e.g. acquiring the requisite technical terms), their lower proficiency counterparts encountered longer-lasting problems (e.g. reduction in content coverage and comprehension), admitting that their disciplinary understanding was never the same as that of their L1 instructed courses. However, the data has also pointed to the complexity of other variables which also affected their perceptions of ease and difficulties for their transitional challenges (e.g. motivation, previous EMI experience, self-study rituals, and prior content knowledge).

Table 42: Matrices of challenges according to IELTS level

Participants	EMI IELTS 6.5 (n=2)	EMI IELTS 7.0 (n=4)	EMI IELTS 7.5 (n=1)	EMI IELTS 8.0 (n=3)	EMI English L1 (n=3)
Main & Sub-themes					

4. Adjustment to the new content learning environment	6	4.1	2	2.4	1.7
Educational culture	3	1.8	1	1.7	1.7
Persistent problems	3	2.3	1	0.7	0

6.9 Summary of research question 3.2

In summary, the qualitative analysis has involved analysing the EMI students' perceptions of their learning experience according to the different levels of English proficiency. The data has offered insights to the realities of students' learning experience, indicating that learners experience greater challenges according to their English level. That is, the students with lower proficiency expressed considerably more onerous challenges than their counterparts with higher proficiency. Seeing that neither students with IELTS 6.5 nor IELTS 7.0 in the current study were able to reach higher content test scores of 12 point or above, it could be argued that English proficiency was paramount to higher levels of success in EMI.

Another noteworthy point was that the qualitative analysis has allowed an exploration of a number of possible explanations to cases where the participants with similar level of English proficiency still experienced markedly different challenges. For instance, Marii (post-test 1) and Minori (post-test 9.5) both possessed IELTS 7.0 but disclosed very different degrees of motivation for learning Chemistry, which could be the very reason for the differences in their learning experience. Minori was a highly motivated and also more able learner, which resulted in far fewer challenges than Marii. These contrasting experiences shared by Marii and Minori have clarified that some of the challenges (e.g. educational culture, reading speed, reliance on translation) were impacted not only by proficiency but also by myriads of other factors, including previous exposure to EMI education in high school, prior content knowledge, motivation, prior content knowledge, course difficulty, revision and preparation, and learning strategy. Therefore, while some of the challenges were more prevalent particularly among less or more proficient English learners, others were common across all participants regardless of English proficiency or whether they were learning Chemistry through English as their first or second language, as summarised below:

- Some challenges (e.g. concentration fatigue, general academic vocabulary, English grammar) were encountered only by the students with lower L2 proficiency (i.e. below IELTS 8.0).
- Some challenges (e.g. efficiency of content learning through EMI) were experienced by all of the students with L2 English regardless of their English proficiency. None of the L1 English students, however, expressed these challenges.
- Some challenges (e.g. learning new technical terms) were faced by all participants including both L2 and L1 English students (i.e. general academic problems).

Finally, this qualitative section concludes with an overview summary of the challenges discussed according to the main and subthemes associated with learning Chemistry through EMI (see Table 42).

Table 43: Overview of EMI challenges

First theme: Comprehension of the course textbook	
Vocabulary knowledge	Proficiency was crucial to mitigating vocabulary challenges. IELTS 8.0 and L1 students had difficulty with technical terms, not general academic English words.
Reading speed	There was a relationship between proficiency and reading speed, but prior content knowledge also affected reading speed.
Reliance on translation	There was a relationship between proficiency and reliance of translation but educational background and whether students were L1 or L2 English speakers also affected their challenges.
Inefficient reading	All L2 English participants reported this challenge regardless of proficiency. Conversely, the L1 English students did not report this challenge.

Second theme: Comprehension of lectures	
Concentration fatigue	This was a general academic challenge reported by all students, but the lower proficiency students suffered a shorter attention span than the higher proficiency students who could be selective about key information.
Humour and jokes	The lower proficiency students struggled more with this challenge due to their inadequate English proficiency and prior EMI experience in high school.
Content and language issues	Lower proficiency students struggled with having to deal with both language and content issues when listening to lectures through L2 English. Prior content knowledge however helped them overcome the challenges.
Technical vocabulary	This was a general academic challenge reported by all students, but those from a Japanese medium school background had to learn basic technical terms in English.

Third theme: Final Grade	
Managing workload	Lower L2 students had to spend longer hours on the EMI course. The higher L2 proficiency students could spend longer hours on other more demanding courses.

Final course exam	This was a general academic challenge reported by all students (anxiety, lack of disciplinary understanding, time management), but there were problem specific to low L2 proficiency students (misunderstanding concepts and exam instructions).
Comment Worksheet	Lower proficiency students had difficulty in formulating questions and solving quizzes. Higher proficiency students could easily share opinions with content teachers. However, previous exposure to EMI, motivation, lecture comprehension, revision and preparation time also mattered.

Fourth theme: Adjustment to the new content learning environment	
Educational culture	All participants faced general school-university transitional challenges. The school teaching medium was possibly a more important factor than proficiency.
Persistent problems	Most of the L2 English students' adjustment challenges never dissipated by the end of the term. Learning through their L1 was more effective than L2 English.

6.10 Chapter summary

The quantitative analyses firstly explored the relationships between the EMI students' English language proficiency (IELTS score), academic English challenges (EMI Challenges Scale), prior content knowledge (pre-test score) and academic success in EMI (post-test score). The finding seems to suggest that challenges were more likely to be apparent among the less proficient English learners than their counterparts with higher proficiency. It was however found that those who reported fewer challenges were not necessarily more successful in gaining content knowledge through EMI than their counterparts who experienced more considerable challenges. The interview data involving a range of student English proficiency was subsequently explored to elicit students' perceptions of content learning-related challenges when studying through EMI. This chapter concludes with the following five key findings:

1. *The importance of English language proficiency in mitigating challenges*
   All students interviewed endorsed the importance of English proficiency contributing to success in EMI, agreeing that the higher the English proficiency, the greater the ease with which the students experienced EMI.

2. *A myriad of factors affecting EMI challenges*
   The magnitude of challenges were affected not only by English proficiency but also by other factors, such as previous exposure to EMI education in high school, motivation, prior content knowledge, perceived course difficulty, revision and preparation, learning strategy, and relationship with other class members and lecturers.

3. *Non-linear relationship between challenges and success*

   The quantitative results (i.e. correlation and regression analyses) did not reveal a linear relationship between post-test scores (as a measure of academic success in EMI study) and challenges. The interview evidence concurred with the quantitative findings. Some students faced various linguistic challenges; yet they still rose to these problems, ultimately succeeding in content knowledge development through EMI. In contrast, others faced fewer linguistic problems but were not necessarily successful EMI learners.

4. *IELTS 8.0 as possible linguistic threshold*

   The L1 English and IELTS 8.0 students reported far fewer language-related challenges compared to their counterparts with lower proficiency. This finding could indicate that IELTS 8.0 or above could be seen as a cut-off line where the students started overcoming their language-related challenges in the current context.

5. *General academic challenges in the EMI Chemistry course*

   While some of the linguistic challenges (e.g. understanding L2 English grammar structures) were distinctive to the lower proficient English learners, some of the very high proficiency students (e.g. Miyazaki and Emma) still expressed general issues associated with learning new academic disciplines at university (e.g. time management) and learning Chemistry in general (e.g. difficulty in learning specific Chemistry terms), some of which therefore were general academic challenges facing all participants regardless of proficiency.

6. *The importance of prior content knowledge and proficiency in EMI content learning*

   Two separate simple linear regression analyses were conducted to reveal that proficiency and pre-test score statistically significantly predicted post-test score, explaining 20.6% and 17.2% of the variation in the post-test score, respectively. A multiple regression also accounted for 35.3% of the variation in the post-test score. Hence, both the quantitative and qualitative findings seem to suggest that the students with higher English proficiency and prior

content knowledge were more likely to be successful in gaining content knowledge through EMI than their counterparts with lower proficiency and prior content knowledge.

In conclusion, the current mixed methods chapter draws on both the direct measures (i.e. EMI Challenges Scales and post-test scores) and indirect measures (i.e. perceived challenges and content learning achievement) of success and challenges in EMI to capture the complicated nature of interactions between various factors that could affect the development of content knowledge through EMI. It highlighted that English language proficiency was key to success in EMI but other factors also emerged as crucial facilitators to mitigate the students' EMI challenges. Improving other factors was considered specifically crucial for the lower proficiency students as they may need to compensate for their inadequate English knowledge to effectively rectify their challenges. Unsurprisingly, improving these factors, whether it be increased motivation, longer study hours or stronger support from lecturers and course mates was not only a large contributing factor to success in EMI but also in their ability to overcome general academic challenges associated with learning new academic disciplines at university.

Chapter 7: Discussion of findings

7.1 Introduction

This chapter seeks to discuss and synthesise the findings of the current study in reference to previously reviewed literature in the field. In the previous three findings chapters I have offered new empirical insights into the complex and multifaceted nature of student content learning and how their achievement is affected by their chosen language of instruction. In this chapter, I set out to consider each of the research questions in turn, reviewing the results of my analysis and presenting my own interpretations in relation to a number of key findings to demonstrate how my research takes the field forward. Having evaluated the research questions individually, the chapter concludes with a summary of the theoretical and methodological contributions of the study, attempting to situate its findings within the fields of EMI, and within the broad literature outlined in Chapter 2.

7.2 Research question 1

7.2.1 Content learning differences between the mediums of instruction

The first research question proposed for this study was: "Is there an impact of the medium of instruction on students' content learning?" To engage with this research question, I first examined the extent to which the achievement of directly measured content knowledge differed between students enrolled in an English and Japanese mediated introductory Chemistry course. The findings of this quantitative evidence were presented in Chapter 4, revealing no measurable quantitative differences in the development of their understanding of Chemistry between the two groups when controlling for their prior content knowledge. The comparison of gain scores between the groups also showed no significant mean differences, indicating that the groups were not significantly different in terms of the gains in their content development. Such positive results concur with previous EMI research on educational outcomes (e.g. Park, 2007; EL-Daou & Abdallah, 2019; Dafouz et al., 2014), indicating that content learning through EMI in this research context does not appear to negatively affect participants' academic performance compared to that of their JMI counterparts.

In addition to the comparison of content test scores between the two mediums of instruction, the current study also explored the proportion of high and low performing students according to the top (the values above 14.6 out of 17) and bottom (the values below 5.40 out of 17) end of the post-test score distribution. The descriptive statistics revealed that 19.7% of the higher proportion of JMI students fell within this top range of the distribution compared to their EMI counterparts. This indicates that while the direct comparison of test scores show no overall mean differences, the observation of individual test scores provide some evidence to indicate differences in the proportion of successful students according to their mediums of instruction.

The qualitative findings of the current study pertaining to students' perceptions of content learning also add to this evidence, showing some differences between EMI and JMI students in terms of their perceptions of ease and challenges towards EMI study. Interview data point to various language-related challenges which are exclusive to those learning through L2 English (e.g. grammar comprehension issues in L2 English, inadequate understanding of exam questions), especially those with lower English proficiency. This suggests that EMI may negatively affect *some* participants' content learning experience according to English language proficiency. This negative evidence seems to corroborate with previous research evidence. For example, a Rapid Evidence Assessment (REA) carried out by the Applied Linguistics Research Group at Oxford, suggests that EMI is not detrimental to highly proficient students but it can negatively affect content learning of those with lower English language proficiency (Murphy et al., 2020). Similarly, many other qualitative studies incorporating interviews (e.g. Başıbek et al., 2014; Kim & Yoon, 2018) also indicate that students often express lower levels of satisfaction in content learning while also being doubtful about the improvement of their overall English language proficiency.

Interview findings also seem to suggest that EMI has a possible long-term negative effect on student' content learning. That is, while most of the JMI students are able to successfully overcome many of their academic challenges by the end of a semester (e.g. conforming to the specialised disciplinary conventions), some of their EMI counterparts feel unable to fully achieve the same level of content learning as that of their L1 medium instruction. This finding seems to suggest that many challenges that manifest at the start of the semester are long-lasting rather than short-term teething

problems. Interestingly, this negative finding stands in contrast to previous longitudinal research (e.g. Evans & Morrison, 2011; Pessoa et al., 2014) which has found that EMI students develop effective learning strategies (e.g. recording unfamiliar vocabulary in a notebook) and become more attuned to learning through L2 English, suggesting that their challenges largely dissipate by the end of an EMI program. One crucial methodological difference between the current and previous studies is that in the aforementioned studies took place over the course of one academic year rather than one semester, suggesting that undergraduate students in many EMI contexts may need a full academic year or even a longer period of time to fully become accustomed to EMI education. Further longitudinal research carried out over the course of an entire university degree is thus needed to better understand how participants in the current research context combat their challenges beyond the a single Chemistry program.

The central aim of this study was adopted from Macaro's (2018) cost-benefit analysis of EMI to investigate whether "learning content through L2 English leads to at least as good learning of academic content as learning content through the students' L1" (p. 154). In response to this central aim, the current study generated inconclusive evidence of a link between the medium of instruction and students' learning of Chemistry. That is, no discernible differences in the overall mean test scores were observed between the two mediums of instruction. At the surface level, therefore, EMI does not appear to have a negative effect on students' learning of content in this context. However, when other sources of data, such as interviews and descriptive statistics, are incorporated, more complicated and negative evidence emerged, which seems to suggest that EMI may have a long-lasting deleterious effect on students' content learning (e.g. slowing down of content coverage, reduction in the understanding of academic concepts). Thus, on the face of it, there are no differences between the two groups in terms of the average group scores of the content test result. However, many of the EMI students appear to experience greater difficulties in reaching the same learning outcomes, suggesting that pathways (process) to success (outcome) differ significantly between the two mediums of instruction.

In light of this conflicting evidence, the present study points to a wide variety of ways in which each stakeholder involved in the delivery of the EMI program could evaluate

this outcome, namely the university, students and teachers. From the viewpoint of the university, no discernible differences in the content test scores between the two groups could mean that EMI had no detrimental effect on students' academic performance. As the universities funded by the TGUP are expected by the government to further increase the number of EMI courses towards the end of the 2023 funding period (MEXT, 2018), this quantitative finding (outcome) may be a favourable result for the current case study university, which also strives to increase the number of EMI programs. Conversely, teachers (and students) may be more interested in understanding individual students' learning experience, thus focusing more on the qualitative evidence (process) than quantitative evidence (outcome), which suggested that EMI may have potentially long-lasting negative outcomes on academic performance. Both students and teachers should focus on the ways in which they could alleviate their challenges associated with EMI content learning. Therefore, stakeholders with different educational aims may interpret the current research evidence differently from their own viewpoint, highlighting the complexities surrounding who benefits and who loses when EMI is put into place.

The findings of the current study presented in Chapter 4 largely adopted a narrow definition of EMI, 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 that go beyond the mere 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). Thus, we now turn to examine "other benefits" of EMI which were identified in the current study, as will be covered in the next section, and whether the positive outcomes of EMI could exceed its long-term costs to content learning (e.g. reduction in content coverage).

7.2.2 Broadening the definition of success in EMI

By definition, the acquisition of content knowledge is considered to be a benchmark of successful EMI in the current context of Japanese higher education (MEXT, 2018). This corroborates the contemporary definition of academic success in higher education which involves the traditional sense of being proficient in academic

disciplines and the fundamental sense of being literate in the specialised academic discourses (Norris & Phillips, 2003). In this thesis, this definition of EMI with predominant content learning aims informed the methodology, which primarily focused on the academic outcomes of EMI in relation to Chemistry students' content test scores.

In this study, while gaining substantial content knowledge appeared to account for a significant proportion of whether EMI was successful in the Chemistry program, interview data revealed a range of other broader benefits, suggesting that successful EMI in this context concerns more than just exam scores. The findings suggest that participants who study Chemistry through EMI reported wider benefits, such as improved English proficiency, improved motivation, access to a broader range of learning resources through EMI, and long-term career advancement (see section 4.11 in Chapter 4). These perceived advantages have also been observed in other educational contexts in previous EMI studies (e.g. Curle et al., 2020; Radjai & Hammond, 2021), suggesting that many of these benefits (e.g. access to English medium resources) were tightly linked to EMI education rather than its L1 medium instruction counterpart. Although EMI learners in this study, especially those with lower English proficiency appear to have endured a considerably more taxing induction into content learning than their JMI counterparts as discussed above, they also seem to be able to counterbalance these negative effects of EMI with perceived benefits (e.g. gains in intercultural communication skills).

Thus, widening the definition of success in EMI to include its broader educational benefits invokes a theoretical discussion of whether the academic outcome of EMI should necessarily be compared with that of the L1 medium comparison group. Among the 23 EMI outcome studies reviewed in Chapter 2, 12 studies involved a control group of EMI with instruction conducted in the respective L1 of the context under investigation (e.g. Rashed et al., 2018; Dafouz et al., 2014; Yang et al., 2019). While control groups might be useful to reveal differences in content learning acquisition attributed to the medium of instruction, they do not tap into ancillary benefits gained only by the EMI group. This study highlights the fact that 'success' in EMI is unique in its own right, offering a range of additional educational benefits.

Accordingly, it may be more pertinent to evaluate the value of EMI in its own right, rather than an association with its L1 medium instruction counterpart.

To move EMI outcome studies forward in the field, more research is needed to better understand whether the advantages of EMI outweigh its disadvantages (e.g. slower reading speed) and to what extent EMI leads to these expected positive outcomes that go beyond the content learning outcomes. To evaluate the trade-off between benefits and costs of EMI, future research needs to widen the definition of EMI and include a larger number of outcome measures (e.g. long-term career advancement after receiving EMI education). This methodological suggestion is repeatedly mentioned 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).

7.3 Research question 2

7.3.1 Key contextual factors in EMI

The second research question was: "In what ways do the two mediums of instruction differ? What context-related differences do the students ascribe their challenges and achievement of content learning?" In order to respond to this second research question, the current study triangulated multiple data sources to achieve a holistic view of EMI content learning, unveiling a number of potential contextual and institutional factors that participants ascribed to their academic success in EMI. The findings presented in Chapter 5 revealed that while participants largely endured a testing initiation into content learning through EMI, they were also able to combat many of their challenges through accessing a range of institutional support provisions, such as: library resources; L1 medium materials and textbooks; the academic writing centre; the teaching assistant; the Comment Worksheet; and the in-sessional academic English language programs.

Such findings concur with many of the recent EMI studies (e.g. Galloway & Ruegg, 2020; Galloway & Rose, 2021) which have highlighted the importance of institutional support provision in assisting EMI university students in their academic work. Seeing that EMI learners face more onerous experiences depending on the level of their English language proficiency, many of these previous studies have largely focused on the role of academic language support provision, such as ESP and EAP courses.

Nevertheless, the current study draws on richer sets of data, including observations and interviews to shed new light on other types of support measures that go beyond the mere language support (e.g. the teaching assistant provision and L1 medium library resources). It is possible that these broader institutional support schemes may have accounted for an even larger proportion of positive factors that lead to success in EMI than the language support provision alone.

Besides the institutional support provision, the current study also highlights a myriad of other context specific factors, such as small class size, student-teacher interaction, peer feedback, group discussion, mixed ability learning environment and participatory educational culture. These factors are highly context dependent but repeatedly shown to be paramount to students' academic success in both this study as well as many other broader education studies (e.g. Spyros & Wei, 2012), whether it be primary, secondary, or tertiary educational sectors. Thus, the present study demonstrates that participants navigate their EMI studies through a combination of effective institutional support provision (e.g. academic language support), teacher support and – not least – supportive and collaborative peer networks.

There has been a growing body of research in the area of student-teacher interaction in EMI classrooms (e.g. Lo & Macaro, 2012 in Hong Kong; Macaro & Tian, 2020 in China; Macaro et al., 2020 in Italy), many of which seem to suggest that interaction in EMI classrooms is likely to be dominated by teachers, leading to low levels of student participation. Unlike this common assumption that EMI lectures are unidirectional, the current study showed that the EMI content teacher achieved interactive lectures. While the study identified a number of factors (e.g. the language proficiency of the students, whether the teacher valued pedagogical benefits of interaction for content learning, and the educational culture being introduced), one key factor driving the frequent teacher-student interaction in the current study appeared to be small class size. Both the observation and interview data identified small class size as a key factor which enabled the EMI content teacher to adopt more interactive lectures. That is, in the class size of 18 students, the EMI students were easily able to interact with both the teacher and peers in classroom interaction.

In contrast, in the large class size of 31 students, the JMI teacher, despite willing to introduce more interactive lectures, was restricted to adopting more monologic

teaching. The fieldnotes also revealed that students were seated distant from each other in a lecture theatre, which was another obstacle for opening up more student participation in classroom interaction. Thus, the current study offered some evidence that small class size was a key driver behind the high level of student-teacher interaction in the present EMI lectures. As the pedagogical benefits of interaction in EMI settings have been theoretically justified both in terms of content learning (Alexander, 2004) and language learning (Gass, 2018), this frequent interaction may have, therefore, contributed to the conditions that underpinned success in content learning for the EMI students.

Another key contextual factor to consider is that the present study was situated in an introductory Chemistry program primarily designed for first year students rather than second, third or fourth year students. Participants regarded that the magnitude of their challenges would become far greater when progressing to more advanced level programs where more linguistically and disciplinarily demanding assessments would be introduced, such as laboratory reports, essays, presentations, tutorials, and seminars. In the same vein, in order to embrace ecological validity, the EMI content teacher devised the content test specifically for the purpose of this study, which consisted of authentic questions closely mirroring the course learning objectives. Consequently, the content test primarily comprised of single-word questions, multiple choice questions in addition to tasks asking participants to draw chemical structures and solve chemical equations. In other words, the language of instruction (whether learning took place in Japanese or English) may have had greater effects on students' academic performance if the study had been situated in a more advanced level program. Students would have been evaluated based on more linguistically demanding tasks, such as oral presentations, essays and laboratory reports. Previous EMI research into science education (e.g. Pun & Jin, 2021; Eriksson, 2018) has found that science students express a greater challenges associated with productive tasks (e.g. presenting information to laboratory technicians, communicating with scientists and practitioners, and producing written course assignments) than receptive tasks (e.g. listening to teacher's explanation of a science topic, reading lesson materials). Thus, the current research context in which students were rarely required to demonstrate productive skills to undertake discipline-specific activities might have in turn limited the type of challenges which were observable to the researcher.

To summarise, in response to the second research question, this section has outlined various key contextual and institutional factors which appear to ensure a highly facilitative environment for content learning in EMI. It is worth emphasising that the findings of the present study are limited to the current context in which students were given access to a range of institutional support schemes and context specific factors greatly facilitated their EMI studies. As noted by Galloway and McKinley (2019), EMI research should take a contextually sensitive approach that accounts for individually unique research contexts, according to factors, such as institutions, countries, course levels, student/teacher profiles, and academic disciplines. More research is thus needed to examine what contextual and institutional factors are held accountable for successful content learning in EMI across different research settings, whether it be rigorous language support provision, smaller class size or frequent teacher-student interaction.

7.4 Research question 3

7.4.1 English proficiency and key learner factors in EMI

The third research question for this study was: "Are there relationships between the EMI students' English language proficiency, academic English challenges, prior content knowledge, and content knowledge achievement?" In response to this third research question, the study explored the qualitative and quantitative relationships between key learner variables in EMI. Interview findings presented in Chapter 6 revealed the intricate relationship between English proficiency and academic success in EMI, suggesting that those with lower English proficiency face a more testing transition into EMI than their counterparts with higher proficiency. They spend longer hours on preparation and revision for lectures to increase their likelihood of success, including looking up unfamiliar vocabulary, and reading a translated version of course materials and textbooks. Much of the research reviewed in Chapter 2, especially those incorporating perception-based instruments, have also identified clear differences in students' learning experience according to the level of L2 proficiency (e.g. Evans & Morrison 2011; Lei & Hu, 2014).

The present study also incorporated regression analyses to explore English proficiency as a predictor of academic success in Chapter 6. The linear regression

findings of the current study show that English proficiency measured using IELTS accounts for 20.6 % of the variance in Chemistry content test scores. Such findings appear to be consistent with some of the previous studies outlined in Chapter 2, for example, with Rose et al. (2019) showing a similar R^2 value of 0.25 for a Social Science subject, such as International Business, and Sánchez-Pérez (2021) showing a slightly higher but similar R^2 value of 0.48 for a STEM subject, such as mechanical Engineering. However, other studies have shown a much smaller R^2 value or even a non-significant finding, for example, with Cho and Bridgeman (2012) revealing a smaller R^2 value of 0.04 in a large-scale study involving students from ten universities (n= 2,594), and Yan and Cheng (2015) discovering that English proficiency was not a significant predictor of undergraduate GPAs (n=138). In light of these conflicting results, while English proficiency was a significant factor facilitating or hindering successful content development through EMI, different research has adopted different English language proficiency measures (e.g. TOEFL, IELTS), academic outcome measures (e.g. GPA or final course test scores), student demographics (e.g. undergraduates or postgraduates), and academic disciplines (e.g. STEM or Humanity). Therefore, these differences in methodology highlight the need to interpret each of the research findings carefully according to their highly contextualised research contexts and methodology.

The current study also explored the extent to which English proficiency predicted academic success in EMI while taking other factors into account, such as prior content knowledge and the magnitudes of language-related challenges. The results of the multiple regression analyses presented in Chapter 6 show that the model including English language proficiency and pre-test scores as independent variables accounted for 35.3% of the variance in the Chemistry content test score, which largely mirrored those reported in other EMI prediction studies (e.g. Curle, Yuksel, Soruç & Altay, 2020; Thompson et al, 2021) which also controlled for key learner variables, such as L2 motivation, self-efficacy and previous EMI learning experience. Interview findings in this study also suggest that the possible interactions between L2 proficiency and other learner factors impinged on academic performance in EMI, such as a high degree of motivation, rigorous study rituals, and previous content knowledge. In other words, in order to achieve greater ease towards EMI studies, students may have to compensate their lower language skills with other skills beyond those assessed by

standard proficiency tests, such as independent learning hours, learning strategies, prior content knowledge, and previous EMI experience. Recognising the constellation of these key learner factors, recent studies (e.g. Thompson et al., 2019) suggest that EMI students should improve both their academic English language skills and other broader academic skills to better prepare for EMI studies (e.g. confidence in learning content through EMI) via language support courses, such as EAP and ESP courses. Thus, this study supports previous research which suggests that success in EMI may be largely dependent on the complex interactions of all manner of key learner factors.

7.4.2 English language threshold in EMI

As this study highlights, and as an increasing number of research studies have suggested (e.g. Martirosyan et al., 2015; Hu & Wu, 2020), EMI students should reach a certain level of English language proficiency to undertake content learning sufficiently. Despite this significant role that L2 proficiency plays in content development in EMI, quantitative evidence of the current study appears to detect no clear discernible linguistic threshold at which students stopped reporting language-related challenges and achieved a higher level of performance. Nevertheless, when exploring EMI students' challenges according to the level of L2 proficiency, there appeared to be some interview evidence to suggest that students with IELTS 8.0 or more are more likely to experience greater ease towards studying through EMI than their counterparts without IELTS 8.0. In particular, certain language-related challenges (e.g. reliance on translation, inadequate general academic vocabulary, and grammar comprehension issues) were only reported by those below IELTS 8.0, indicating that none of the participants reaching IELTS 8.0 experience these academic English language-related issues. This result echoes previous findings that L2 English students with average IELTS band score of 7.0 (ranging between 6.5 and 7.5) are still not able to fully combat their linguistic challenges in English medium studies in higher education in the UK (Trenkic & Warmington, 2019), suggesting that IELTS 7.0 in their research is still not high enough for learners to fully alleviate their challenges in an English medium degree program. In the current study, interview evidence suggests that studying through EMI requires a high level of L2 proficiency and those without reaching a certain proficiency level can face a considerably more taxing induction into EMI studies than those with higher L2 proficiency. However, it

is not clear from the current evidence as to where exactly this threshold may lay based on students' perceptions of what constituted as "enough proficiency". Accordingly, in addition to qualitative research, more quantitative research is needed to examine where exactly the discernible differences may occur in students' perceived level of challenges according to L2 proficiency.

7.4.3 Challenges and academic success in EMI

In response to the third research question, the current study has also explored the relationship between challenges and academic success in EMI. Interestingly, the correlation and regression analyses of the study did not necessarily detect a linear relationship between language-related challenges measured using the EMI Challenges Scale and post-test scores as a measure of academic success in EMI. That is, successful students who performed well on the content test did not necessarily experience greater ease with EMI studies. Interview findings also found that participants, regardless of the level of their challenges, can overcome their challenges and ultimately succeed in content learning through EMI. In contrast, those who faced fewer linguistic challenges were not necessarily successful EMI learners, in terms of their content test scores. Such findings seem to stand in contrast with previous research (e.g. Kamasak, Sahan & Rose, 2020) which has suggested that a greater magnitude of challenges results in lower level of academic success in EMI, demonstrating a strong positive correlation between EMI challenges and success. These contradictory findings suggest that while content test scores served as a proxy for academic outcomes, the level of students' perceived challenges was affected by a combination of many factors (e.g. motivation, prior content knowledge, and L2 proficiency). Thus, the current study suggests that students with certain learner characteristics (e.g. higher L2 proficiency, motivation, prior content knowledge) might perform particularly well irrespective of their perceived levels of challenges in EMI.

In light of this non-significant finding between challenges and success, one of the explanations lies in the way in which the magnitude of students' challenges was operationalised in the current study. In the quantitative research, challenges were measured using the EMI Challenges Scale, which was originally developed by Evans

and Morrison (2011) in the English Language Centre who researched on L2 English students' language-related challenges in Hong Kong, where their focus was to conduct needs analysis to better understand how they could address and support their students' language problems. This indicates that items included as part of the questionnaire were largely connected to challenges associated with academic English language with an extensive focus on tasks connected to Social sciences and Humanities, such as essay writing and oral presentation. In this study, the questionnaire was localised (see section 3.9.4 in Chapter 3) to exclude items that were not relevant to the current research context of the Chemistry program (e.g. writing a bibliography and references section). However, many of the questions that were geared towards challenges more prevalent in Social sciences and Humanities may not have been able to capture challenges in the current STEM discipline. More context specific research instruments (e.g. EMI challenges questionnaire specifically designed for STEM subjects) should be devised to identify more accurate accounts of discipline specific challenges in future EMI research.

Similarly, another potential explanation for the non-significant finding could be the gap between students' perceived challenges and actual performance. This study triangulated perception-based data and directly measured test scores to better understand the potential impact of medium of instruction on students' content learning in the Chemistry program. The study shows that there is not necessarily a clear alignment between participants' actual performance and perceptions of ease towards EMI studies. For example, interviews showed that Saika, who learned though L1 English, was largely enthusiastic about her own perceived academic performance, but her actual post-test score of 7.5 was below the group mean score of 10.01. Her final exam score and final grade for the Chemistry course were 36.5% and "C", respectively. This inconsistency between students' perceived and actual performance seems to indicate that some of the L1 instruction students may have been more inclined to report greater ease than difficulty, speculating that learning through L1 would ensure trouble-free performance. Hence, using both perception-based data and directly measured test scores offers a more solid and holistic picture of students' learning experience, and can potentially explain why students' perceived challenges might not have been correlated with their actual performance in EMI.

Thus, while students expressed different types of challenges according to L2 proficiency, the level of their challenges do not necessarily determine how successful they are in terms of their final exam or content test scores in the Chemistry program. In the context of the current study, the magnitude of students' challenges and success in EMI seems to be largely dependent on an array of other key learner factors.

7.4.4 Language-related challenges in EMI

The findings of this study presented in Chapter 6 also offer some evidence that participants learning Chemistry through L2 English experience challenges due to their inadequate English language proficiency. When exploring their challenges according to L2 proficiency, it became clear that learning through L2 English exacerbated the magnitude of their challenges in EMI. The current study identified academic English language-related challenges which were exclusively reported by those learning through L2 English. For example, participants had to search for general academic English words, such as "involve", "polarised", "charged" and "law" which frequently appeared in their textbook. Similarly, those with IELTS 6.5 had particular difficulty in understanding humour and jokes used by their EMI content teacher, and understanding exam questions and instructions accurately and quickly. Such findings seem to support previous EMI studies (e.g. Aizawa et al., 2020; Vladusic et al., 2016; Kim et al., 2017), which have found that learning through L2 English aggravated the magnitude of students' academic challenges due to their language-related challenges (e.g. difficulty in writing essays and delivering oral presentations and participating in tutorial and seminar discussions). Interviews of the current study also revealed that vocabulary-related issues in part stemmed from the opaque English language spelling system. Participants were unable to identify or look up key technical words used by their professors during EMI lectures due to their difficult spelling, such as "tetrahedral molecular" and "electron configuration".

Thus, this study showed that learning through L2 English inflicted additional language-related challenges on EMI students, revealing a number of linguistic challenges only reported by those learning through L2 English. Thus, these students were likely to grapple with a "double whammy" of content and linguistic issues. The possibilities of "dual loss" were repeatedly discussed in previous studies, such as

Yang (2015) and Kim and Yoon (2018) who suggested that students were not only proficient enough in the English language to achieve sufficient understanding of academic concepts in their chosen academic disciplines but they were also highly skeptical about the expected positive impact of EMI on enhancing their overall English language proficiency.

So far I have argued that the magnitude of students' challenges in EMI are negatively affected by the low level of L2 proficiency. However, it is not necessarily straightforward to determine whether the challenges stem from the language issues or content issues in an EMI setting where learners grapple with the new academic concepts while also having to overcome an additional burden to carrying out their studies in L2 English.

Macaro (2020) observed the difficulty in separating general academic vocabulary from technical vocabulary, suggesting the need to unpack the notion of "unknown lexical items" (p. 269). The current study revealed that students had difficulty in learning technical terms, such as "tetrahedral molecular" and "electron configuration" in their EMI lectures. As Macaro (2020) has observed, one of the causes of their difficulty in this study was stemming from their lack of vocabulary knowledge as these terms contained particularly difficult spellings and unfamiliar lexical items (such as "tetrahedral"). Another cause of their challenge was associated with understanding scientific concepts attached to the terms. Macaro (2020) suggests that achieving a "masterful" understanding of a definition of new technical terms involves understanding the definition of the terms, further unknown lexical items used to explain the definition, and all the other non-discipline-specific items (e.g. "charge" and "law") that contain a discipline-specific meaning. Once the terms are introduced in lectures, students are then expected to explain the lexically complex definition in their own words and use the technical terms in their own essays and presentations, which would require the use of further technical and academic language.

Thus, it is the combination of both lexical and disciplinary knowledge that is required to fully master the acquisition of new technical terms. However, in EMI settings, students are presented with an extra linguistic hurdle coupled with a demanding task of learning a new academic discipline. As Macaro (2020) observed, it might be the

language used to explain the scientific concepts that is the causes of challenges for many of the L2 speakers of English in the study.

In the current study, students had difficulty with technical terms (e.g. “tetrahedral molecular” and “electron configuration”) because these words contained particularly difficult spellings and unfamiliar English lexical items (such as “tetrahedral molecular”). However, the reason for their challenge might have also stemmed from the difficulty associated with the complicated academic concepts behind these terms as well as the discipline-specific meanings attached to them. Therefore, I now turn to discuss discipline specific challenges below to establish that content and language issues are inextricably intertwined in EMI.

7.4.5 Discipline specific challenges in EMI

So far a large proportion of previous EMI research (e.g. Yıldız, et al, 2017; Siegel, 2020) has mainly focused on students’ general academic English challenges (e.g. understanding teachers’ accents, and acquiring general academic vocabulary), rather than discipline specific challenges. Nevertheless, the findings of this study, as one of its contributions, unveiled a number of students’ challenges specific to the learning of Chemistry in EMI. The current study has thus challenged the notion that EMI students faced challenges simply because content learning occurred through their L2 English. When exploring their challenges according to English proficiency, it became clear that those learning through L1 English also experienced many discipline specific challenges, such as acquiring an abundance of technical vocabulary and solving chemical equations. In particular, the study found that lacking technical vocabulary in English presented challenges for both L1 and L2 English students to solve chemical problem sets, such as calculating equations and electron configurations (e.g. 1s2 2s2 2p6 3s2 3p1) and determining the valencies of Chemical elements (e.g. Neon=0, Phosphorus=3). Similarly, students also lacked a rich repertoire of vocabulary which contained both scientific and everyday meanings used specifically in the science domain, such as “shell”, “element”, “activity”, “stability”, and “capacity”. Such findings concur with previous research into vocabulary issues in EMI. For example, Cassels and Johnstone (1983) and Gardner (1972) examined secondary school students’ comprehension of scientific vocabulary in EMI, revealing that many of their

L1 English EMI participants encountered difficulty in understanding non-discipline specific words that also contained scientific meanings, such as “evacuate”, “sequence” and “accumulate”. Similarly, Chan (2015) in Hong Kong also found that their L2 English EMI participants faced both the general language-related challenges and challenges specific to the learning of mathematics, in particular technical vocabulary distinctive to its academic domain (e.g. “equilateral”, “equiangular”, “isosceles” and “isometric”), showing that the major challenge facing L2 English learners in many EMI settings is understanding the discipline specific language used in science discourse.

Thus, while the low level of English proficiency was still a significant barrier to academic performance, learning Chemistry at tertiary level required highly specialised disciplinary and lexical knowledge (e.g. tetrahedral molecular), and there were many discipline specific challenges which were common for all participants irrespective of English proficiency, even for the L1 English participants. Thus, the discipline specific challenges were not necessarily confined to L2 English learners. Seeing that each academic discipline accompanies different types of challenges, future research is needed to explore more discipline specific challenges according to each of the different academic disciplines.

To summarise, EMI students struggled to acquire content knowledge due to their low English language proficiency. Nevertheless, Chemistry as an academic subject itself was also a significant factor aggravating the magnitudes of their challenges, suggesting that language and content challenges are inextricably intertwined in EMI.

7.4.6 Factors associated with challenges in EMI

This study has pointed to three particular factors contributing to the language-related and discipline-specific challenges facing students in the current EMI context, namely 1) the absence of explicit English language criteria, 2) the absence of language development in the overall instructional aim of EMI, and 3) the contextual differences in how support provision was put into place in the EMI and JMI programs.

Firstly, interviews and observations unveiled that the current case study university published neither clear nor explicit English language requirements for its EMI degree programs. It was largely left open for individual students to decide whether or not they should enrol in the EMI program, which in turn resulted in admitting those with potentially inadequate levels of English proficiency for their chosen academic subject. While the EMI Chemistry teacher encouraged those who were not confident in their language ability to register for the JMI program at the start of the semester, those with IELTS 6.5 (Hikaru and Rina) could still undertake the EMI program. As a result of their inadequate L2 proficiency, the study repeatedly showed that Hikaru and Rina struggled with the language issues in their EMI studies (e.g. lack of academic vocabulary, reading comprehension issues). The absence of such explicit language criteria for EMI degree programs has been observed in previous EMI studies (e.g. Aizawa & Rose, 2019). Ali (2013), for instance, identified vague criteria for students' English language proficiency on one of the EMI universities' policy documents in Malaysia, claiming that students must have achieved "excellent English proficiency", and "the ability to communicate in English through their academic courses" (p. 80) in order to be admitted to EMI programs. More recently, in conducting content analysis of websites of purposively sampled universities in 52 "Official Development Assistance" (ODA) countries, Sahan, Mikolajewska, Rose, Macaro, Searle, Aizawa, Zhou and Veitch (2021) found that only half of the universities (n=57) reported clear English language proficiency requirements for entry to EMI programs on their admission websites, which accounted for 54.8% of the 104 websites they reviewed in their study. They also found that universities reviewed in their study reported English language criteria on their websites in a variety of ways, some of which were standardised international English language proficiency tests (e.g. IELTS, TOEFL), while others were high school grades or more localised placement test scores. Therefore, these studies have found a lack of consistency in how universities around the world publicise their language requirements, indicating a need to introduce more explicit and clearer language criteria for EMI programs. It is also important to offer instruction as to how students can compare their test scores with the approximate equivalent score of another test by offering clear score conversion tables. In the current study, the absence of such clear language requirements contributed to the admission of students with inadequate English proficiency who encountered the greater level of challenges in EMI.

Secondly, there is often no explicit focus on language development in the overall instructional aim of EMI (Airey, 2012), which differentiates itself from other comparable terms, such as CLIL which often accompany the dual language and content goals. Some empirical attempts have been made to document students' English language development through EMI due to their long-term exposure to English (e.g. Rogier, 2012) but little empirical evidence for this yet exists. In the current study, EMI by its very definition paid little to no attention to students' language development (MEXT, 2015). At the research site university, it was assumed that language learning should occur prior to undertaking EMI programs. As a result, language learning was not the explicit aim in the EMI Chemistry course. Interview evidence presented in Chapter 5 also confirmed that the EMI content teacher in the Chemistry program was only responsible for the delivery of content knowledge and hardly regarded himself as a language teacher, paying little attention to students' language-related challenges or language development. This detachment from language learning and teaching was one of the very reasons which led to students' language-related challenges in the current study. As it was not the expectation to improve English proficiency while learning through EMI, learners should be encouraged to undertake content courses with language support courses built into their regular content programs to better combat their proficiency concerns.

Finally, the third reason for the prevalence of students' challenges related to the differences in how certain institutional support provision was integrated into the EMI and JMI programs. While the EMI and L1-instructed parallel programs were supposedly identical, certain support schemes were only available to one of the programs. In particular, there was a practical issue in recruiting a bilingual graduate teaching assistant who could assist students in both Japanese and English. Interview data presented in Chapter 5 noted that the JMI content teacher recruited the teaching assistant to support the JMI program. Although the curriculum of the parallel course was identical, it was not possible to offer the same support provision across the two programs, indicating the real-world complexities in offering a parallel degree program. In this context, the content teachers were not able to arrange a bilingual teaching assistant who could support both the programs. Thus, the lack of support

provision in the EMI stream might have been another reason which contributed to the greater level of students' challenges.

To summarise, this section has discussed key factors associated with challenges in EMI. Due to the absence of clear language requirement for EMI studies, students with the insufficient L2 proficiency in this context still made their way into the EMI degree program, resulting in the greater level of language-related challenges. Similarly, due to the predominant focus on the development of content knowledge, little focus was placed on the language learning aim, which meant that students received little language support from their content teacher in the EMI program. Furthermore, EMI students were not able to access the teaching assistant because the university was not able to make a practical arrangement to offer certain institutional support scheme, which also contributed to the greater level of challenges.

7.5 Chapter conclusion

This chapter has sought to integrate and synthesise the results of the current study presented in Chapters 4 to 6 to highlight the significance of the findings and present my interpretations in reference to previously reviewed literature in the field. The discussion of the results has been organised under various findings of the study. I conclude the chapter with a summary of the following empirical contributions of the study in an attempt to situate its findings within the research field of EMI, outlined in Chapter 2.

1. Wider conceptualisation of success in EMI
   EMI offers broader educational benefits that go beyond mere content learning outcomes, suggesting the need to widen the conceptualisation of success in EMI in future research.
2. Long-lasting challenges in EMI
   Students in this study were not able to fully overcome many of their challenges by the end of the semester. Longitudinal research is needed to observe how students overcome their challenges beyond the current program.
3. Language and content challenges in EMI
   While learning through L2 English negatively affects EMI content learning,

the learning of Chemistry itself also introduces discipline specific challenges. Students in such contexts may face a "double whammy" of content and language problems in EMI.

4. English proficiency and key learner factors in EMI
   While L2 proficiency is key to success in EMI, a constellation of other learner factors play a vital role in leading to greater ease in EMI.
5. English language threshold in EMI
   While those with IELTS 8.0 seem to fully overcome their challenges, this study does not provide sufficient evidence to ascertain any clear proficiency threshold where challenges start to level-off and/or dissipate.
6. Challenges and academic success in EMI
   The level of perceived challenges does not necessarily determine success in EMI. That is, successful students who perform well on the content test do not necessarily experience greater ease with EMI.
7. Key contextual and institutional factors in EMI
   Students ascribe various contextual and institutional factors to their academic success in EMI. Students benefit from institutional support (e.g. language support provision) and context specific factors (e.g. L1 medium resources).

In light of methodological contributions, this study is one of the first attempts in the field of EMI outcome research to conduct a longitudinal, quasi-experimental study via both perception-based and direct measures. Consequently, the methods used in the current study improve on the previous research literature in the following ways:

1. Longitudinal design to explore experiences over the duration of a an EMI course (three months)
2. Direct measures (pre-post content knowledge test)
3. Triangulation of data sources to assess student learning (pre-test scores, post-test scores, gain scores and final grades)
4. Direct comparison group (L1 medium instruction group)
5. Interviews and observations as a follow-up method of direct measures to capture quasi-experimental differences
6. High level of ecological validity (authentic EMI and JMI programs)

Finally, the study emphasises the prevalence of students' challenges in EMI. As a large proportion of factors contributing to successful content learning is a good command of the English language, the study highlighted the importance of clear English language requirement for EMI studies and the rising demand for well-integrated English language support, organised alongside their regular EMI content courses.

Chapter 8 – Conclusion

8.1 Introduction

The current study examined whether the learning of Chemistry through EMI influenced students' achievement of content learning. In doing so, I undertook a direct comparison between students studying introductory Chemistry in an EMI course and those studying the same curriculum in a JMI course. The essential conclusion of the study is that while there were no quantitative differences in content test outcomes according to the medium of instruction, qualitative data showed that there were clear differences in their perceived learning experience between the two groups, highlighting that outcome (directly measured content test scores) were the same but pathway (perceived learning experience) were different according to the medium of instruction. The study also found that EMI students with lower English language proficiency faced considerably more onerous challenges than their counterparts with higher proficiency when learning through EMI. Meanwhile, an array of other factors also emerged as drivers of success in EMI, such as previous EMI experience, motivation, revision and preparation, and learning strategy. In this chapter, I critically evaluate the results of my analysis and discuss the extent to which the findings provide satisfactory answers to the research questions. The chapter then presents some methodological and theoretical reflections on the limitations of this study. I then continue the chapter with recommendations for possible areas of future research, before discussing some pedagogical implications for curriculum planning in EMI programs. Finally, the conclusions drawn from the study will be presented at the end, highlighting the contribution of this thesis to the broader field.

8.2 Research question one

The first research question for this study was: "Is there an impact of the medium of instruction on students' content learning?" This research question consisted of two sub-research questions, and will be addressed separately in section 8.2.1. and 8.2.2.

8.2.1 To what extent does the achievement of content learning of the EMI and JMI students differ?

This first sub-research question examined the extent to which directly measured test scores varied between students enrolled in an English and Japanese medium course.

8.2.1.1 EMI and JMI students improved their content knowledge over the course of one semester

Firstly, pre-test scores and post-test scores for EMI and JMI students were compared separately using a Wilcoxon Signed Rank test. The result showed that EMI students improved scores by 50.5% (from Mean 1.80 to 10.40) and JMI students improved scores by 41.8% (from Mean 2.54 to 9.64). This indicates that EMI and JMI students on average both improved their content knowledge over the course of one semester.

8.2.1.2 There were no differences in prior content knowledge between EMI and JMI students

Secondly, pre-test scores between EMI (Md=0) and JMI students (Md=2) were compared using a Mann-Whitney U Test. The result suggested that the two groups were not statistically significantly different in terms of their prior content knowledge at the start of the semester.

8.2.1.3 There were no differences in post-test scores between EMI and JMI students

Thirdly, post-test scores between EMI and JMI students were compared while controlling for the pre-test scores. The result of an ANCOVA showed no significant differences between the two groups in terms of their post-test scores. It also showed that only 2.7% of the variance in the dependent variable (post-test scores) was explained by the independent variable (medium of instruction). This result suggests that the academic outcomes between the two groups did not significantly differ at the end of the semester when accounting for their prior content knowledge.

8.2.1.4 There were no differences in gain scores between EMI and JMI students

Fourthly, the gain scores (gain = post-test–pre-test) were compared between EMI students (M=8.59) and JMI students (M=7.10) using an independent samples t-test. No significant differences were found in gain scores between the two groups.

8.2.1.5 19.7% of the higher proportion of JMI students fell within the top range of the post-test score distribution

Finally, the proportion of high and low performing students was explored based on the top (the values above 14.6 out of 17) and bottom (the values below 5.40 out of 17) end of the post-test score distribution. The descriptive statistics revealed that a 19.7% of the higher proportion of JMI students were represented within this top range of the distribution compared to their EMI counterparts. This result indicated that there were some descriptive differences in the proportion of successful students according to the medium of instruction.

To summarise, although the results of an ANCOVA indicated no overall group mean differences between EMI and JMI students, observation of their individual post-test scores showed that a higher proportion of JMI students were observed in the top range of this post-test score distribution. Thus, this result highlights a need for qualitative examinations to tease out the perceptions and learning experiences of individual students.

8.2.2 How do the EMI and JMI students perceive the medium of instruction to affect their challenges and achievement of content learning?

The second sub-research question examined how the mediums of instruction affected students' perceptions of their academic achievement based on qualitative analysis of interviews. Below, findings will be discussed in relation to four particular challenges facing EMI and JMI students when they study Chemistry.

Perceived challenges in EMI

8.2.2.1 Insufficient comprehension of the course textbook

The study highlighted several factors behind students' low comprehension of the course textbook. Firstly, EMI students struggled to comprehend both technical and general academic terms used in their textbook (e.g. 'involve', 'polarised', 'charged' and 'law') whereas JMI students only lacked discipline specific terms (e.g. 'Molecular Orbital Theory'). Secondly, although EMI students struggled to read the textbook accurately and quickly due to their poor understanding of English grammar structures, many of the JMI students could read it at higher speeds without sacrificing their

comprehension. Thirdly, EMI students who had previously learned Chemistry in Japanese prior to university all placed great reliance on translation. Finally, EMI students whose learning occurred through L2 English could not achieve the same level of comprehension as that of their L1, noting that EMI learning was cumbersome and inefficient. Due to these challenges, they sacrificed clarity, retained less information, were forced to simplify concepts, and lacked in nuanced understanding. However, JMI students attributed their challenges primarily to their limited ability to understand the subject matter, suggesting that challenges were caused by different factors (language or content) depending on the medium of instruction.

8.2.2.2 Insufficient comprehension of the course lectures

The study indicated a range of factors that impeded students' comprehension of the course lectures. Firstly, EMI students were unaccustomed to listening to long stretches of English mediated-lectures, whereas JMI students could largely maintain a high level of concentration. Secondly, EMI students described that humour and jokes were less accessible than disciplinary concepts because they generally emerged more spontaneously and were not closely linked to the content of the lecture. Conversely, none of the JMI students reported this challenge. Thirdly, many EMI students struggled to listen to, understand and process academic concepts in English. This extra linguistic hurdle coupled with a demanding task of learning a new academic discipline impeded their lecture comprehension. However, their JMI counterparts did not experience any of these problems. Finally, EMI students had to relearn the English versions of Chemical elements in the periodic table which they had already studied in Japanese at school, such as 'hydrogen', 'helium' and 'nitrogen'. They also lacked an understanding of vocabulary containing both scientific and everyday dual meanings used in the technical domain, such as 'shell', 'element', 'activity', and 'stability'. JMI students also struggled to learn a large amount of technical terms but they already knew many of the words from their school Chemistry lessons.

8.2.2.3 Low performance on the final course exam and Final Grade

The study found several factors that contributed to students' low performance on the final course exam and Final Grade. Firstly, EMI students had to spend long hours preparing for the EMI course, as it was more linguistically demanding than other courses taught in Japanese. Their heavy workload prevented them from managing

time effectively, which in turn negatively affected their grades for both the Chemistry and other courses. In contrast, JMI students did not necessarily have to prioritise the Chemistry course but perceived that other courses were more difficult due to factors, such as course level, professors, assignments and their own prior content knowledge. Secondly, EMI students could not understand exam questions and instructions accurately and quickly. Their linguistic challenges (e.g. slow reading speed) had persisted throughout the semester, leading to reduction in the understanding of disciplinary concepts and a slowing down of content coverage. JMI students also expressed some similar problems (e.g. poor time management, low comprehension of concepts), but their performance was affected by non-linguistic factors (e.g. their own ability to understand subject matters). Finally, EMI students struggled to solve problem sets and formulate comments in English due to their low lecture comprehension and insufficient L2 English writing skills. Conversely, JMI students were enthusiastic about the Comment Worksheet without highlighting any issues.

8.2.2.4 Difficulty in adjusting to the new academic and cultural conventions

The study indicated two particular factors behind students' school to university transitional challenges. Firstly, EMI students struggled to adapt to more student-centred learning culture, which was specific to EMI. They faced the additional difficulty of having to be more critical and analytical. JMI students also had to conform to the new academic culture (e.g. long lecture duration, independent learning, fast-paced lectures, and advanced content level). Yet, they regarded their teaching method to be teacher-centred, which mirrored how they were previously taught at high school. Secondly, EMI students' challenges were long-lasting rather than short-term teething problems. Many of their problems expressed at the start of the study persisted throughout the semester (e.g. slow reading speed, insufficient lecture comprehension, and reduction in understanding). In contrast, JMI students successfully adjusted to the new academic culture, specific requirements of assessment and evaluation, and the disciplinary norms and practices. Many of their academic challenges fully dissipated by the end of the semester.

Thus, the study found some qualitative differences in students' perceived difficulty in being successful in the learning of Chemistry depending on the medium of

instruction. That is, EMI students appeared to have endured a considerably more taxing induction into content learning than their JMI counterparts.

Perceived successes in EMI

In answering the second sub-research question, the study also explored students' perceptions of positive impacts of content learning through EMI, going beyond the mere educational outcomes of achieving successful test scores. Students expressed the following four benefits of EMI education.

8.2.2.5 Improved English proficiency

Although the primary aim of learning Chemistry was to achieve a sufficient mastery of content knowledge, many EMI students regarded the development of L2 skills as an expected outcome of EMI. While this was a particular benefit exclusive to EMI students, one JMI learner reported an improvement in technical vocabulary size in English. Some of his reading assignments were in English and the lecture slides also introduced English translations alongside the original Japanese terms.

8.2.2.6 Improved motivation

Although EMI introduced an additional linguistic hurdle, many students recognised the English medium nature of the course as a driving force behind their strong motivation and determination because they were intrinsically motivated to learn through English. Similarly, JMI students were also highly motivated to improve their content knowledge as many of them wanted to pursue Chemistry at university.

8.2.2.7 Access to a wider range of learning resources through EMI

EMI students were able to access information bilingually using both English and their L1. Due to the wider range of online learning resources available in English (e.g. YouTube channels) than their L1s alone, EMI participants attributed their success to the use of the English language. In contrast, JMI students reported the excessive reliance on Japanese based materials.

8.2.2.8 Long-term career advancement

Many students believed that EMI allowed them to achieve internationally oriented career goals. Lisa in particular valued the opportunities to interact with overseas

students in English. Conversely, JMI students were motivated to solely improve their disciplinary knowledge because their main long-term aim was to become a leading scientist in Japan rather than overseas.

In summary, while the quantitative findings revealed that EMI had at least no detrimental impact on students' content learning, the qualitative evidence revealed that EMI students encountered the greater level of challenges than their JMI counterparts. However, EMI students also derived much benefit from EMI education. Thus, the study highlighted that EMI students may have been able to compensate the costs to content learning with these subsidiary benefits which were linked to EMI.

8.3 Research question two

The second research question for this study was: "In what ways do the two mediums of instruction differ? What context-related differences do the students ascribe their challenges and achievement of content learning?" In answering this question, the study drew on student and teacher interviews triangulated with observations to explore an array of contextual differences between the two courses which students ascribed to their achievement of content learning. In answer to this question, findings are organised in the following bullet point format for ease and clarity of discussion.

8.3.1 Strong presence of L1 medium resources among EMI students

EMI students took advantage of both English and L1 medium learning resources to maximise all of their available linguistic resources, such as library books, journal articles, online materials and textbooks from high school. In contrast, JMI students restricted themselves to Japanese medium learning resources. Thus, this learning condition where learners used English and L1 was one of the crucial factors which may have contributed to EMI students' successful content learning.

8.3.2 Peer learning as a useful pedagogical tool to ensure content mastery in EMI

EMI students derived much benefit from peer learning, such as clarifying academic concepts, establishing close friendships, and creating long term support networks. Benefits of peer learning appear to have significantly improved the speed and ease of the transition into EMI education. A mixed ability learning environment where learners with different levels of disciplinary understanding also provided them with

opportunities to learn to communicate and breakdown complex concepts. In contrast, none of the JMI students reported the presence of peer learning, suggesting that peer learning was vital to academic success in EMI.

8.3.3 Support provision was put in place differently in EMI and JMI

Although a range of support provision was put into place, there were also contextual differences in how the support provision was offered. Firstly, the teaching assistant provision was made only available to JMI students. The teaching assistant had limited English proficiency and was entirely absent in EMI lectures. Secondly, the Comment Worksheet was introduced in both the courses. However, the frequency of its use was not consistent across the two courses. Finally, the benefits of the English language program was only reported by EMI students. Thus, the two courses unveiled many contextual differences in how the support provision was integrated, and these contextual factors may have influenced the progress of their content learning.

8.3.4 Interactive and participatory educational culture in EMI

EMI students ascribed the academic success to the interactive educational culture of the EMI course. This educational culture provided them with a range of benefits to aid learning of the academic concepts. Students were able to interact with their content teacher, clarify issues with their course members, gain more confidence in seeking help, and organise their thoughts by explaining concepts using their own words. In contrast, JMI learners were introduced to more teacher-centred learning culture. They had little or no chance to engage with their content teacher and course members. Overall, these differences in educational culture between the two courses could be another crucial factor which contributed to students' academic success in EMI.

8.4 Research question three

The third research question for this study was: "Are there relationships between the EMI students' English language proficiency, academic English challenges, prior content knowledge, and content knowledge achievement?" This research question consisted of two sub-research questions. The first section will draw on descriptive and inferential statistics of the relationship between key quantitative variables, including English proficiency, challenges, prior content knowledge, and the achievement of EMI content knowledge in section 8.4.1. Subsequently, the second part will involve

student interviews to shed light on a more nuanced understanding of how these factors may interact and influence students' perceptions of difficulties and ease towards academic study in EMI in section 8.4.2.

8.4.1 To what extent factors, such as the EMI students' English language proficiency and perceived EMI challenges are associated with content knowledge achievement?

This first sub-research question examined the extent to which students' achievement in content learning was associated with proficiency, prior knowledge and challenges.

8.4.1.1 Descriptive analysis: Proficiency, pre-test, post-test scores and challenges

Firstly, descriptive statistics showed that EMI students' post-test scores varied according to proficiency. That is, those with higher proficiency, on average, achieved higher scores than their counterparts with lower proficiency. An examination of the observations below 15.8% and above 84.2% of the distribution as the threshold of lower and higher scores showed that the top and bottom end of the distribution was exclusive to those who reached IELTS 8.0 or above and IELTS 7.0 or below, respectively. Thus, there were differences in the proportion of successful students represented at the top and bottom end of this distribution based on proficiency.

Secondly, the EMI Challenges Scale scores also varied according to proficiency. This suggests that students at different levels of proficiency expressed differences in the perceptions of challenges. An exploration of the observations above 84.2% and below 15.8% of the distribution indicated that the top end of the distribution was exclusive to those who reached IELTS 8.0 and above whereas the bottom end of the distribution represented more of the lower proficiency students who reached IELTS 6.5 and 7.0. Thus, those with IELTS 8.0 or above were more likely to report greater ease towards EMI studies than their counterparts with IELTS 7.0 or below.

Thirdly, EMI students' post-test scores appeared to vary according to the levels of challenges. This indicates that the higher the score on the EMI Challenges Scale, the greater the students performed on the post-test; the lower the challenges score, the lower they scored on the post-test. However, a close-up examination of the top and

bottom range of the observations revealed no clear link between post-test scores and the level of challenges. Thus, the extent to which challenges were perceived to be more or less difficult may not necessarily be associated with post-test scores.

8.4.1.2 Correlations: Proficiency, pre-test, post-test scores and challenges

The correlation analyses suggested that students with higher proficiency were more likely to achieve greater content test scores and face fewer challenges. More successful students were more likely to be equipped with higher prior content knowledge and were possibly able to mitigate academic English challenges. However, possessing considerable prior content knowledge was not necessarily associated with the achievement of proficiency or greater ease for challenges in EMI.

8.4.1.3 Regression analysis: Proficiency, pre-test, post-test scores and challenges

Three linear regression revealed that proficiency and pre-test scores were both significant predictors of post-test scores while challenges score did not significantly predict post-test scores. This suggests that students might have expressed greater challenges, but they were still able to achieve successful content test performance. A multiple regression model included proficiency and pre-test scores, which explained 35.3% of the variance in post-test scores. As proficiency and pre-test scores in simple linear regression models only explained 20.6% and 17.2%, respectively, this multiple regression model accounted for a larger proportion of variance in post-test scores.

8.4.2 How do the EMI students perceive English language proficiency to affect their perceptions of challenges towards EMI study?

The second sub-research question explored how students perceived the magnitude of their challenges differently according to English proficiency. Below, findings will be discussed in relation to four particular challenges facing EMI students.

8.4.2.1 Insufficient comprehension of the course textbook

The study highlighted several factors that impeded students' comprehension of the course textbook. Firstly, EMI students from all proficiency levels experienced difficulty with insufficient vocabulary knowledge. In particular, those with IELTS 6.5 had issues with high frequency general academic English words, such as 'inhale',

‘hydrogen’ and ‘compute’. Conversely, students reaching IELTS 8.0 or above appeared to only face issues with technical words that contained unfamiliar academic concepts. Secondly, students with lower proficiency struggled with English grammar structures and expressed a slower reading speed than that of their L1. Conversely, those with higher proficiency could understand the textbook accurately without slowing down the reading speed. Nonetheless, the level of prior content knowledge also appeared to greatly influence their reading speed. Thirdly, those with IELTS 6.5 in particular had to constantly translate unfamiliar English words into Japanese. L2 proficiency influenced the reliance of translation but educational background and whether learners were L1 or L2 English speakers also appeared to affect whether they needed to rely on translation. Finally, participants from all English proficiency levels perceived that learning through their L1 would be more efficient than learning through their L2 English. Even those with higher proficiency still perceived their comprehension to be somehow inferior to that of their L1.

8.4.2.2 Insufficient comprehension of the course lectures

Results indicated various challenges regarding students’ low comprehension of the course lectures. Firstly, EMI students from all proficiency levels appeared to experience concentration fatigue when attending university lectures as this was a general academic challenge reported by all students. Nonetheless, those with lower proficiency in particular were not accustomed to listening to long stretches of EMI lectures. Their counterparts with higher proficiency, however, could be more selective about key information. Secondly, students below IELTS 8.0 struggled to understand humour and jokes during EMI lectures due to their low listening skills and absence of prior experience of attending lectures in English. In contrast, those with IELTS 8.0 or above were capable of understanding humour and jokes more effortlessly. Thirdly, students with lower proficiency struggled to deal with both language and content issues when listening to, comprehending and processing lectures in English. Unsurprisingly, having a higher level of prior content knowledge also helped them boost their lecture comprehension. Finally, learning new technical words was a general academic challenge reported by all students regardless of proficiency. Even those with high proficiency still faced this problem. However, students from Japanese medium school backgrounds in particular had to confront an extra hurdle of relearning English versions of many of the basic Chemical terms previously taught in Japanese.

8.4.2.3 Low performance on the final course exam and Final Grade

The study indicated several factors behind students' low performance on the final course exam and Final Grade. Firstly, EMI students with lower proficiency noted that their ability to manage workload undermined their final grades. Those with IELTS 7.5 or below had to spend particularly long hours preparing for the Chemistry course due to the English medium nature of the course, which in turn negatively affected their final grades for both the Chemistry and other enrolled courses. Conversely, students with higher proficiency could spare more time studying for other courses, which were more demanding for other reasons, such as course level, assignments, and their own disciplinary background. Secondly, low performance of the final course exam was a general academic challenge reported by all students regardless of proficiency. However, there were also specific challenges reported exclusively by those with low proficiency (e.g. misunderstanding concepts and exam instructions). Finally, while all students valued the Comment Worksheet, those with lower proficiency expressed particular difficulty in formulating questions and solving problems in English. In contrast, their counterparts with higher proficiency did not report any of these challenges. Meanwhile, previous exposure to EMI education, motivation, lecture comprehension, revision, and preparation were shown to be also important factors influencing their performance on the Comment Worksheet and final grades.

8.4.2.4 Difficulty in adjusting to the new academic and cultural conventions

The study indicated two particular factors behind students' school to university transitional challenges. Firstly, students from all proficiency levels appeared to grapple with general school-university transition, such as conforming to the new academic culture and conventions of a new institution and discipline. However, students with low proficiency in particular had to contend with the acculturation into the new academic culture specific to EMI, which was more interactive and participatory. Besides proficiency, the school educational culture and teaching medium at school was also a factor that impacted on their ability to adjust to EMI education. Secondly, students with higher proficiency could fully overcome many of their challenges reported at the start of the study (e.g. longer lecture duration, advanced technical vocabulary). However, those with lower proficiency expressed that many of their adjustment issues persisted through the term (e.g. small vocabulary

size, slow reading speed, and grammar issues). This suggests that achieving a certain level of proficiency was a prerequisite for being successful in EMI.

In conclusion, interview data indicated that students with lower proficiency expressed considerably more onerous challenges than their counterparts with higher proficiency. That is, many of the challenges were more likely to be apparent among less proficient English learners in EMI. While proficiency was key to success, other factors also emerged as crucial facilitators for success, such as previous exposure to EMI education, motivation, prior content knowledge, revision and preparation, learning strategy, and relationship with course members and lecturers.

8.5 Limitations of the study

8.5.1 The sample size

While the descriptive statistics of the current study revealed some differences in EMI students' post-test scores and challenges according to English language proficiency, the sample size of the EMI group (n=28) did not allow the use of t-tests or one-way ANOVA. Without multivariate analyses, it was not possible to examine whether these group mean differences occurred between any specific groups or whether students with different IELTS band scores (e.g. IELTS 6.5 and 7.0) were statistically different from each other in terms of their post-test scores or language-related challenges. The current study could only suggest that EMI students with lower proficiency had greater challenges and lower post-test scores than those with higher proficiency. Thus, future quantitative research is needed with a large sample size to statistically examine to what extent language-related challenges expressed by EMI learners may differ according to English level (Kline, 2016; Hair, Hult, Ringle & Sarstedt, 2017).

8.5.2 Generalisability of the study

This study embraced the "messiness" of EMI implementation as part of its ecological validity—that is a myriad of contextual factors (e.g. university ranking, disciplines, L1 use, teacher profiles, class size, pedagogy, course level) were considered part of the real-world representation of EMI programs, thus the findings are interpreted within the current research context. While it was not the intent to generalise the findings of this study to a larger student population across wider EMI contexts, some

of the following factors were particularly unique to the current research context and must be acknowledged as they may have specifically limited the generalisability of the study.

8.5.2.1 English proficiency

Firstly, the advanced level of L2 proficiency among the EMI students in this study may have greatly contributed to successful content learning in EMI. As data were collected from one of the top ranked Japanese universities participating the TGUP, the EMI participants (n=28) achieved the median IELTS band score of 7.0 ranging from 6.5 to 8.0, which stood in contrast to CEFR level of B2 (IELTS band score ranging from 5.5 to 6.5) described by Brown and Bradford (2018) as a commonly accepted threshold required for success in EMI in Japan. This high level of English proficiency among the EMI students may have been one of the crucial factors that contributed to the conditions that underpinned success in EMI content learning and particularly limited the generalisability of the study.

8.5.2.2 Course level

Secondly, another key contextual factor was that the study was conducted in an introductory course. This suggests that the content test adopted was primarily made up of single-word questions, multiple choice questions, and problem sets asking students to draw chemical structures and solve chemical equations. These questions were linguistically less demanding than other types of questions, such as essay-based exams. Similarly, the introductory-level course may have also limited the types of challenges reported by students, where certain types of assessments were absent from the course curriculum, such as laboratory reports, essays, presentations and seminars. In exploring EMI students' language-related challenges, Evans and Morrison (2011) found that academic writing was perceived to be the most challenging aspect of university study (e.g. writing in an appropriate academic style, writing extended academic texts), whereas the current study found no occurrences of such challenges associated with academic writing. As such, the introductory nature of the current Chemistry program might not have allowed the researcher to capture the full extent of how the medium of instruction could affect students' academic performance.

8.5.2.3 Self-enrollment

Thirdly, self-enrollment of students' preferred language of instruction may have introduced a range of differences in demographic factors between the EMI and JMI groups. As a result, a larger proportion of students with higher English proficiency and previous EMI learning experience who were also favourably versed in EMI education were present in the EMI group. This study successfully controlled for students' pre-test scores between the two groups to ensure that the participants had a similar baseline of prior content knowledge at the start of the study. However, future outcome research could statistically control for a larger number of key learner variables and recruit an EMI group and its L1 medium instruction control group who share as similar learner factors as possible.

8.5.3 Duration of data collection

In unpacking the concept of "cost to content learning", Macaro (2018) observes that potential reduction and slowing down of content learning in EMI could be acceptable as long as it is "a short-term, documented and carefully controlled negative effect" (p. 157). Accordingly, it is important that future research investigates whether students' challenges in EMI are long-lasting problems or short-term teething problems. While this study revealed that students were not able to fully alleviate their challenges over the course of one academic semester, it is possible that many of their challenges could dissipate by the end of their four-year degree program, as they progress in their content learning and English proficiency. Dörnyei (2007) recommends the use of longitudinal research as an approach to address these issues, defining it as a procedure where data collection is completed several times over a period of time. Thus, more longitudinal research is needed in future to better understand whether students' challenges persist and dissipate during their university careers and identify specific areas of learning that needs any additional support.

8.6 Recommendations for future research

In light of limitations of the current study, the previous section has highlighted various emerging issues that can be addressed in future research. This section offers some tentative suggestions for possible areas of future research that deepens our understanding of implications of EMI in higher education. I recommend future research in relation to four broad research themes: 1) the interplay between students'

perceptions of challenges and success in EMI; 2) EMI content learning differences between academic disciplines; 3) factors contributing to success in EMI; and 4) broader educational outcomes of EMI that exceed the content and language outcomes.

Firstly, I propose that future research should explore success and benefits of EMI in addition to its challenges and difficulties. As EMI programs expand rapidly, scholarly attention to date has been placed mainly on its challenges and difficulties than its subsidiary successes and benefits. Much of the previous research incorporating qualitative data in particular has found negative evidence that EMI content learning can be ineffective, compared to its L1 medium instruction counterpart (e.g. Bradford, 2012; Cho, 2012; Pun & Jin, 2021). For example, Chan (2014) claims that the constant use of dictionary and translation is a particular area of EMI content learning that is time-consuming for many low proficient students. On the contrary, the current study suggested that participants viewed the additional time required for revision and preparation for EMI studies as a double-edged sword. On the one hand, EMI was time-consuming and laborious; on the other hand, EMI achieved more extensive and in-depth content learning. For example, learners had to rely on additional learning resources, such as dictionaries, online resources, and supplementary learning materials in their L1, which expanded their content knowledge further. They also developed new learning strategies, such as making an separate notebook to write down unfamiliar English words and concepts. Many of them emphasised a need to make an additional effort to make their learning more efficient to be successful in EMI. Thus, seeing that a switch in the medium of instruction may have both positive and negative effects on students' content learning, I propose that future research should consider the interplay between students' perceptions of both challenges/difficulties and success/benefits when they learn through EMI.

Secondly, this study highlighted that many of the challenges reported by students were distinctive to the learning of Chemistry, such as difficulty in calculating the formula of a compound. Seeing the discipline-specific nature of EMI, some have examined whether the effects of EMI vary according to subject areas. For example, Dafouz et al. (2014) revealed no differences in EMI students' academic performance between Accounting (Social Sciences) and History (Humanities). Conversely, Coxhead and Boutorwick (2018) found that Humanities requires a larger amount of

low frequency vocabulary knowledge than STEM, suggesting that EMI students may experience different levels of challenges according to academic subjects. As this study was predicated entirely on a single case of an undergraduate introductory Chemistry course, future projects could examine how EMI affects students' content learning differently according to academic subjects.

Thirdly, this study adopted in-depth qualitative methods of student and teacher interviews and lecture observations to shed light on an array of positive factors that contributed to the speed and ease of successful content learning in EMI (e.g. class size, educational culture, peer learning, teacher-student interaction, and language support). In particular, many EMI students proactively took advantage of both Japanese and English medium resources to maximise their linguistic resources. The triangulation of data sources thus achieved nuanced contextual data to improve the transparency of the inevitably messy research context in which a range of contextual differences existed between the EMI and JMI courses. Future research should therefore adopt qualitative approaches to ensure they have a contextually rich and detailed picture of EMI students' learning experience, within which to interpret findings. This would help us better understand what factors drive greater levels of successful content learning in EMI, and how we can ensure that key positive factors are present to create context sensitive support systems that can facilitate students' content learning in EMI.

Finally, in regard to a number of broader benefits of EMI education found in this study (e.g. access to a wider range of learning resources, improved linguistic and intercultural competence), I suggest that future research projects could investigate the broader educational benefits of EMI that go beyond the mere language and content outcomes. Recently, some evidence has emerged in terms of the development of softer skills through EMI, including how EMI facilitates students' long-term career outcomes (e.g. Huang & Curle, 2021; Altay & Erçin, 2020). However, there is not yet sufficient empirical evidence to establish causal links between the introduction of EMI and its expected outcomes. Thus, future research may attribute success of EMI to the development of learners' more soft skills. This would help us better understand whether EMI achieves its full potential to deliver subsidiary benefits that exceed the language and content outcomes.

In summary, this section has suggested gaps in research literature which can be further addressed in future research in light of the limitations of the current study. Table 43 below offers an overview of recommendations for possible areas of future research on outcomes of EMI, which are presented alongside suggestions of data set and analysis procedures used to answer them.

Table 44: Summary of recommendations for future research

Future research themes	Specifics	Instruments & Data
1. The interplay between students' perceptions of challenges and success in EMI	- Investigate challenges and success between academic disciplines - Explore students' perceptions according to proficiency - Compare findings between higher and lower ranked universities	- Classroom observation - Student interviews
2. EMI content learning differences between academic disciplines	- Include delayed measurements to examine whether there are differences in long-term content acquisition and retention - Extend sample sizes of students	- Classroom observation - Student and teacher interviews - Content test outcomes - Challenges questionnaires
3. Factors contributing to content learning success in EMI	- Compare findings between higher and lower ranked universities - Investigate key factors between academic disciplines	- Classroom observation - Student and teacher interviews - Background questionnaires - Motivation questionnaires
4. Broader educational outcomes of EMI that exceed the content and language outcomes	- Investigate career outcomes based on the medium of instruction - Examine how EMI impacts university ranking, strategic policy, and research output	- Student and teacher interviews - Questionnaires for alumni and employers - Policy documentation

8.7 Recommendations for curriculum planning in EMI

The positive evidence found in this study was largely attributed to a context where learners had access to a wide variety of institutional support (e.g. library resources, L1 medium materials, the teaching assistant, and the English language program), and where they could establish supportive and collaborative peer networks to combat their challenges. In order to create a learning condition where these positive factors are present, this study yields a number of pedagogical implications.

Firstly, EMI students with high English proficiency may also benefit from undertaking preparatory language courses. The current study found that EMI students regardless of their English language proficiency faced an array of general academic challenges when learning a new academic subject at university (e.g. conforming to the academic culture, understanding academic concepts, and managing workload). Previous studies (e.g. Charles & Pecorari, 2016; Arnó-Macià, Aguilar-Pérez & Tatzl, 2020; Thompson et al., 2019) have suggested that preparatory language courses, such as EAP and ESP courses can facilitate students' ability to develop general academic skills required for discipline specific activities at university, such as critical thinking, academic writing, communication, presentation, bibliographic and referencing skills. Therefore, students with higher levels of English proficiency may also benefit from undertaking such preparatory language courses where they can develop boarder academic skills required for university study.

Secondly, preparatory language courses may be offered in the form of ESP or ESAP courses rather than EAP courses to provide students with language support that is geared towards specific academic disciplines. The current study found that learners faced a number of discipline specific challenges that were highly relevant to the learning of Chemistry (e.g. acquiring an abundance of technical vocabulary and academic concepts, and solving chemical equations). Accordingly, EMI students may benefit from receiving subject-specific language support that is directly pertinent to their specific academic subject needs, rather than general academic English language programs.

Thirdly, EMI teachers may wish to play more prominent role in supporting students' language problems in EMI programs. The current study found that the content teacher in the EMI course felt only responsible for the delivery of content knowledge and hardly paid attention to students' English language issues, a finding mirrored in previous studies (Costa, 2012; Airey, 2012; Wilkinson, 2015). However, the current study also discovered that students in particular transitioning from a Japanese-medium high school were unaccustomed to listening to long stretches of English medium lectures, resulting in a range of language-related challenges (e.g. concentration fatigue, misspelling of both general academic and technical English vocabulary, and insufficient understanding of humour and jokes). In light of these challenges, EMI

teachers may wish to write down important class announcements on a whiteboard, avoid using idiomatic and colloquial expressions, and offer a list of key terminologies to better support EMI students.

Finally, universities may wish to arrange different support schemes for international and local students. While the presence of overseas students is vital to the development of EMI provision spurred by the recent higher education initiatives (Radjai & Hammond, 2021), overseas students in the current study were not necessarily able to access support schemes in the same way as their local student counterparts, in particular when support was only offered in Japanese (e.g. library resources and teaching assistants). To this end, universities may wish to organise support schemes specifically catered for overseas students (e.g. local language support).

8.8 Conclusion

In conclusion, the current research examined the extent to which the medium of instruction affects undergraduate students' learning of Chemistry in Japan. In doing so, I undertook a direct comparison between students studying introductory Chemistry through EMI and those studying the same curriculum through JMI. An analysis was undertaken of students' learning outcomes measured by pre-post course content tests. These measurable test differences were triangulated with student and teacher interviews, in addition to lecture observations, in order to identify further differences in their perceived content learning experience. While the quantitative results did not detect a statistically significant difference in test outcomes between the groups, the qualitative evidence suggested that EMI students experienced a considerably more taxing induction into their learning of Chemistry. Given that research to date has primarily focused on either positive or negative outcomes of EMI, this study adopted a two-pronged, quantitative and qualitative approach to unpack the complicated nature of students' content learning to build on knowledge in the field of EMI research. Methodologically, this study highlighted the importance of teasing out the experiences and perceptions of individual students and their learning experiences using qualitative methods, rather than simply computing and contrasting group mean test scores through quantitative methods.

This study also responded to the still largely unheeded call for quasi-experimental

research into academic outcomes of EMI. In doing so, the study collected data in real learning environments with real teachers and learners, rather than sanitised research conditions where contextual, institutional and learner factors are sidelined, rather than brought to the forefront to aid data interpretation. To identify an array of uncontrolled factors that influenced content learning, quantitative measures of test scores were substantiated with in-depth qualitative methods, including student and teacher interviews, and lecture observations. To capture the changing nature of students' learning experiences, the study drew on a pre–post research design. As a result, the study offered real-world implications that had direct relevance to students' educational experiences under investigation. It recommended that triangulation of quantitative, qualitative and longitudinal research is central to future EMI research.

Finally, the study unveiled a number of differences in terms of students' perceived benefits of studying Chemistry through EMI and JMI, highlighting that many of the benefits were distinctive to EMI education and exceeded the language and content outcomes (e.g. improved English language proficiency, long-term career advancement). To better assess the broader educational benefits of EMI, this study has aided in theory building of EMI outcome research that benefits of EMI may not necessarily be investigated in an association with its L1 medium instruction counterpart but should be treated in its own right. Thus, this study has contributed to theoretical and methodological knowledge in this unexplored realm of EMI outcome research.

Appendices

Appendix 1: The number of classes taught in foreign languages or in English

	2013	2016	2019	2023
Number of classes taught foreign language (A)	248	353	449	568
Number of classes taught in English (B)	240	345	441	560
Number of classes taught at the university (C)	1,451	1,400	1,400	1,400
Proportion (B/C)	16.5%	24.6%	31.5%	40.0%

Source: 'Top Global University Project' (2016)

Appendix 2: Interview participants

Name	Language of instruction	Gender	IELTS	First Language	Pre-test	Post-test	Final Grade
1 Hikaru*	EMI	Male	6.5	Japanese	2	5	C
2 Rina*	EMI	Female	6.5	Japanese	0	7.5	C
3 Marii*	EMI	Female	7.0	Japanese	0	1	C
4 Ayaka*	EMI	Female	7.0	Japanese	0	8.5	C
5 Minori*	EMI	Female	7.0	Japanese	0	9.5	B
6 Lisa	EMI	Female	8.0	Japanese	3	12.5	A
7 Miyazaki*	EMI	Female	8.0	Japanese	13.5	15	A
8 Charlotta	EMI	Female	8.0	German	0	9.5	C
9 Tikamon	EMI	Female	7.5	Thai	5.5	12.5	A
10 Kiley	EMI	Female	7.0	Swazi	1	4	E
11 Maxwell	EMI	Male	-	English	0	8	D
12 Saika	EMI	Female	-	Bilingual	0	7.5	C
13 Emma	EMI	Female	-	Bilingual	0	12.5	A
14 Oda*	JMI	Male	5.0	Japanese	2	15	B
15 Kyoka*	JMI	Female	5.5	Japanese	5	16	A
16 Nao	JMI	Female	7.5	Japanese	0	3	C
17 Shizuka	JMI	Female	6.0	Japanese	2	6	C

Notes: Bilingual students self-reported to obtain both Japanese and English as their first languages; Grading scale for the pre-post-test ranges from 0 to 17, and Final Grade ranges from A to E (A, B, C, D, E) where A represents the highest possible grade, and E the lowest possible pass grade (fail); Student with an asterisk (*) participated in both the first and second round interviews.

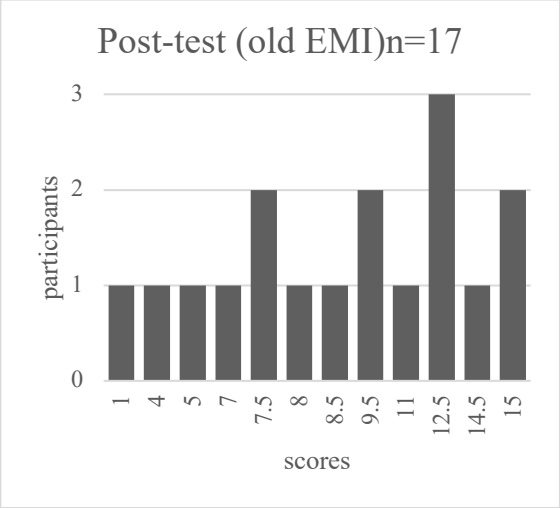
Appendix 3: Combining the two sets of the EMI groups

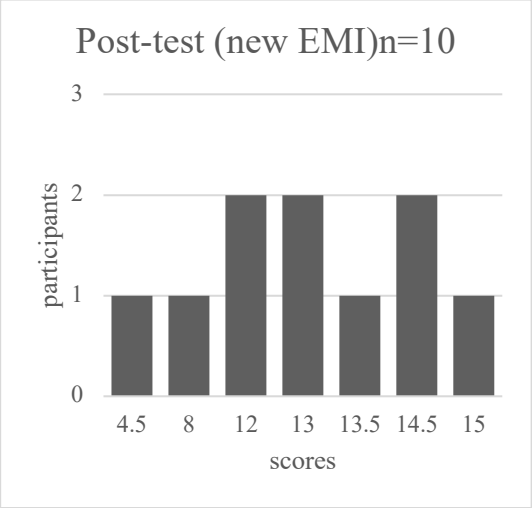
The study combined the two separate datasets of the EMI participants in the 2018-19 (old) and 2019-20 (new) sets. The results below show no significant difference in the variables (i.e. English proficiency, pre-post-test scores and gain scores) between the sets. Due to the scale type and normality of each data, an independent samples t-test was used for post-test scores and gain scores which are ratio variables and normally distributed, while a Mann Whitney U test was used for L2 proficiency which was an ordinal variable and pre-test scores which were ratio but not normally distributed.

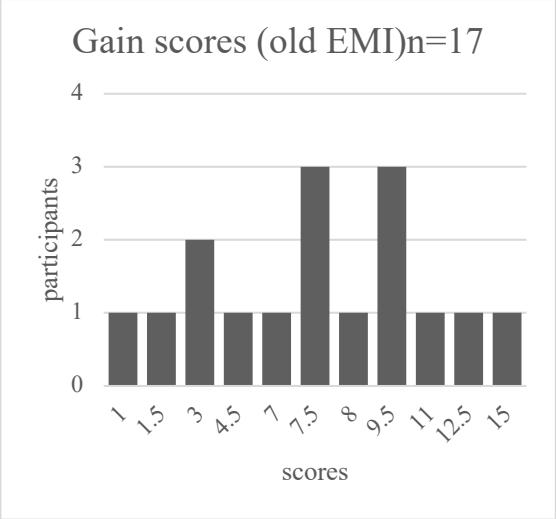
Post-test scores and gain scores

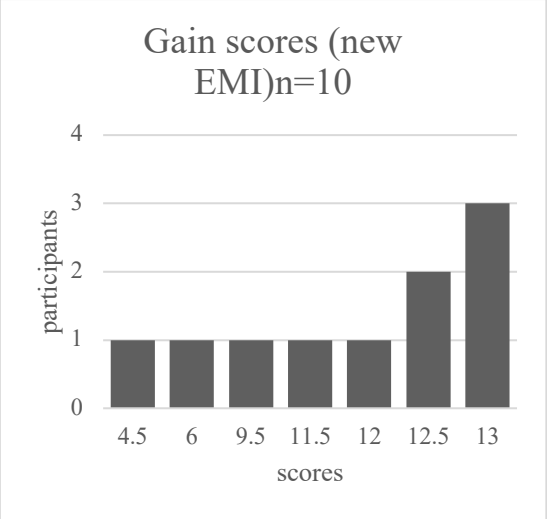
The assumptions of equal variances via Leven's test were met F (.951)=0.339; (.212)=0.649. An two-tailed independent samples t-test suggest that the mean differences in post-test and gain scores are not significant between the old and new EMI groups ($t_{[25]} = -1.709$; $p = .10$) and ($t_{[25]} = -2.403$; $p = .086$), respectively. The magnitude of the differences in the means for post-test (mean difference = -2.6, 95% *CI*: -.564 to .524) was small (eta squared =.015); and for gain scores (mean difference = -3.42, 95% *CI*: -.634 to -.489) was small (eta squared =.011).

	Group	N	Min	Max	Mean	Median	SD	*Z skewness*	*Z kurtosis*
Post-test scores	old	17	1	15	9.44	9.500	3.99	-0.65	-0.32
	new	10	4.5	15	12	13	3.30	-1.38	1.70
Gain scores	old	17	1	15	7.32	7.500	3.82	0.09	-0.24
	new	10	4.5	13	10.75	12.75	3.12	-1.07	0.52





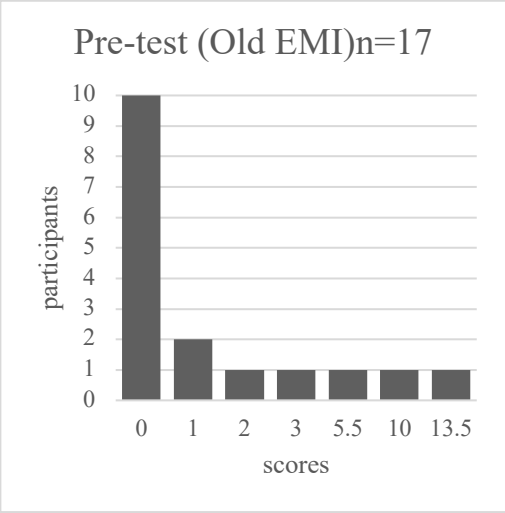


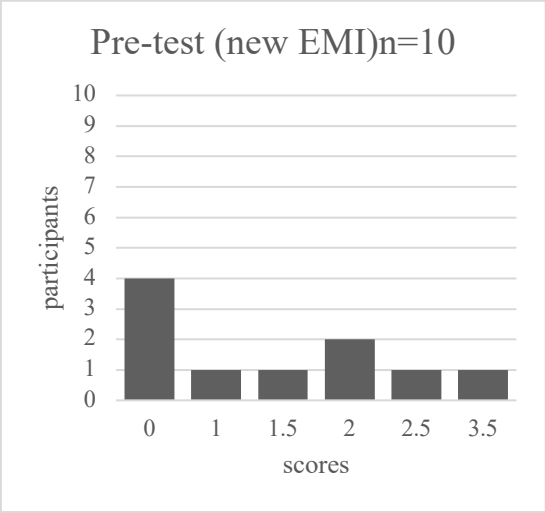


L2 proficiency and pre-test scores

A Mann-Whitney U Test examined the potential differences in the students' pre-test and IELTS scores between the old and new EMI groups, showing no significant difference in the pre-test scores of the old (Md=0, n=17) and new groups (Md=1.25, n=10), U=75, z=-.542, p=.588, r=.022; and in the IELTS scores of the old (Md=7.5, n=17) and new groups (Md=8, n=10), U=72.5, z=-.650, p=.515, r=.052.

	Group	N	Min	Max	Mean	Median	SD	*Z skewness*	*Z kurtosis*
Pre-test scores	old	17	0	13.5	2.12	0	3.96	3.94	3.85
	new	10	0	3.5	1.25	1.25	1.25	0.66	-0.70
IELTS scores	old	17	6.5	8.0	-	7.5	-	-	-
	new	10	7.0	8.5	-	8.0	-	-	-





Appendix 4: Background Questionnaire Form (English translated Version)

Thank you for taking the time to take part in this study which investigates the effects of English taught courses on students' learning in Japanese universities. The information you provide will be treated in the strict confidence.

①Year ・major : _____year_____major. （Please state A.S. if not decided）

②ELA Stream : _____.

③English language proficiency（IELTS, TOEFL, TOEIC）:

test_____score（L,R,WS）_____.

④Have you had any previous experiences to study subjects in a foreign language before high school? : Yes・No .

If so, where and what kind of school did you go? :_____.

⑤Which language did you use to learn academic subjects in high school? : _____.

⑥Do you have study abroad experience? : Yes・No .

If so, which country, which level and how long did you study? : _____.

⑦Are you taking English medium classes? : Yes・No .

If yes, which classes are you taking? : _____.

⑧How much time on average did you spend on class preparation for each class period? （Time 2 only）

: _____**hours**_____**minutes.**

⑨How much time on average did you spend on English learning outside English medium courses per day? Include time you spent on EAL course work. （Time 2 only）

: _____**hours**_____**minutes.**

⑩There were 14 classes this term? Please circle any class that you did NOT attend. （Time 2 only）

: 1 2 3 4 5 6 7 8 9 10 11 12 13 14

Email address : _____@_____.

※If you wish to take part in an interview (30 minutes) and/or to be informed of the findings of the research, please feel free to contact the researcher any time of the research, Ikuya Aizawa.

My Email Address is ikuya.aizawa@education.ox.ac.uk

Appendix 5: Challenges in your EMI classes (English version)

Part1 Academic Writing Skills

Please answer each item by circling the appropriate number on the scale from 1 to 5.		Very Difficult		-		Very easy
1	Planning written assignments	**1**	**2**	**3**	**4**	**5**
2	Expressing ideas in correct English	**1**	**2**	**3**	**4**	**5**
3	Revising written work	**1**	**2**	**3**	**4**	**5**
4	Using appropriate academic style	**1**	**2**	**3**	**4**	**5**
5	Writing a bibliography/ references section	**1**	**2**	**3**	**4**	**5**
6	Proofreading written work	**1**	**2**	**3**	**4**	**5**
7	Referring to sources in written work	**1**	**2**	**3**	**4**	**5**
8	Summarizing/ paraphrasing ideas in sources	**1**	**2**	**3**	**4**	**5**
9	Organizing ideas in coherent paragraphs	**1**	**2**	**3**	**4**	**5**
10	Expressing ideas clearly and logically	**1**	**2**	**3**	**4**	**5**
11	Linking ideas from different sources	**1**	**2**	**3**	**4**	**5**
12	Writing the introduction to an assignment	**1**	**2**	**3**	**4**	**5**
13	Writing the body of an assignment	**1**	**2**	**3**	**4**	**5**
14	Writing the conclusion to an assignment	**1**	**2**	**3**	**4**	**5**
15	Liking sentences smoothly	**1**	**2**	**3**	**4**	**5**

Part 2 Academic Reading Skills

Please answer each item by circling the appropriate number on the scale from 1 to 5.		Very Difficult		-		Very easy
1	Understanding specific vocabulary	**1**	**2**	**3**	**4**	**5**
2	Working out the meaning of difficult words	**1**	**2**	**3**	**4**	**5**
3	Reading carefully to understand a text	**1**	**2**	**3**	**4**	**5**
4	Reading quickly to find specific information	**1**	**2**	**3**	**4**	**5**
5	Identifying supporting ideas and examples	**1**	**2**	**3**	**4**	**5**
6	Reading quickly to get overall meaning	**1**	**2**	**3**	**4**	**5**
7	Identifying the key ideas of a text	**1**	**2**	**3**	**4**	**5**
8	Taking brief, relevant notes	**1**	**2**	**3**	**4**	**5**
9	Using your own words when taking notes	**1**	**2**	**3**	**4**	**5**
10	Understanding the organization of a text	**1**	**2**	**3**	**4**	**5**

Part 3 Academic Speaking Skills

Please answer each item by circling the appropriate number on the scale from 1 to 5.		Very Difficult		-		Very easy
1	Speaking accurately (grammar)	**1**	**2**	**3**	**4**	**5**
2	Speaking clearly (pronunciation)	**1**	**2**	**3**	**4**	**5**
3	Presenting information/ ideas	**1**	**2**	**3**	**4**	**5**
4	Participating actively in discussion	**1**	**2**	**3**	**4**	**5**
5	Communicating ideas fluently	**1**	**2**	**3**	**4**	**5**
6	Speaking from notes	**1**	**2**	**3**	**4**	**5**
7	Asking questions	**1**	**2**	**3**	**4**	**5**
8	Answering questions	**1**	**2**	**3**	**4**	**5**
9	Communicating ideas confidently	**1**	**2**	**3**	**4**	**5**
10	Using visual aids (e.g. PowerPoint)	**1**	**2**	**3**	**4**	**5**

Part 4 Academic Listening Skills

Please answer each item by circling the appropriate number on the scale from 1 to 5.		Very Difficult		-		Very easy
1	Understanding the main ideas of lectures	**1**	**2**	**3**	**4**	**5**
2	Understanding the overall organization of lectures	**1**	**2**	**3**	**4**	**5**
3	Understanding key vocabulary	**1**	**2**	**3**	**4**	**5**
4	Taking brief, clear notes	**1**	**2**	**3**	**4**	**5**
5	Identifying supporting ideas and examples	**1**	**2**	**3**	**4**	**5**
6	Understanding lecturers' accents	**1**	**2**	**3**	**4**	**5**
7	Following a discussion	**1**	**2**	**3**	**4**	**5**
8	Identifying different views and ideas	**1**	**2**	**3**	**4**	**5**
9	Understanding questions	**1**	**2**	**3**	**4**	**5**
10	Understanding classmates' accents	**1**	**2**	**3**	**4**	**5**

Adopted from 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.'

Appendix 6: EMI Student Interview Schedule (Japanese written after English)

1. **Effects of EMI on content learning**
   (英語で専門教科を学習する経験)

   - You mentioned in the questionnaire that your first language is _____. Do you think that you would understand the course differently if _____ were the medium of instruction?
   - （もし母語が使用されていたら、もっと簡単に授業を理解していたと思いますか。）
   - Does the language of instruction affect the level of your comprehension of the lecture and textbook?
   - （教授言語が異なることによって専門知識を学習するのに影響があると思いますか。）
   - Does the language of instruction affect your final grade and GPA?
   - （教授言語が英語になることは成績にも影響すると思いますか。）
   - Does the language of instruction affect the efficiency of your learning?
   - （英語で専門教科を学習する効率性についてどう思いますか。）
   - 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?
   - （母語で勉強した時の方が授業の準備と復習時間は短いですか。）

2. **Effects of EMI on English language learning**

   (英語で専門教科学習する経験)

   - Do you think that learning chemistry in English helps you improve your English level?
   - （英語力を英語開講の授業を通して向上させることができましたか。）
   - Which academic English skills do you improve through EMI? Reading, writing, listening, speaking?
   - （英語開講の授業を通して英語のどの側面がもっとも向上したと思いますか。）
   - Do you think that you improve your academic and technical vocabulary knowledge in English?
   - （語彙力(専門用語とアカデミック用語)が伸びたと思いますか。）
   - Do you think that EMI helps you improve your English skills as much as ESP courses?
   - （アカデミック英語の授業と英開講の授業はどちらが英語を伸ばすのに役に立ちましたか。）
   - You indicated your IELTS scores are_____. Is your English level enough to study the course in English?
   - （現在の IELTS 点数で学習をする際に十分だと思いますか。）

3. **EMI challenges**

   (英語で専門教科を勉強する際に直面する問題)

   - Do you face any challenges when learning chemistry in English?
   - （化学を学習する際に直面する問題は何ですか。）
   - What about challenges regarding the following academic English skills?
   - （４技能についてはどのような問題に直面していますか。）

1. writing 2. listening 3. speaking 4. reading 5. academic and technical vocabulary

   - Do you face any other general challenges when learning chemistry? Not English related-challenges?
   - （英語に関係なく化学の学習で困ることはありますか。）

- When you face challenges, what form of support do you use? Reading textbooks, using dictionaries, and asking teachers?
- （問題時にはどうしますか。辞書を使う、教科書を事前に読む、先生に聞く等行動しますか。）
- Do you use any materials in your first language to help you learn chemistry better?
- （授業外で日本語を使用して授業の内容を理解しようとしましたか。）

4. **Attitudes towards EMI**

   (EMI に対する姿勢)

- Can you share your general attitudes towards learning in English? Are you in favour of EMI programs?
- （EMI に賛成ですか。）
- What do you think would be the benefits of EMI programs?
- （EMI の何が一番の利益ですか。）
- What do you think would be the disadvantages of EMI programs?
- （EMI の何が欠点ですか。）
- What was your motivation to take the EMI chemistry course when there was also the Japanese option?
- （日本語ではなく英語開講を取ろうと思う理由は何ですか。）

5. **Post-course interview questions (only 'TIME 2')**

   (学期末インタビュー)

- Now you have finished the semester. How did you like the course through the semester?
- （一学期終了しましたが、コースの感想を教えてください。）
- Do you think you improved content knowledge though EMI as envisaged?
- （専門教科の学習は思うようにできましたか。）
- What do you think about your performance on the post-test, compared to the pre-test?
- （学期頭のテストと学期末テストで学習の伸びをはかることができましたか。）
- You said you had problems in the area of xx and xx (e.g. listening to teacher's accent), were you able to overcome these challenges towards the end of the semester?
- （初めのインタビューで問題になっていた点は克服することができましたか。）
- If you were to take the same course again, would you choose English again as the language of instruction?
- （同じコースをとるとしたらまた英語開講を受けますか。）

Appendix 7: JMI Student Interview Schedule (Japanese written after English)

1. **Effects of language of instruction on content learning**
   (英語で専門教科を学習する経験)

- You mentioned in the questionnaire that your first language is Japanese. Do you think that you would understand the course differently if English were the medium of instruction?
- （もし英語が使用されていたら、授業を理解するのにもっと困難を覚えていたと思いますか。）
- Does the language of instruction affect the level of your comprehension of the lecture and textbook?
- （教授言語が異なることによって専門知識を学習するのに影響があると思いますか。）
- Does the language of instruction affect your final grade and GPA?
- （教授言語が英語になることは成績にも影響すると思いますか。）
- Does the language of instruction affect the efficiency of your learning?
- （英語で専門教科を学習する効率性についてどう思いますか。）

- Do you think that the amount of time you spend on preparation and revision for the course would become longer if you were learning in English?
- （母語で勉強した時の方が授業の準備と復習時間は短いですか。）

2. JMI challenges
(英語で専門教科を勉強する際に直面する問題)

- Do you face any challenges when learning chemistry?
- （化学を学習する際に直面する問題は何ですか。）
- What about challenges regarding the following academic skills in your lectures?
- （授業中に４技能についてはどのような問題に直面していますか。）

1. writing
2. listening
3. speaking
4. reading
5. academic and technical vocabulary

- When you face challenges, what form of support do you use? Reading textbooks, using dictionaries, and asking teachers?
- （問題時にはどうしますか。辞書を使う、教科書を事前に読む、先生に聞く等行動しますか。）
- Do you use any materials in English to help you learn chemistry better?
- （授業外で英語を使用して授業の内容を理解しようとしましたか。）

3. Attitudes towards EMI
(EMI に対する姿勢)

- Can you share your general attitudes towards learning in English? Are you in favour of EMI programs?
- （EMI に賛成ですか。）
- What do you think would be the benefits of EMI programs?
- （EMI の何が一番の利益ですか。）
- What do you think would be the disadvantages of EMI programs?
- （EMI の何が欠点ですか。）
- You seem very (positive/negative) towards EMI. What was your motivation to take the chemistry course in Japanese when there was also the English option?
- （日本語ではなく英語開講を取ろうと思う理由は何ですか。）

4. Post-course interview questions (only 'TIME 2')
(学期末インタビュー)

- Now you have finished the semester. How did you like the course through the semester?
- （一学期終了しましたが、コースの感想を教えてください。）
- Do you think you improved content knowledge though JMI as envisaged?
- （専門教科の学習は思うようにできましたか。）
- What do you think about your performance on the post-test, compared to the pre-test?
- （学期頭のテストと学期末テストで学習の伸びをはかることができましたか。）
- You said you had problems in the area of xx and xx (e.g. understanding technical vocabulary), were you able to overcome these challenges towards the end of the semester?
- （初めのインタビューで問題になっていた点は克服することができましたか。）
- If you were to take the same course again, would you select Japanese again as the language of instruction?
- （同じコースをとるとしたらまた英語開講を受けますか。）

Appendix 8: EMI and JMI Teacher Interview Schedule (Japanese written after English)

1. Teaching practices

(授業に関して)

- What is the proportion of Japanese and English use in your course?
- （授業での日本語と英語の使用頻度を教えてください。）
- What is your typical teaching style? Can you go over how you conduct lectures, discussions, student-led presentations, pop quizzes, Comment Sheet, and teacher feedback?
- (どのような形式で授業を行いますか。講義、プレゼン、ディスカッション)
- Would you teach differently depending on the language of instruction?
- (EMI と JMI でに何か大きな異なりはありますか。)
- How often do you organise a faculty meeting with your team teaching partner?
- (どの頻度でチームティーチングの先生と摺り合わせをしますか。)
- Do you simplify your language to help the students understand you better?
- （学生にわかってもらうために言語を簡単にしますか。）
- Do you reduce content coverage because teaching is conducted in (Japanese or English)?
- （学生の理解を考えて本来よりも教える内容を減らすことはありますか。）

2. Effects of language of instruction on content learning
(専門教科を学習する経験)

- Does the language of instruction affect the level of your students' lecture comprehension?
- （教授言語が異なると専門知識を学習するのに影響があると思いますか。）
- Does the language of instruction affect your students' final grade and GPA?
- （教授言語が異なると成績にも影響すると思いますか。）
- Does the language of instruction affect the efficiency of your students' learning?
- （教授言語が異なると学習効率性にも影響すると思いますか。）
- If their first language were used, do you think your students would have understood your courses better?
- （もし母語が使用されていたら、学生は簡単に授業を理解していたと思いますか。）

3. Effects of EMI on English language learning (only used for the EMI teacher)
(英語で専門教科学習する経験)

- Do you think your students improve their English language proficiency through your EMI course?
- （英語力を英語開講の授業を通して学生が向上させることができましたか。）
- Which academic English skills do your students should improve? Reading, writing, listening, speaking?
- （英語開講の授業を通して英語のどの側面がもっとも向上したと思いますか。）
- Do you think your students have improved their academic and technical English vocabulary knowledge?
- （語彙力(専門用語とアカデミック用語)が伸びたと思いますか。）
- Do you think ESP programs help your students' English level?
- （アカデミック英語の授業は英語を伸ばすのに役に立ちましたか。）
- How much English do you think your students would need to effectively learn in EMI courses?
- （どのくらいの英語力が英語開講をとるためには必要だと思いますか。）

4. Student challenges
(専門教科を勉強する際に直面する問題)

- What do you think would be the challenges faced by your students in your course?
- （化学の学習する際に直面する最も顕著な問題は何ですか。）
- Do your students face any challenges regarding the following academic English skills:
- （４技能についてはどのような問題・困難に直面していますか。）(only used for the EMI teacher)

1. writing
2. listening
3. speaking
4. reading
5. academic and technical vocabulary

Appendix 9: Content test: post-test questionnaire for pilot study purpose only

テスト後のアンケート調査

Your Student number _____

学籍番号

I would like to ask you some quick questions about your experience of taking this test.

テストについて以下の質問があります。

1. The test was… *(please tick one box)*

テストについて（以下選択ください）

Very easy ❑ Quite easy ❑ Neither easy nor difficult ❑ Quite difficult ❑ Very difficult ❑

とても簡単❑ 簡単❑ ちょうどいい❑ 難しい❑ とても難しい❑

2. The test took… *(please tick one box)*

テストを受ける際にかかった時間について（以下選択ください）

Not enough time ❑ About the right amount of time ❑ Too much time ❑

十分に時間がない❑ ちょうどいい時間❑ ありまる時間❑

3. The questions toward the beginning of the test were easier than the questions toward the end of the test *(please tick one box)*

テストの前半に質問の方が後半の質問に比べて簡単であった（以下選択ください）

True ❑ False ❑ I'm not sure ❑

正しい❑ 正しくない❑ わからない❑

4. I didn't guess any of the answers *(please tick one box)*

どの質問も推測せずに回答した（以下選択ください）

True ❑ False ❑

正しい❑ 正しくない❑

5. I feel confident in… *(please tick one box)*

テストの結果の自信は（以下選択ください）

About 20% ❑ About 40% ❑ About 60% ❑ About 80% ❑ About 100% ❑

約２０％❑　約４０％❑　約６０％❑　約８０％❑　約１００％❑

…of my answers.

6. I gave up trying to answer the questions after a certain point *(please tick one box)*

途中で回答するのを諦めた（以下選択ください）

True ❑　　False ❑

正しい❑　正しくない❑

6a. If you answered 'true' above, *when* did you give up trying to answer the questions? *(please tick one box)*

上記の６で「正しい」を選択した場合にはテストのどの部分で諦めましたか。（以下選択ください）

Near the start ❑

開始直後❑

About a quarter of the way through ❑

４分の１程度❑

Near the middle ❑

ちょうど真ん中程度❑

About three quarters of the way through ❑

４分の３程度❑

Near the end ❑

終了間際❑

7. Please use this space to write any further comments that you'd like to make about the test

他に何かコメントがありましたら、お願いします。ありがとうございます。

Appendix 10: Marking Criteria

5 段階評価（パターン1）	最優	優	良	可	不可
5 段階評価（パターン2）	A	B	C	D	F
5 段階評価（パターン3）	100～90点	89～80点	79～70点	69～60点	59点以下
GPA	4	3	2	1	0

Adopted from the content teacher via email

Appendix 11: The syllabus of “Concepts in Chemistry” (EMI version)

icuMAP

2018 Winter Term Concepts in Chemistry 化学概論
CHM101

Language of Instruction: E

コウ ジュリアン R. (KOE, JULIAN R.)

Credit (単位) : 2

Period(s) 時限数	Lec.(講義)	Sem.(演習)	Lab.(実験実習)	Exe.(実技)	Intensive(集中講義)
	2				

Description (概要)

Origin and occurrence of the chemical elements, atomic structure and the periodic table. Atomic approach to the structure of materials and fundamental properties of inorganic and organic compounds.

化学における物質観,元素の起源と分布,原子構造と周期律,物質の構造と化学結合,基本的な無機・有機物質の構造・物性・化学反応などについて,主として原子論的立場から考察する。

Chemistry is the study of substances. What they are, how they change and how they can be controlled. We are surrounded by substances and our lives depend on chemical changes of substances. Chemistry concerns the whole of nature including biological activity, the environment, pharmaceuticals and new materials. As our lives depend increasingly on various substances, it is important to develop our scientific literacy so that we can understand these substances using the Scientific Method from a chemical point of view, not only for chemistry majors but everyone. The courses Concepts in Chemistry and General Chemistry (Foundation of Chemistry courses I and II) provide students with the broad fundamentals of chemistry, and are especially important for chemistry majors as they provide the basis for the more advanced courses.

化学は物質の学問です。物質とはどういう物か、どのように変化し、どのように我々がコントロールできるかを学ぶ 学問です。私たちは物質に囲まれています。たとえば生命活動はまさに物質の化学変化そのものです。化学は生体活 動、自然環境、薬剤や新素材など、私たちを取り巻くすべてに関わり、その依存度は益々高まっています。従って、 科学的に物質を理解するためのリテラシーは化学専修生のみならず、すべての学生にとって必要だと考えています。 化学概論と基礎化学の2コースは化学の一般的な基礎のI とIIとして位置づけられ、次に続くより専門的なコースの基 本になる重要な科目です。

Learning Goals (学修目標)

Contents (内容)

To understand the fundamental principles and knowledge of chemistry as a foundation of the natural sciences.

自然科学の基礎の一つとして、化学の基本的概念や基本知識を理解する。

Electrons in Atoms Periodic Table and Atomic Properties Chemical Bonding I: Lewis theory, covalent bonding, polarity, resonance, molecular shapes, bond order length and energy Chemical Bonding II: valence bond theory, hybridisation, multiple bonds, molecular orbital theory, delocalisation, bonding in metals Acids and Bases: Arrhenius, Bronsted, pH, strong/weak acids and bases ions, Lewis Organic Chemistry

Electrons in Atoms 原子の中の電子 Periodic Table and Atomic Properties 周期律表と原子の性質 Chemical Bonding I: Lewis theory, covalent bonding, polarity, resonance, molecular shapes, bond order length and energy 化学結合I:ルイス論、共有結合、極性、共鳴構造、分子構造、結合次数、結合距離、結合エネルギー Chemical Bonding II: valence bond theory, hybridisation, multiple bonds, molecular orbital theory, delocalisation, bonding in metals 化学結合II:原子価結合論、混成化、多重結合、分子軌道論、極在化、金属の結合 Acids and Bases: Arrhenius, Bronsted, pH, strong/weak acids and bases ions, Lewis 酸と塩基:アーレニウスの定義、ブロンステッドの定義、pH、酸性・塩基性の強さ、ルイス酸・塩基Organic Chemistry 有機化学

Language of Instruction (教授言語の詳細)

Lecture: English Readings/Materials: English Tests/Quizzes/Assignments: Both English Discussions/Presentations/Other learning activities: English Communication with the instructor: Both Japanese and English

Grading Policy (成績評価基準)

Expected study hour outside class (授業時間外学習)

References (参考文献)

Course work, Final Exam, Attendance 宿題、期末試験、出席

140 minutes per week.

Chemistry: The Central Science Global Edition, 13th edition Brown et al. Pearson

ISBN-10 1-292-05771-8 ISBN-13 978-1-292-05771-2

Notes (注意事項)

Schedule (スケジュール)

The entire course (including homeworks and exam) will be carried out in English. The text book is also in English (but a Japanese edition is also available).

3/TU,3/TH

Appendix 12: The syllabus of “Concepts in Chemistry” (JMI version)

icuMAP

2018 Winter Term

Concepts in Chemistry 化学概論

CHM101

Language of Instruction: J

Description (概要)

久保 謙哉 (KUBO, KENYA)/田 旺帝 (CHUN, WANG JAE)/河野 泰朗 (KAWANO, YASURO)

Credit (単位) : 2

Period (s) 時限数	Lec.(講義)	Sem.(演習)	Lab.(実験実習)	Exe.(実技)	Intensive(集中講義)
	2				

Origin and occurrence of the chemical elements, atomic structure and the periodic table. Atomic approach to the structure of materials and fundamental properties of inorganic and organic compounds.

化学における物質観,元素の起源と分布,原子構造と周期律,物質の構造と化学結合,基本的な無機・有機物質の構造・物性・化学反応などについて,主として原子論的立場から考察する。

Chemistry is the study of substances. What they are, how they change and how they can be controlled. We are surrounded by substances and our lives depend on chemical changes of substances. Chemistry concerns the whole of nature including biological activity, the environment, pharmaceuticals and new materials. As our lives depend increasingly on various substances, it is important to develop our scientific literacy so that we can understand these substances using the Scientific Method from a chemical point of view, not only for chemistry majors but everyone. The courses Concepts in Chemistry and General Chemistry (Foundation of Chemistry courses I and II) provide students with the broad fundamentals of chemistry and are especially important for chemistry majors as they provide the basis for the more advanced courses.

化学は物質の学問です。物質とはどういう物か、どのように変化し、どのように我々がコントロールできるかを学ぶ 学問です。私たちは物質に囲まれています。たとえば生命活動はまさに物質の化学変化そのものです。化学は生体活 動、自然環境、薬剤や新素材など、私たちを取り巻くすべてに関わり、その依存度は益々高まっています。従って、 科学的に物質を理解するためのリテラシーは化学専修生のみならず、すべての学生にとって必要だと考えています。 化学概論と基礎化学の2コースは化学の一般的な基礎のI とIIとして位置づけられ、次に続くより専門的なコースの基本になる重要な科目です。

Learning Goals (学修目標)

Contents (内容)

To understand the fundamental principles and knowledge of chemistry as a foundation of the natural sciences.

自然科学の基礎の一つとして、化学の基本的概念や基本知識を理解する。

(WD = Wang-Jae Chun, KK = Kenya Kubo, JK = Junji Kobayashi) (WD) 原子の電子構造 (WD) 元素の周期的な性質 (WD) 化学結合(WD) 分子の形と結合理論 (KK) 酸塩基平衡 (KK) 有機化学

Language of Instruction (教授言語の詳細)

Lecture: Japanese Readings/Materials: Japanese Tests/Quizzes/Assignments: Japanese Discussions/Presentations/Other learning activities: Japanese Communication with the instructor: Both Japanese and English

Grading Policy (成績評価基準)

Expected study hour outside class (授業時間外学習)

Coursework, Final Exam, Active participation 宿題、期末試験、Active participation

References (参考文献)

ブラウン 一般化学 I(原書13版)、丸善出版(2016) 萩野和子 監訳 ISBN: 978-4621300107

ブラウン 一般化学 II(原書13版)、丸善出版(2016) 萩野和子 監訳 ISBN: 978-4621300114

Notes (注意事項)

Schedule (スケジュール)

URL

Contact (Phone, E-mail, Office Hour)

Plagiarism, cheating, and other forms of academic dishonesty will not be tolerated and will be punished according to ICU regulations.
https://moodle.icu.ac.jp/27/pluginfile.php/47714/mod_resource/content/1/academic_integrity_ej.pdf

3/TU,3/TH

wchun@icu.ac.jp

140 minutes per week (minimum).

Appendix 13: Start list of themes

Main themes and 1st sub-theme	2nd sub-theme	EMI/JMI data
Challenge: Speaking		
Accuracy/ fluency	Grammar	EMI
	Speech flow	EMI
Presentation	Content issue	EMI/JMI
	Language issue	EMI
	Organisation issue	EMI/JMI
	IT literacy skills	EMI/JMI
Asking a question	Content issue	EMI/JMI
	Language issue	EMI
	Confidence issue	EMI/JMI
Classroom discussions	Language issue	EMI
	Confidence issue	EMI/JMI
Challenge: Writing		
Laboratory report	Content issue	EMI/JMI
	Language issue	EMI
Essay	Language issue	EMI
	Organisation issue	EMI/JMI
	References and citations	EMI/JMI
Vocabulary	Technical vocabulary	EMI/JMI
	Academic vocabulary	EMI
Challenge: Reading		
Inadequate vocabulary size		EMI/JMI
Slow reading speed	Technical vocabulary	EMI/JMI
	Grammar	EMI
Reliance on translation	Technical vocabulary	EMI
Inefficient reading	Prior content knowledge	EMI/JMI
	Language issue	EMI
Challenge: Listening		
Low concentration	Missing information	EMI/JMI
Humour and jokes from the lecturer		EMI
Content and language issues	Additional language problems	EMI

	Low understanding of lectures	EMI
Low understanding of technical terms	New technical terms	EMI/JMI
	English spelling issue	EMI
Challenge: Language use		
Reliance on code switching	Teacher L1 use	EMI
	Student L1 use	EMI
Simplified teacher language		EMI
Teachers' English proficiency		EMI
Challenge: Quality of learning		
Simplified content		EMI/JMI
Reduced content coverage		EMI/JMI
Challenge: Others		
Loss of L1		EMI
Inadequate language support scheme		EMI
Inadequate computer literacy		EMI/JMI
Challenge: Grades		
Poor management of workload		EMI/JMI
Low performance of the final course exam	Inadequate content knowledge	EMI/JMI
Low performance of Comment Worksheet	Low content knowledge	EMI
Motivation and effort as mediators of GPA	Diligence and determination	EMI/JMI
Challenge: Adjustment		
Adjustment to the new educational culture	General university learning culture	EMI/JMI
	EMI learning culture	EMI
Persistent problems	Long-lived issue	EMI
	Short-lived issue	EMI/JMI
Success: Content learning		
New modes of thinking		EMI
Knowledge transformation and application		EMI/JMI
Success: English knowledge		
Listening		EMI
Reading		EMI/JMI
Speaking		EMI
Writing		EMI/JMI

Success: Motivation		
Study abroad		EMI
Professional goal		EMI/JMI
In-depth learning		EMI/JMI
Success: Learning resource		
Bilingual resources		EMI
English online resources		EMI
Success: Career advancement		
Increased content knowledge		EMI/JMI
International career goals		EMI
Global competence		EMI
Study abroad		EMI
Success: others		
Value for money		EMI
Effective communication with peers and teachers		EMI/JMI

Appendix 14: Final list of themes

Main themes and 1st sub-theme	2nd sub-theme	EMI/JMI data	Time 1/2 data
Challenge: Comprehension of the course textbook			
Inadequate vocabulary size	Technical vocabulary	EMI/JMI	Time1/2
	Academic vocabulary	EMI	Time1/2
Slow reading speed	Technical vocabulary	EMI/JMI	Time1
	Grammar issue in English	EMI	Time1
Reliance on translation	Technical vocabulary	EMI	Time1
	Whole paragraphs	EMI	Time1
Inefficient reading	Prior content knowledge	EMI/JMI	Time1
	Language issue	EMI	Time1/2
Challenge: Comprehension of the course lectures			
Low concentration	Missing information	EMI/JMI	Time1
	Low understanding of concepts	EMI	Time1

Low comprehension of humour and jokes from the lecturer		EMI	Time1/2
A 'double whammy' of content and language issues	Additional language problems	EMI	Time1
	Low understanding of lectures	EMI	Time1
	Low note taking performance	EMI	Time1
Low understanding of technical terms	New technical terms	EMI/JMI	Time1/2
	Dual meaning terms	EMI	Time1
	English spelling issue	EMI	Time1
Challenge: Final Grade			
Poor management of workload		EMI/JMI	Time1
Low performance of the final course exam	Inadequate content knowledge	EMI/JMI	Time2
	Misunderstanding exam instructions/ questions	EMI	Time2
	Insufficient exam time	EMI	Time2
Low performance of Comment Worksheet	Low content knowledge	EMI	Time1
	L2 English writing issue	EMI	Time1
Motivation and effort as mediators of Final Grade	Diligence and determination	EMI/JMI	Time1
	Exam strategy	EMI/JMI	Time1
Challenge: Adjustment to the new academic and cultural conventions of the Chemistry course			
Adjustment to the new educational culture	General university learning culture	EMI/JMI	Time1
	EMI learning culture	EMI	Time1
Persistent problems	Long-lived issue	EMI	Time2
	Short-lived issue	EMI/JMI	Time2
Success: English knowledge			
Listening		EMI	Time1/2
Reading		EMI/JMI	Time1/2
Speaking		EMI	Time1/2
Vocabulary		EMI/JMI	Time1/2
Success: Motivation			
Study abroad		EMI	Time1
In-depth learning		EMI/JMI	Time1

Success: Learning resource			
Bilingual resources		EMI	Time1
English online resources		EMI	Time1
Success: Career advancement			
Increased content knowledge		EMI/JMI	Time1/2
International career goals		EMI	Time1/2
Global competence		EMI	Time1/2
Study abroad		EMI	Time1/2

Appendix 15: Sample email requesting permission to conduct study

Dear Professor/ Dr. [name] at xx university in Japan,

My name is Ikuya Aizawa. I am currently undertaking a DPhil in Education at Department of Education at Oxford University, working under Dr. Heath Rose and Dr Jess Briggs. I am writing to request permission to collect data at your university for a study that I am conducting on a cost-benefit investigation of the English-medium instruction in Japan. My dissertation project specifically aims to examine the following questions: 1. Is there any effect of the medium of instruction on students' content knowledge acquisition? 2. What are the differences (if any) between successful and unsuccessful EMI learners? 3. What is the relationship between students' English language proficiency and content knowledge acquisition? 4. What language-related and content-related learning challenges do EMI students encounter?

If you permit me to collect data, it will involve me contacting students who are currently/ have enrolled in your university to take part in the study. The students will be asked to complete a questionnaire about self-perceived issues with academic English use and their English language use in EMI lectures. Additionally, students who have agreed to take part in an interview will be asked to participate in a 30-minute interview about their English use. In total, the study will take approximately 40 minutes of each student's time (a questionnaire and interview). Classroom observations will also be carried out during the semester. Please see the Participant Information Leaflet attached to this email for further details on the study, which will be provided to all potential participants.

Permission to conduct this study is voluntary. Even if you grant permission to collect data now, you can withdraw this permission at any time without any consequences of any kind. If you chose not to give permission it will have no adverse consequences for your relationship with the Department of Education at the University of Oxford.

If you have any questions about this research, please do not hesitate to contact me.

Yours sincerely,

Ikuya Aizawa

DPhil in Education

Department of Education

University of Oxford

15 Norham Gardens

Oxford, OX2 6PY

TEL: +44-75829-50701 (UK) ; +81-90-5813-8353 (Japan)

Appendix 16: Participant Information Leaflet

1. Study title: Effects of English as Medium of Instruction (EMI) on content knowledge acquisition in Japanese Higher Education

2. Background to the study: You are invited to participate in a research project that is designed to investigate the effects of English medium instruction in EMI universities in Japan. The study is being carried out by Ikuya Aizawa, who is a current postgraduate student enrolled DPhil in Education in the Department of Education at the University of Oxford.

3. What have I been invited? You have been invited to take part in this research because you are currently studying at xx university in Japan, where English is used as a medium of instruction. The study includes students and from 1st to 4th year who use English to learn academic subjects, who speak Japanese as their mother tongue.

4. Do I have to take part? Your participation is voluntary. You are also very welcome to ask the researcher, Ikuya Aizawa, for more information at any time. Even if you agree to participate now, you can withdraw at any time, without giving a reason, and without any consequences or any academic penalty. If you withdraw, your data will be removed from the study. I hope that you will agree to take part in to this research as your contribution and information will be very useful to the project and will provide pedagogical application to EMI education policy in Japan.

5. What will happen in the study? If you agree to participate, this will involve you completing a questionnaire about self-perceived issues with academic English use in an EMI context and another questionnaire about English use in an EMI lectures and content and vocabulary knowledge test. Additionally, if you agree to participate in an interview, you will be asked to undertake a 30-minute interview about your academic English use.

6. Are there any potential risks? If you elect to take part in both the interview and the questionnaires, the study may take up to 40 minutes of your time, which may be mentally tiring for you. However, you will be asked to take part in an interview on a different day from the day that you will be asked to take a questionnaire.

7. What happens to the research data provided? Any information or data that I obtain from you during this research which can be identified with you will be treated confidentially. I will do this by replacing your name with a number (e.g. participant 1), and deleting any identifying information from

written transcripts of the interview. All data, including the original audio recordings of the interviews, will be kept on the researcher's password-protected computer, personal password-protected online storage folders, or password-protected USB storage devices. The original digital audio recording and all copies will be available only to the researcher for transcription purposes. Any hard copies of data, such as consent forms, and research notes, will be kept confidentially. If you should wish to see the transcripts of your interview, you can request them at any time. Excerpts from the transcription may be used in future conference presentations and publications, providing they do not contain sensitive or personally identifying information. Audio recordings will be kept for three years following any publications from the project, and then deleted.

8. Will the research be published? The research might be published in academic journals and conference papers in the future, which may be publicly available. The researcher will take steps ensure that any data from you that appears in a publication will not be identifiable. Steps may include use of pseudonyms, the changing of place names, and deletion of personal information. You may request to be notified when data from this project is published.

9. Who has reviewed this project? This project has been reviewed by, and received ethics clearance through, the University of Oxford Central University Research Ethics Committee.

10. Who do I contact if I have a concern about the study or I wish to complain?

If you have a concern about any aspect of this project, please speak to the researcher (Ikuya Aizawa, ikuya.aizawa@education.ox.ac.uk, +44-(0)75829-50701), who will do his best to answer your query. The researcher should acknowledge your concern within 10 working days and give you an indication of how he intends to deal with it. If you remain unhappy or wish to make a formal complaint, please contact the chair of the Research Ethics Committee at the University of Oxford (using the contact details below) who will seek to resolve the matter in a reasonably expeditious manner:

- Chair, **Social Sciences & Humanities Inter-Divisional Research Ethics Committee**; Email: ethics@socsci.ox.ac.uk; Address: Research Services, University of Oxford, Wellington Square, Oxford OX1 2JD

Appendix 17: Consent form

STUDY TITLE: Effects of English as Medium of Instruction (EMI) on content knowledge acquisition in Japanese Higher Education

RESEARCHER DETAILS: Ikuya Aizawa, who is a current postgraduate student enrolled DPhil in Education in the Department of Education at the University of Oxford

PURPOSE OF STUDY: You are invited to participate in a research project that is designed to investigate the effect of the recent EMI initiative in EMI universities in Japan

		Initial here
1	I have read the information sheet, and had the opportunity to ask questions. If I asked questions, I have received satisfactory answers.	
2	I understand that my participation is voluntary and I am free to withdraw at any time, without reason, and without any adverse consequences or any academic penalty.	
3	I understand audio recordings of my interview will only be made available to the researcher for transcription purposes	
4	I understand that my data will be stored in password protected electronic files, and that original recordings will be deleted three years after publication of the research	
5	I understand that data from the project may appear in academic journals and conference papers in the future, but the researcher will take steps to ensure this data will not be identifiable to me	
6	I understand how to raise concerns with the researcher or make a complaint directly to the Research Ethics Committee at the University of Oxford	
7	I consent to being audio-recorded	
8	I understand that excerpts from my interviews may be used in publications and conference presentations in the future	
9	I agree to take part in the study	

Name of Participant: _____

Signature: _____ Date: _____

Name of researcher: _____

Signature: _____ Date: _____

Appendix 18: Z-score table for skewness and kurtosis

EMI (n=27), JMI (n=26), total sample (n=53)

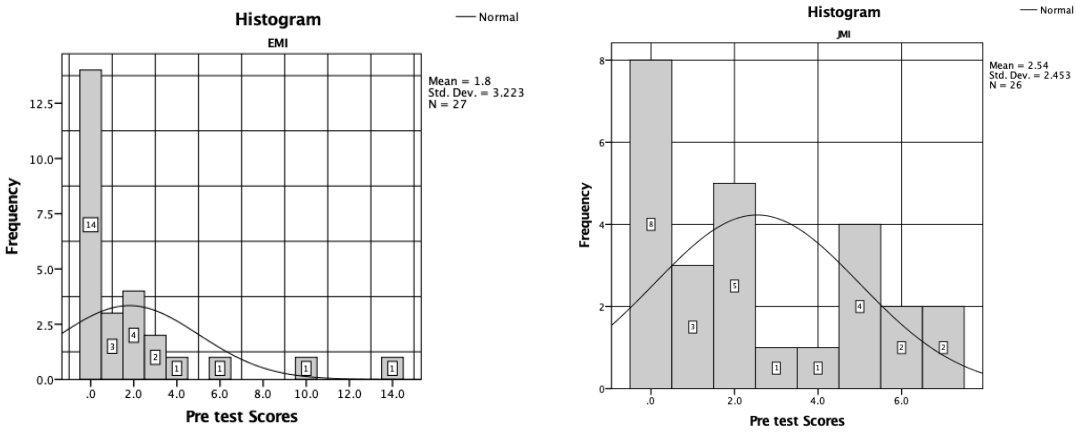
	MOI	N	*Z skewness**	*Z kurtosis**
Pre-test scores	EMI	27	5.90	8.29
	JMI	26	1.16	-1.33
	Total	53	0.18	0.16
Post-test scores	EMI	27	-1.52	-0.42
	JMI	26	-0.03	-1.70
	Total	53	-1.12	-0.57

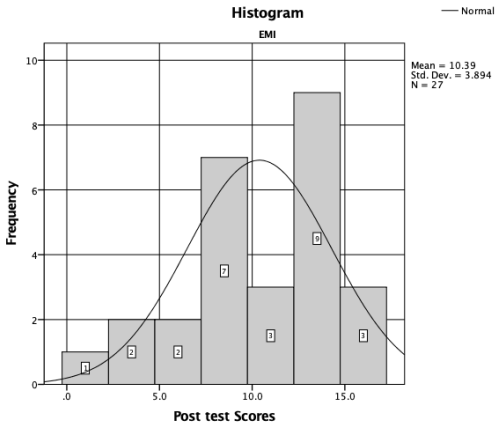
EMI (n=20), JMI (n=26), total sample (n=46)

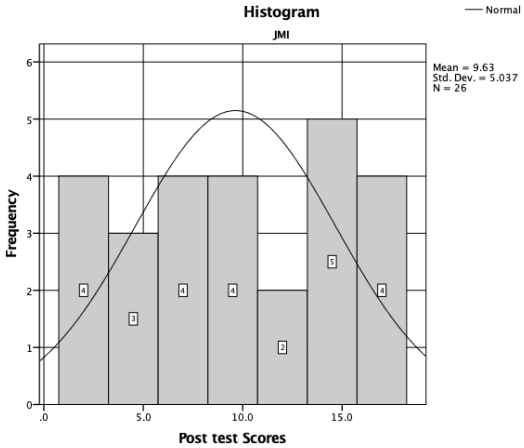
	MOI	N	*Z skewness**	*Z kurtosis**
Pre-test scores	EMI	20	4.56	5.25
	JMI	26	1.16	-1.33
	Total	46	4.95	5.04
Post-test scores	EMI	20	-1.29	0.72
	JMI	26	-0.03	-1.70
	Total	46	-1.15	-1.60

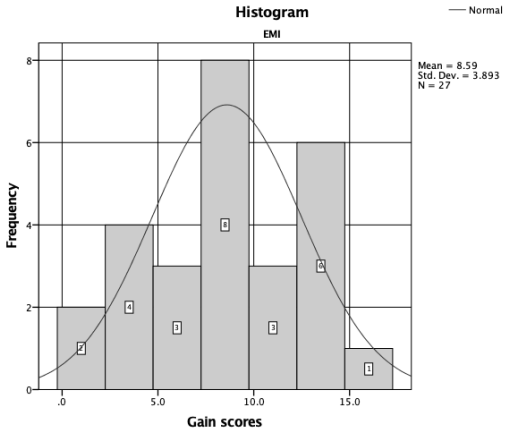
*Z-scores of skewness and kurtosis were computed by dividing the skewness and kurtosis measures by the respective standard error.

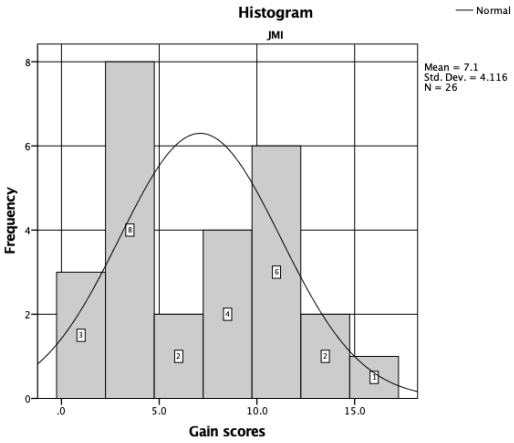
Appendix 19: Histograms: EMI and JMI students' pre-test, post-test and gain scores

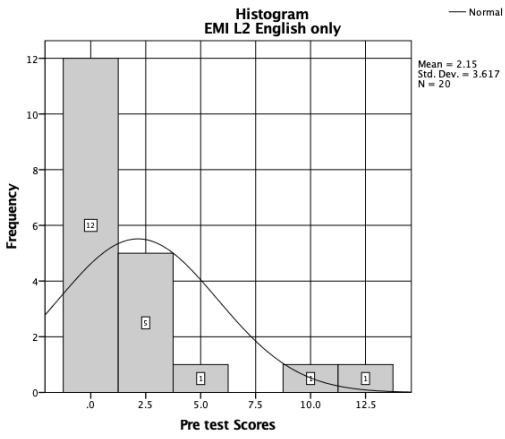


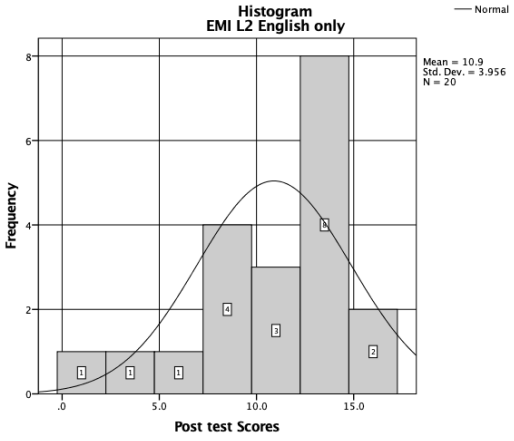

Histogram
EMI
Normal
Mean = 10.39
Std. Dev. = 3.894
N = 27
Frequency
Post test Scores

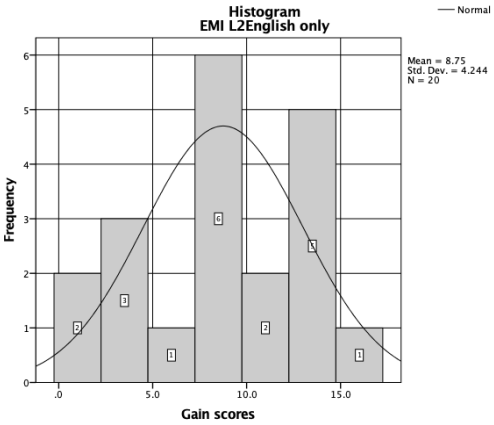

Histogram
JMI
Normal
Mean = 9.63
Std. Dev. = 5.037
N = 26
Frequency
Post test Scores


Histogram
EMI
Normal
Mean = 8.59
Std. Dev. = 3.893
N = 27
Frequency
Gain scores


Histogram
JMI
Normal
Mean = 7.1
Std. Dev. = 4.116
N = 26
Frequency
Gain scores


Histogram
EMI L2 English only
Normal
Mean = 2.15
Std. Dev. = 3.617
N = 20
Frequency
Pre test Scores


Histogram
EMI L2 English only
Normal
Mean = 10.9
Std. Dev. = 3.956
N = 20
Frequency
Post test Scores



Appendix 20: ANCOVA assumption- Linearity for ANCOVA

Figure: Scatterplot (whole sample)

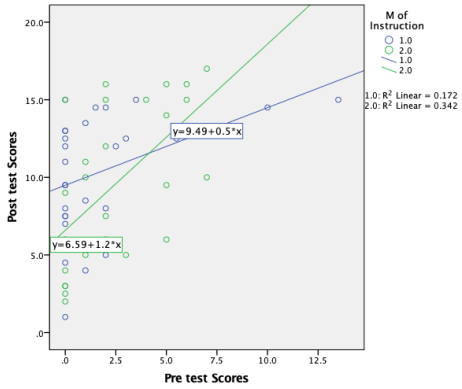
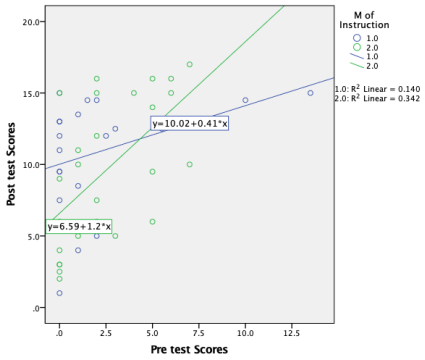


Figure: Scatterplot (L2 English only)



Appendix 21: Gain score analyses

Between groups comparison: whole sample

The gain scores (i.e. gain = post-test–pre-test) were compared to determine whether there were significant differences in knowledge gains between the groups. The independent variable was a nominal variable with dichotomous categories (English or Japanese). The dependent variable (gain scores) was measured on a ratio scale (continuous variable) as equal intervals were obtained between all the properties being measured (Field, 2018).

Prior to conducting the analyses, normality of the dependent variable was inspected using numerical methods through a z-score of skewness and kurtosis (see the table below) and visual methods of the distribution curve the approximate shape of histograms (Sullivan & Artino, 2013). Based on these inspections, gain scores were reasonably normally distributed.

Table: Descriptive statistics – gain scores for the EMI (n=27), JMI (n=26) and total sample (n=53)

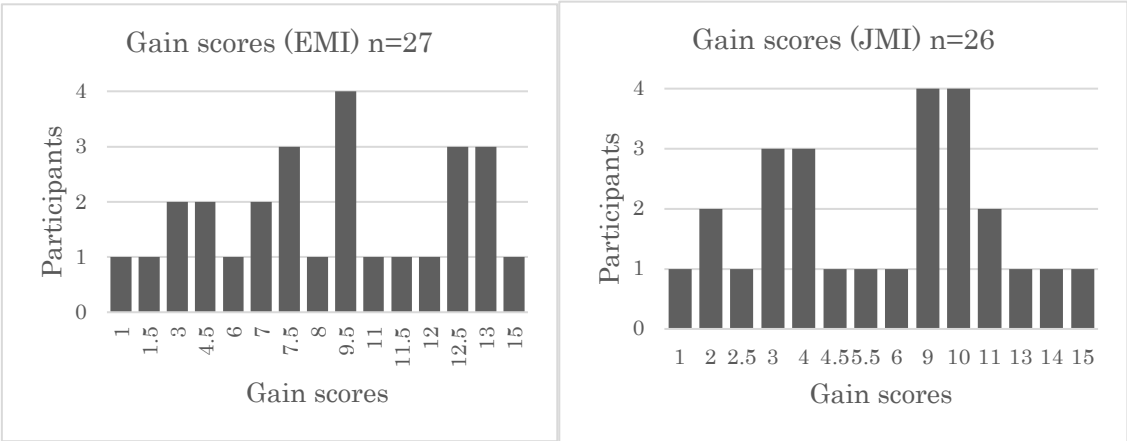
	MOI	**N**	**Min**	**Max**	**Mean**	**Median**	**SD**	***Z skewness***	***Z kurtosis***
Gain scores	EMI	27	1	15	8.59	9.500	3.15	-0.83	-0.93
	JMI	26	1	15	7.10	7.500	4.16	0.50	-1.32
	Total	53	1	15	7.89	7.500	4.03	-4.04	-0.55

As I did not specify which group would attain higher gain scores in a particular direction, a two-tailed non-directional t-test is used to examine whether the difference lies within either of the two tails of the distribution. The statistical hypotheses are as follows:

• H_0 is that there are no mean-level differences in gain scores of JMI and EMI groups **(H_0: $\mu_{EMI\ students} - \mu_{JMI\ students} = 0$)**.

• H_1 is that there are mean-level differences in gain scores of JMI and EMI groups (**H_1: $\mu_{EMI\ students} - \mu_{JMI\ students} \neq 0$**).

Levene's test confirmed the equality of variances in the samples (homogeneity of variance), F (.698)=0.407, n.s. The EMI and JMI group achieved a mean gain score of 8.59 (SD=3.15) and 7.10 (SD=4.16), respectively, indicating that the EMI students had higher gains of 8.7% than the JMI cohort. However, an independent samples t-test indicated no evidence that the null hypothesis was false; H_0 was not rejected and H_1 could not be accepted; ($t_{[51]} = 1.36$; $p = .18$, two-tailed). The mean differences in gain scores were not significant between the groups. Thus, there were no significant differences in content knowledge gains between EMI and JMI students.



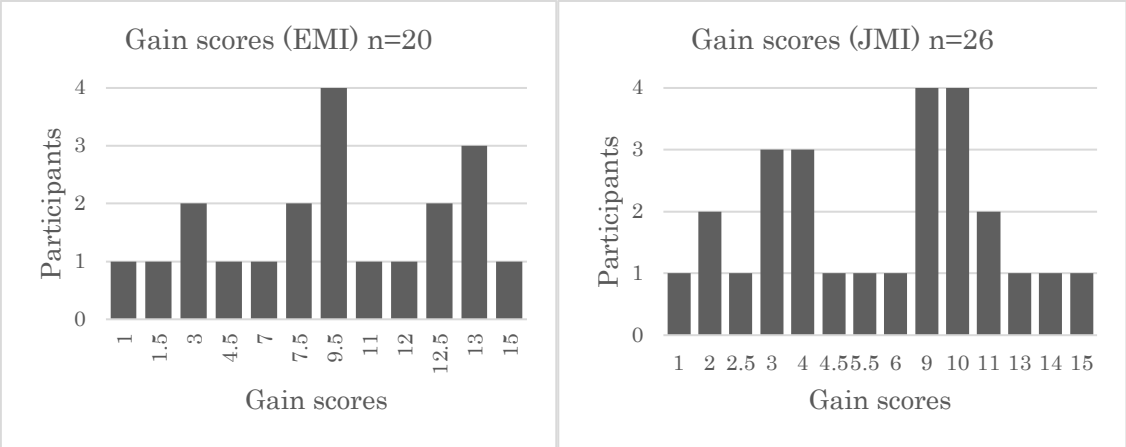
Between groups comparison: L2-English only

The gain scores were also compared to investigate whether there were significant differences in score gains between the EMI (excluding those who learnt through L1 English) and JMI groups. Adopting the above-mentioned numerical and visual methods, normality of the gain scores was confirmed. Levene's test confirmed the equality of variances in the samples (homogeneity of variance), F (.088)=0.769, n.s. The mean gain score for the 20 EMI students was 8.75 while the mean score for the 26 JMI students was 7.10, suggesting higher gain scores for the EMI than the JMI group. However, an

independent samples t-test demonstrated that the mean differences in gain scores were not significant between the groups ($t_{[44]} = 1.333$; $p = .189$, two-tailed). Therefore, the L2 English students in the EMI cohort and their JMI counterparts taught in their L1 were not significantly different in terms of the gain in their content knowledge.

Table: Descriptive statistics – test scores for the EMI (n=20), JMI (n=26) and total sample (n=46)

	MOI	N	Min	Max	Mean	Median	SD	*Z skewness*	*Z kurtosis*
Gain scores	EMI	20	1	15	8.75	9.50	4.24	-1.02	-0.89
	JMI	26	1	15	7.10	7.50	4.16	0.50	-1.32
	Total	46	1	15	7.03	7.5	4.06	-0.24	-1.84



Appendix 22: 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 23: Tests of normality

Formal test of normality check (including L1 English students n=27 EMI, 26 JMI)

	M of Instruction	Shapiro-Wilk		
		Statistic	df	Sig.
Pre test Scores	1.0	.615	27	.000
	2.0	.862	26	.002
Post test Scores	1.0	.920	27	.060
	2.0	.910	26	.076
Gain Scores	1.0	.949	27	.203
	2.0	.925	26	.071

Formal test of normality check (excluding L1 English students n=20 EMI, 26 JMI)

	M of Instruction	Shapiro-Wilk		
		Statistic	df	Sig.
Pre test Scores	1.0	.647	20	.000
	2.0	.862	26	.002
Post test Scores	1.0	.875	20	.054
	2.0	.910	26	.026
Difference	1.0	.921	20	.105
	2.0	.925	26	.061

Appendix 24: Matrices of pre-test score, challenges and post-test scores

Table: Matrices of post-test scores according to pre-test scores (n=13 EMI)

Prior knowledge (Pre-test scores) / Post knowledge (Post-test scores)	**Post-test low (1-6)** (n=3)**	**Post-test medium (6-12)** (n=6)**	**Post-test high (12-17)** (n=4)**
Pre-test 0 (n=7)	Marii (Pre 0; Post 1)*	Rina (Pre 0; Post 7.5)* Saika (Pre 0; Post 7.5) Maxwell (Pre 0; Post 8) Minori (Pre 0; Post 9.5)* Charlotta (Pre 0; Post 9.5)	Emma (Pre 0; Post 12.5)
Pre-test 1-2 (n=3)	Kiley (Pre 1; Post 4) Hikaru (Pre 2; Post 5)*	Ayaka (Pre 1; Post 8.5)*	
Pre-test 3-4 (n=1)			Lisa (Pre 3; Post 12.5)
Pre-test 5-6 (n=1)			Tikamon (Pre 5.5; Post 12.5)

Pre-test 7-17 (n=1)			Miyazaki (Pre 13.5; Post 15)*

* Those who participated in both the first and second round of interviews
** Post-test scores: Low (0-40%), Medium (40-70%) and High (70-100%) adopted from the university's grading criteria

Table: Matrices of post-test scores according to challenges scores (n=13 EMI)

Mean challenges scores (speaking, reading, listening) / Content learning (Post-test scores)	**Post-test low (1-6)** (n=3)**	**Post-test medium (6-12)** (n=6)**	**Post-test high (12-17)** (n=4)**
Very difficult 2 (n=2)	Hikaru (Ch 2.84 ;Post 5)*	Rina (Ch 2.84 ;Post 7.5)*	
Difficult 3 (n=6)	Marii (Ch 3.65 ;Post 1)* Kiley (Ch 3.96 ;Post 4)	Ayaka (Ch 3.00 ;Post 8.5)* Minori (Ch 3.88 ;Post 9.5)* Charlotta (Ch 3.77 ;Post 9.5)	Tikamon (Ch 3.68 ;Post 12.5)
Easy 4 (n=4)		Maxwell (Ch 4.78 ;Post 8) Saika (Ch 4.56 ;Post 7.5)	Emma (Ch 4.67 ;Post 12.5) Lisa (Ch 4.88 ;Post 12.5)
Very easy 5 (n=1)			Miyazaki (Ch 5.00 ;Post 15)*

* Those who participated in both the first and second round of interviews
** Post-test scores: Low (0-40%), Medium (40-70%) and High (70-100%) adopted from the university's grading criteria
*** Challenges: The original challenges questionnaire in Evans and Morrison's (2011) study was localised to fit in the current research setting and the entire writing part was removed. The questionnaire data "were used to form a composite mean for each skill area; that is, a 'mean of means' derived from the items in each area" (p. 203). Students responded via a five-point Likert scale with answers ranging from 1 (very difficult) to 5 (very easy).

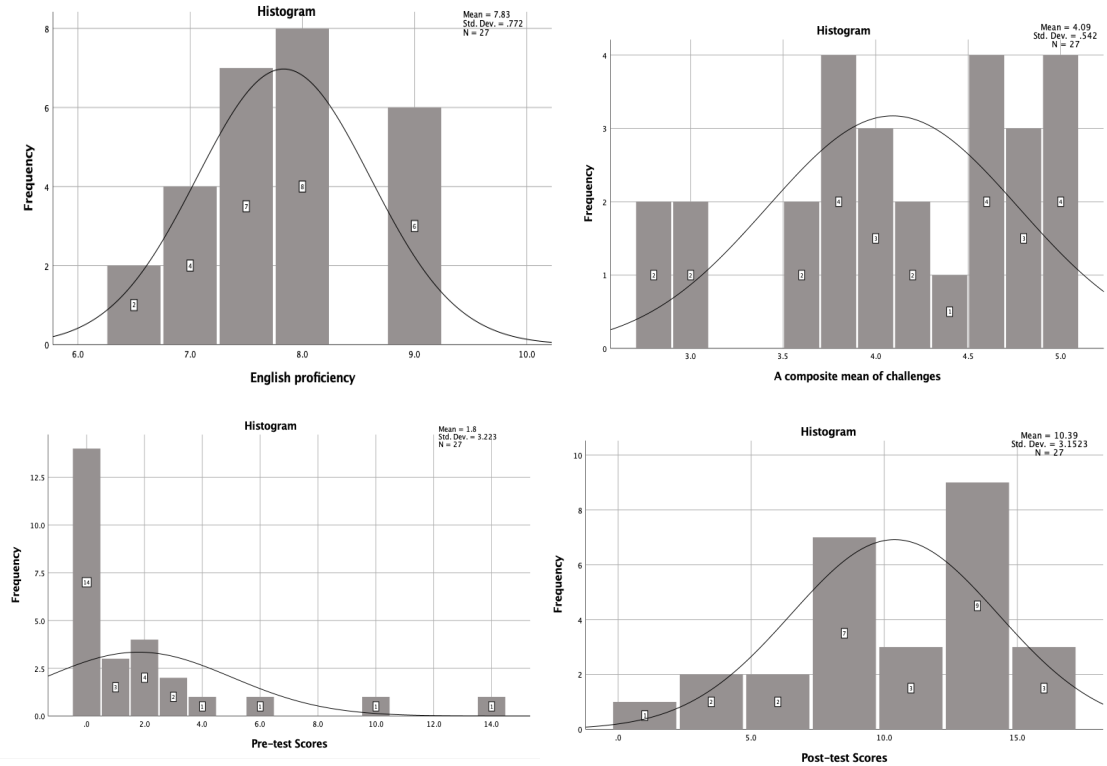
Appendix 25: Z-score table for skewness and kurtosis

The EMI student sample (n=27)

	Z skewness*	***Z kurtosis****
Pre-test scores	5.81	6.18
Post-test scores	-1.91	-0.57
Challenges	-1.01	-0.98
English proficiency	0.45	-0.84

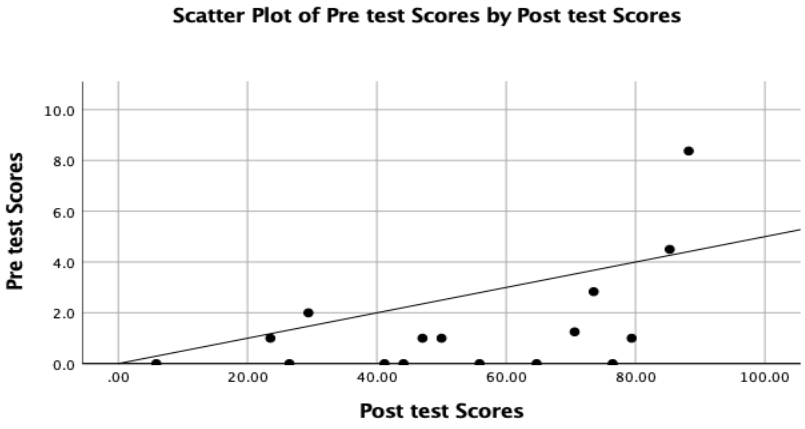
*Z-scores of skewness and kurtosis were computed by dividing the skewness and kurtosis measures by the respective standard error.

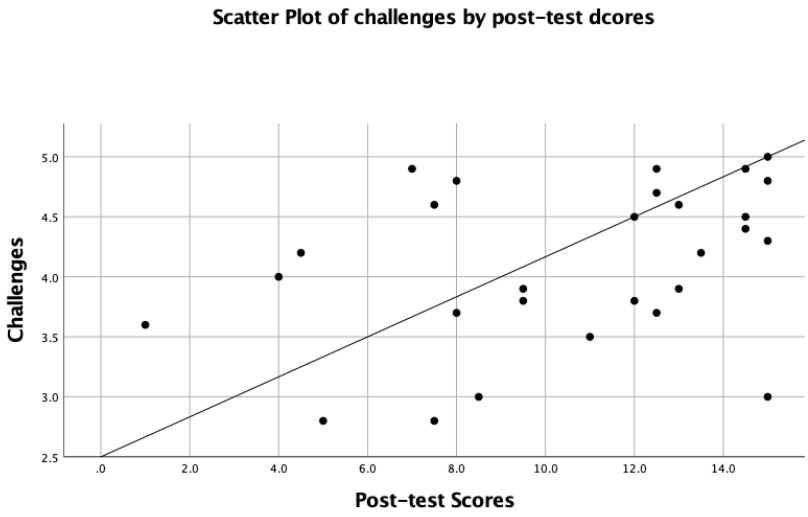
Appendix 26: Histograms of pre-test, post-test scores, challenges and proficiency

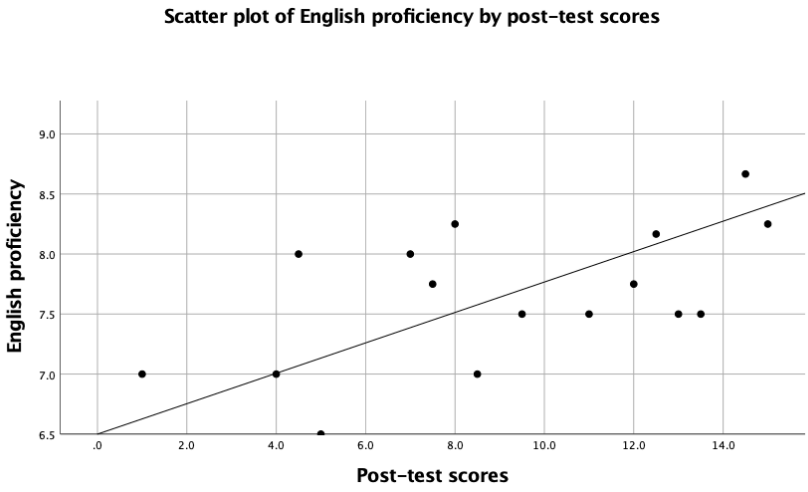


Appendix 27: Linearity assumption

Figure: Scatterplot of pre-test scores, post-test scores and challenges

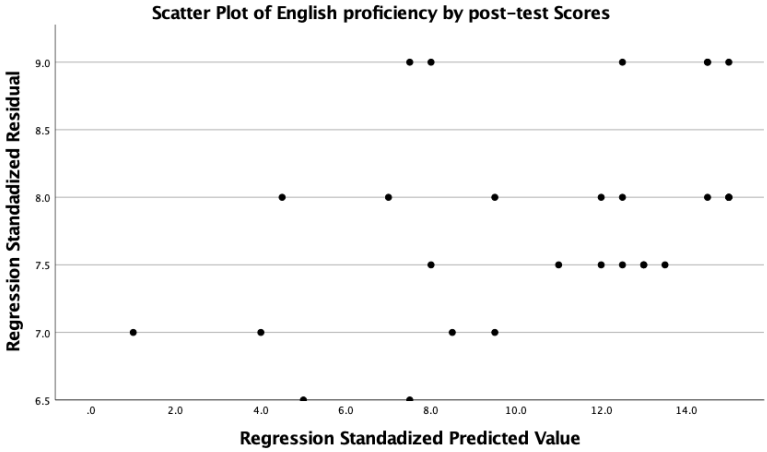


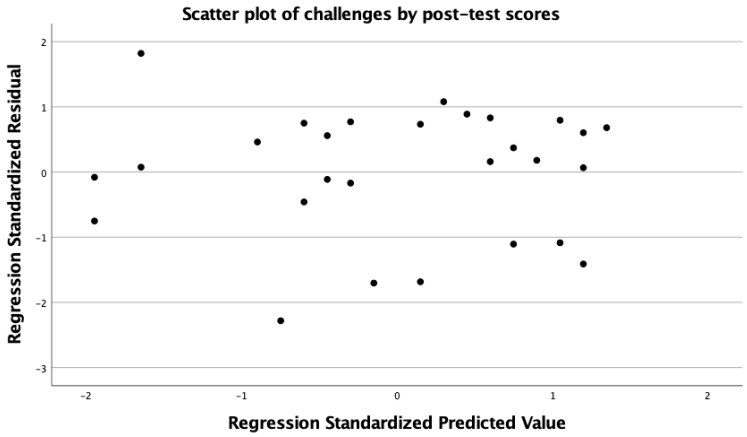


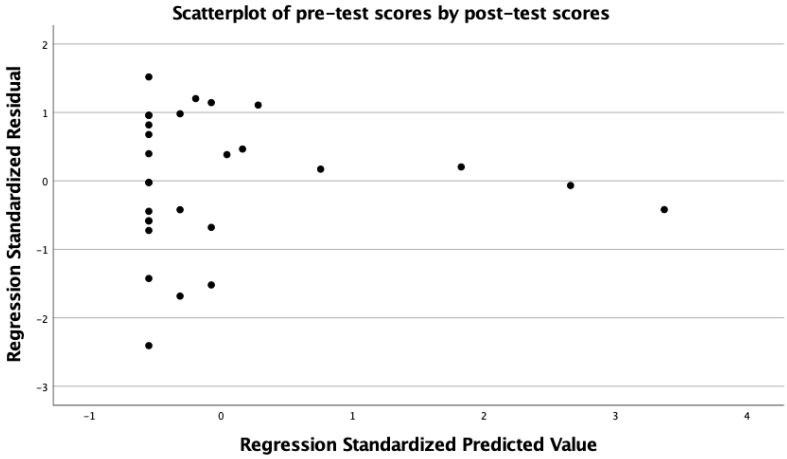


Appendix 28: Homoscedasticity assumption:

Figure: A scatterplot of standardized residuals versus standardized predicted values







Appendix 29: Normality of residuals (errors) assumption:

Figure: A normal probability plot of challenges

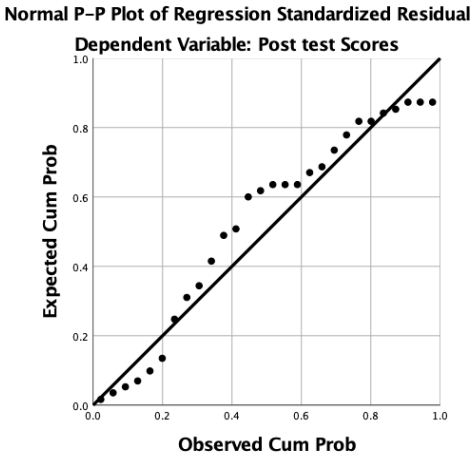


Figure: A normal probability plot of challenges

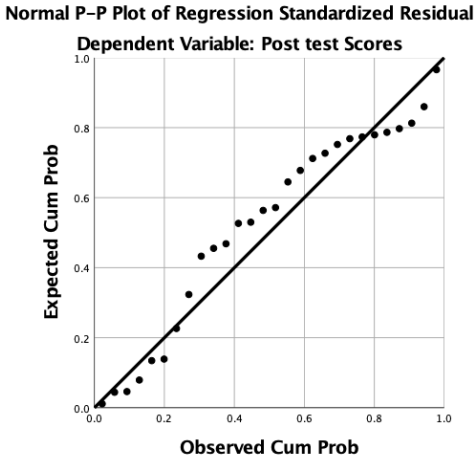
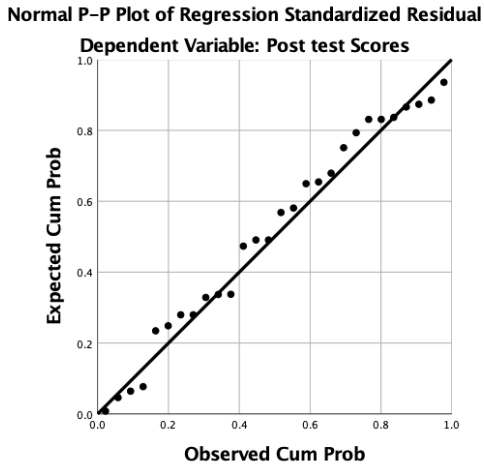


Figure: A normal probability plot of pre-test scores



References

Adams, T. L. (2003). Reading mathematics: More than words can say. *The Reading Teacher*, 56(8), 786–795.

Aguilar, M., & Muñoz, C. (2014). The effect of proficiency on CLIL benefits in Engineering students in S pain. *International Journal of Applied Linguistics*, 24(1), 1-18.

Airey, J. (2012). I Don't Teach Language." The Linguistic Attitudes of Physics Lecturers in Sweden. *AILA Review*, 25, 64–79.

Airey, J., & Linder, C. (2006). Language and the experience of learning university physics in Sweden. European Journal of Physics, 27(3), 553–560.

Aizawa, I., and H. Rose. (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. *Higher Education*, 77, 1125–1142.

Aizawa, I., & Rose, H. (2020). High school to university transitional challenges in English medium instruction in Japan. *System*, 95.

Aizawa, I., Rose, H., Thompson, G., & Curle, S. (Under Review 2020). Beyond the threshold: Exploring English language proficiency, linguistic challenges, and academic language skills of Japanese students in an English Medium Instruction programme. *Language Teaching Research*.

Alcott, B. (2017). Does Teacher Encouragement Influence Students' Educational Progress? A Propensity-Score Matching Analysis. *Research in Higher Education,* 58, 773–804.

Alexander, R. (2004). Towards dialogic teaching: Rethinking classroom talk. Cambridge: Dialogos.

Ali, N.L. (2013). A changing paradigm in language planning: English-medium instruction policy at the tertiary level in Malaysia. *Current Issues in Language Planning*, 14(1), 73–92.

Ali, H. (2020). Lecture comprehension difficulties experienced by Omani students in an English-medium engineering programme. *Cogent Arts & Humanities*, 7(1), Cogent arts & humanities, 01 January 2020, Vol.7(1).

Altay, M., & Erçin, N. (2020). Uncovering the Reflections of English Medium Instruction in Engineering Graduates' Career, *Sakarya University Journal of Education*, 10(3), 577-588.

Ament, J. R., & Pérez Vidal, C. (2015). Linguistic outcomes of English medium instruction programmes in higher education: A study on economics undergraduates at a Catalan university. *Higher Learning Research Communications*, 5(1), 47-67.

Ammon, U., & McConnell, G. D. (2002). *English as an academic language in Europe: A survey of its use in teaching*. Frankfurt am Maim: P. Lang.

Arnó-Macià, E., Aguilar-Pérez, M., & Tatzl, D. (2020). Engineering students' perceptions of the role of ESP courses in internationalized universities. *English for Specific Purposes*, 58, 58-74.

Baker, W., & Hüttner, J. (2017). English and more: A multisite study of roles and conceptualisations of language in English medium multilingual universities from Europe to Asia. *Journal of Multilingual and Multicultural Development*, 38(6), 501-516.

Başıbek, N., Dolmacı, M., Cengiz, B. C., Bür, B., Dilek, Y., & Kara, B. (2014). Lecturers' Perceptions of English Medium Instruction at Engineering Departments of Higher Education: A Study on Partial English Medium Instruction at Some State Universities in Turkey. *Procedia - Social and Behavioural Sciences*, 116, 1819–1825.

Beanland C, & Schneider Z., l. (1999). *Nursing research: methods, critical appraisal and utilisation*. Mosby, Sydney.

Bell, J. (1993). *Doing your research project : a guide for first-time researchers in education and social science*. (2nd ed.). Buckingham: Open University Press.

Benjamin, A. S., & Pashler, H. (2015). The Value of Standardized Testing: A Perspective From Cognitive Psychology. *Policy Insights from the Behavioral and Brain Sciences*, 2(1), 13–23.

Belhiah, H., & Elhami, M. (2015). English as a medium of instruction in the Gulf: When students and teachers speak. *Language Policy*, 14(1), 3-23.

Blaikie, N. W. H. (2000). *Designing social research : the logic of anticipation*. Cambridge: Polity Press.

Blunch, N. J. (2013). *Introduction to structural equation modeling using IBM SPSS Statistics and Amos.* (2nd ed.). London: Sage Publications.

Bradford, A. (2012). Adopting English-Taught Degree Programs. *International Higher Education*, (69), 8-10.

Bradford, A. (2016). Toward a Typology of Implementation Challenges Facing English-Medium Instruction in Higher Education: Evidence From Japan. *Journal of Studies in International Education*, 20(4), 339-356.

Bradford, A., & Brown, H. (2017). *English-Medium Instruction in Japanese Higher Education:Policy, Challenges and Outcomes* (Multilingual Matters). Blue Ridge Summit, PA.

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

Briggs Baffoe-Djan, J., & Smith, S, A. (2020). Descriptive statistics in data analysis. Observations and field-notes: Recording lived experiences. In J. McKinley, & H. Rose (Eds.), The Routledge Handbook of Research Methods in Applied Linguistics (pp. 336-347). Routledge. https://www.routledge.com/The-Routledge-Handbook-of-Research-Methods-in-Applied-Linguistics-1st/McKinley-Rose/p/book/9781138501140.

Brown, H. (2017). Why and why now? Understanding the rapid rise of English-medium instruction in higher education in Japan. *Journal of International Studies and Regional Development*, 8, 1-16.

Brown, H., & Bradford, A. (2017). EMI, CLIL, & CBI: Differing approaches and goals. In P. Clem-ents, A. Krause, & H. Brown (Eds.), *Transformation in language education*. Tokyo: JALT.

Brown, H., & Bradford, A. (2018). Teaching subject content through English: CLIL and EMI courses in the Japanese university. In P. Wadden & C. C. Hale (Eds.), Teaching English at Japanese universities (pp. 103–108). London: Routledge.

Brummet, Q., Masterton, M., and Smith, D. (2014). Evaluation of Commercial School and Teacher Lists to Enhance Survey Frames. CARRA Working Paper 2014-07. Washington, DC: U.S. Census Bureau. Available: https://www.census.gov/content/dam/Census/library/working-papers/2014/adrm/carra-wp-2014-07.pdf [July 2020].

Bryman, A. (2012). *Social research methods*. Oxford: Oxford University Press.

Byun, K., Chu, H., Kim, M., Park, I., Kim, S., & Jung, J. (2011). English-Medium Teaching in Korean Higher Education: Policy Debates and Reality. *Higher Education: The International Journal of Higher Education and Educational Planning*, 62(4), 431–449.

Cameron, L., Wise, S. L., & Lottridge, S. M. (2007). The development and validation of the Information Literacy Test. *College and Research Libraries*, 68(3), 229–236.

Carter, N., Bryant-Lukosius, D., Dicenso, A., Blythe, J., & Neville, A. (2014). The use of triangulation in qualitative research. Oncology Nursing Forum, 41(5), 545-7.

Cassels, J. R. T. & Johnstone, A. H. (1983). The meaning of words and the teaching of chemistry. *Education in Chemistry*, January: 10–11.

Chalmers, H. (2019). *The Role of the first language in English Medium Instruction*. OUP ELT Position Paper. Available at: https://elt.oup.com/feature/global/expert/emi?cc=gb&selLanguage=en

Charles, M., & Pecorari, D. (2016). Introducing English for Academic Purposes. Routledge.

Chan, S. (2015). Linguistic challenges in the mathematical register for EFL learners: linguistic and multimodal strategies to help learners tackle mathematics word problems, *International Journal of Bilingual Education and Bilingualism*, 18(3), 306-318.

Chang, J.-Y., Kim, W., & Lee, H. (2017). A Language Support Program for English-Medium Instruction Courses: Its Development and Evaluation in an EFL Setting. *International Journal of Bilingual Education and Bilingualism*, 20 (5), 510–528.

Chapple, J. (2015). Teaching in English is not necessarily the teaching of English. *International Education Studies*, 8(3)1–13.

Chin, J. (2002). Is There a Scholarship of Teaching and Learning in Teaching Sociology? A Look at Papers from 1984 to 1999. *Teaching Sociology.* 30(1):53–62.

Cho, D. W. (2012). English-medium instruction in the university context of Korea: Tradeoff between teaching outcomes and media-initiated university ranking. *The Journal of Asia TEFL*, 9(4), 135-163.

Cho, Y., & Bridgeman, B. (2012). Relationship of TOEFL iBT scores to academic performance: Some evidence from American universities. *Language Testing*, 29(3), 421–442.

Citro, C. (2014). From multiple modes for surveys to multiple data sources for estimates. *Survey Methodology*, 40(2), 137–161.

Civan, A., & Coşkun, A. (2016). The effect of the medium of instruction language on the academic success of university students. *Educational Sciences: Theory & Practice*, 16, 1981–2004.

Collins, C. S., Lee, M. N. N., Hawkins, J. N., & Newbauer, D. E. (2016). Palgrave Handbook of Asia pacific higher education. New York: Palgrave Macmillan.

Cook, T.D. & Campbell, D.T. (1979) *Quasi-experimentation: design and analysis issues for field settings.* (Boston, MA, Houghton Mifflin).

Cooper, M., & Stowe, R. (2018). Chemistry Education Research-From Personal Empiricism to Evidence, Theory, and Informed Practice. *Chemical Reviews*, 118(12), 6053-6087.

Costa, F. (2012). Focus on form in ICLHE lectures in Italy: Evidence from English-medium science lectures by native speakers of Italian. *AILA Review*, 25, 30–47.

Costa, F., & Coleman, J. A. (2012). A survey of English-medium instruction in Italian higher education. *International Journal of Bilingual Education and Bilingualism*, 16(1), 3-19.

Coxhead, A. (2000). A New Academic Word List. TESOL Quarterly, 34(2), 213–238.

Coxhead, A., & Boutorwick, T. J. (2018). Longitudinal vocabulary development in an EMI international school context: Learners and texts in EAL, Maths, and Science. *TESOL Quarterly*, 52(3), 588– 610.

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

Crandall, J. A., & Christison, M. (2016). *Teacher education and professional development in TESOL: global perspectives*. New York: Routledge.

Creswell, J. W., Fetters, M. D., & Ivankova, N. V. (2004). Designing a mixed methods study in primary care. *Annals of Family Medicine*, 2(1), 7.

Creswell, J. W., & Plano Clark, V. L. (2007*). Designing and conducting mixed methods research*. Thousand Oaks, CA: SAGE.

Creswell, J. W., Plano Clark, V. L., & Garrett, A. L. (2008). *Methodological issues in conducting mixed methods research designs*. In M. M. Berg- man (Ed.), Advances in mixed methods research (pp. 66–83). London: SAGE.

Curdt-Christiansen, X. L. (2020). Observations and field-notes: Recording lived experiences. In J. McKinley, & H. Rose (Eds.), The Routledge Handbook of Research Methods in Applied Linguistics (pp. 336-347). Routledge. https://www.routledge.com/The-Routledge-Handbook-of-Research-Methods-in-Applied-Linguistics-1st/McKinley-Rose/p/book/9781138501140

Curle, S., Jablonkai, R., Mittelmeier, J., Sahan, K., & Veitch, A. (2020). English Medium Part 1: Literature review. In N. Galloway (Ed.) English in higher education (Report No. 978-0-86355-977–8). British Council.

Curle, S., Jablonkai, R., Mittelmeier, J., Sahan, K., & Veitch, A. (2020). English Medium Part 1: Literature review. In N. Galloway (Ed.) English in higher education (Report No. 978-0-86355-977–8). British Council.

Curle, S, Yuksel, D, Soruç, A, & Altay, M. (2020). Predictors of English Medium Instruction academic success: English proficiency versus first language medium, *System*, 95.

Curry, N., & Pérez-Paredes, P. (2021). Understanding Lecturers' Practices and Processes: A Qualitative Investigation of English-Medium Education in a Spanish Multilingual University. In: Carrió-Pastor, M.L., & Bellés Fortuño, B. (eds.). *Teaching Language and Content in Multicultural and Multilingual Classrooms* (pp. 123-156). Palgrave MacMillan.

Dafouz, E., M. Camacho., & E. Urquia. (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.

Dafouz, E., & Camacho-Miñano, M. M. (2016). Exploring the impact of English-medium instruction on university student academic achievement: The case of accounting. *English for Specific Purposes*, 44, 57–67.

Dafouz, E., & Smit, U. (2016). Towards a Dynamic Conceptual Framework for English-Medium Education in Multilingual University Settings. *Applied Linguistics*, 37(3), 397-415.

Dearden, J. (2015). *English as a medium of instruction: A growing global phenomenon*. London, England: British Council.

Dearden, J., & Macaro, E. (2016). Higher Education Teachers' Attitudes towards English Medium Instruction: A three country comparison. *Studies in Second Language Learning and Teaching*, 6, 455-486.

Delucchi, Michael. (2014). Measuring Student Learning in Social Statistics: A Pretest-Posttest Study of Knowledge Gain. *Teaching Sociology*, 42(3), 231-239.

Din, Y. Y., Tsang, W. K., & Cheung, S. P. (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(2), 295-331.

Dippold, D., Heron, M., & Gravett, K. (2021). Language, literacies and learning in the disciplines: A higher education perspective. *Language Teaching*, 1-3.

Dörnyei, Z., & Csizér, K. (2012). *How to Design and Analyze Surveys* in A. Mackey, S. Gass (Ed.), Second Language Acquisition Research in Research methods in second language acquisition: a practical guide. (A. Mackey & S. M. Gass, Eds.). Oxford: Oxford: Wiley-Blackwell.

Dudley-Evans, T., & St. John, M. (1998). *Developments in ESP : A multi-disciplinary approach* (Cambridge language teaching library). Cambridge: Cambridge University Press.

Earl, L & Torrance, N. (2000). Embedding Accountability and Improvement Into Large-Scale Assessment: What Difference Does it Make? *Peabody Journal of Education*, Vol. 75 (4).

Egalite, A. J., & Kisida, B. (2018). The Effects of Teacher Match on Students' Academic Perceptions and Attitudes. *Educational Evaluation and Policy Analysis*, 40(1), 59–81.

Ekoç, A. (2020). English Medium Instruction (EMI) from the Perspectives of Students at a Technical University in Turkey. *Journal of Further and Higher Education*, 44(2), 231-243.

EL-Daou, B., & Abdallah, A. (2019). The impact of CLIL implementation on Lebanese students' attitudes and performance. *Global Journal of Foreign Language Teaching*. 9(1), 001–013.

Elliott, M.R., Raghunathan, T.E. and Schenker, N. (2020). Combining Estimates from Multiple Surveys. In Wiley StatsRef: Statistics Reference Online (eds N. Balakrishnan, T. Colton, B. Everitt, W. Piegorsch, F. Ruggeri and J.L. Teugels).

Eriksson, A. (2018). Redesigning a Discipline-Specific Writing Assignment to Improve Writing on an EMI Programme of Engineering. *Journal of Academic Writing*, 8(2), 48-66.

Evans, S. (2002). The medium of instruction in Hong Kong: policy and practice in the new English and Chinese streams. *Research Papers in Education*, 17(1), 97–120.

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.

Fenton-Smith, B., Humphreys, P., & Walkinshaw, I. (Eds.) (2017). *English medium instruction in higher education in Asia-Pacific: From policy to pedagogy*. New York, NY: Springer.

Field, A. P. (2018). *Discovering statistics using IBM SPSS Statistics*. (5th ed.). London: Sage Publications.

Fitzmaurice, G. M., Laird, N. M., & Ware, J. H. (2004). *Applied longitudinal analysis*. Hoboken, NJ: Wiley.

Fidan Uçar, H., & Soruç, A. (2018). Examining Turkish university students' sense of achievement, motivation, and anxiety: A comparison of the English- and French-medium education systems. *Eurasian Journal of Applied Linguistics*, 4(2), 177-191.

Flick, U. (2009). *An Introduction to Qualitative Research*. London: Sage Publications.

Floris, F. D. (2014) Learning subject matter through English as the medium of instruction: students' and teachers' perspectives, *Asian Englishes*, 16:1, 47-59.

Flowerdew, J. (2016). English for Specific Academic Purposes (ESAP) writing: making the case. *Writing and Pedagogy*. 8(1), 5-32.

Flowerdew, J., & Miller, L. (1992). Student Perceptions, Problems and Strategies in Second Language Lecture Comprehension. *RELC Journal*, 23, 60–80.

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. British Council ELT Research Papers (Vol. 17).

Galloway, N., & McKinley, J. (2019). Englishization of higher education. In Mohebbi, H. & Coombe, C. (Eds.), *Research Questions in Language Education and Applied Linguistics*. Berlin: Springer.

Galloway, N., & Rose, H. (2021). English Medium Instruction and the English Language Practitioner. *ELT Journal*. 75. 10.1093/elt/ccaa063.

Galloway, N., & Ruegg, R. (2020). The provision of student support on English Medium Instruction programmes in Japan and China. *Journal of English for Academic Purposes*, 45.

Gardner, P. L. (1972). Words in Science, Melbourne: Australian Science Education Project, 210.

Gass, S. (2018). Input, interaction, and the second language learner. Routledge.

Gerretson, H., & Golson, E. (2005). Synopsis of the use of course-embedded assessment in a medium sized public university's general education program. *Journal of General Education*, 54(2), 139–149.

Graham, J. G. (1987). English Language Proficiency and the Prediction of Academic Success. *TESOL Quarterly*, 21(3), 505-21.

Grant, A. M., & Wall, T, D. (2009). The neglected science and art of quasi-experimentation: Why-to, when-to, and how-to advice for organisational researchers. *Organisational Research Methods*, 12, 653-686.

Gratch-Lindauer, B. (2003), Selecting and developing assessment tools, in Avery, E.F. (Ed.), Assessing Student Learning Outcomes for Information Literacy Instruction in Academic Institutions, *Association of College and Research Libraries*, Chicago, IL, pp. 22-39.

Gravetter, F., & Wallnau, L. (2000). *Statistics for the behavioral sciences*. Wadsworth / Thomson Learning, Stamford.

Gray, D. (2018). *Doing research in the real world* (4th ed.). Los Angeles.

Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a Conceptual Framework for Mixed-Method Evaluation Designs. *Educational Evaluation and Policy Analysis*, 11(3), 255–274.

Guo, H., Tong, F., Wang, Z., Min, Y., & Tang, S. (2018). English- vs. Chinese-Medium Instruction in Chinese Higher Education- A Quasi-Experimental Comparison. Sustainability, 10(11), 4230.

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2017). *A primer on partial least squares structural equation modeling* (PLS-SEM) (2nd ed.). Thousand Oaks, CA: Sage.

Hernandez-Nanclares, N. & Jimenez-Munoz, A. (2017) English as a medium of instruction: evidence for language and content targets in bilingual education in economics, *International Journal of Bilingual Education and Bilingualism*, 20:7, 883-896.

He, J., & Chiang, S. (2016). Challenges to English-medium instruction (EMI) for international students in China: A learners' perspective: English-medium education aims to accommodate international students into Chinese universities, but how well is it working? *English Today*, 32(4), 63-67.

Hellekjær, G.O. (2010). Language Matters Assessing Lecture Comprehension in Norwegian English-Medium Higher Education, In C. Dalton-Puffer, T. Nikula & U. Smit (Eds), Language Use in Content And Language Integrated Learning (CLIL) Classrooms (pp. 233–258), Amsterdam: John Benjamins.

Henriksen, B., Holmen, A., & Kling, J. (2019). *English medium instruction in multilingual and multicultural universities: Academics' voices from the Northern European context*. Oxford: Routledge/Taylor & Francis.

Hoare, P., & Kong, S. (2008). Late immersion in Hong Kong: Still stressed or making progress? In T. W. Fortune & D. J. Tedick (Eds.), Pathways to multilingualism: Evolving perspectives on immersion education (pp. 242–263). Clevedon, England: Multilingual Matters.

Hu, G., & Duan, Y. (2019). Questioning and responding in the classroom: A cross-disciplinary study of the effects of instructional mediums in academic subjects at a Chinese university. *International Journal of Bilingual Education and Bilingualism*, 22(3), 1-19.

Hu, J., & Wu, P. (2020). Understanding English language learning in tertiary English-medium instruction contexts in China. *System*, 93.

Hua, T-L. (2019). Understanding the Learning Challenges of English-Medium Instruction Learners and Ways to Facilitate Their Learning: A Case Study of Taiwan Psychology Students' Perspectives. *Latin American Journal of Content & Language Integrated Learning*, 12(2), 321-340.

Huberman, A. M., & Miles, M. B. (2002). *The qualitative researcher's companion*. Thousand Oaks, [Calif.] : SAGE.

Hultgren, A. K. (2014). English language use at the internationalised universities of Northern Europe: Is there a correlation between Englishisation and world rank? *Multilingua: Journal of Cross-Cultural and Interlanguage Communication*, 33(3-4), 389-412.

Hu, G., Li, L., & Lei, J. (2014). English-medium instruction at a Chinese university: Rhetoric and reality. *Language Policy*, 13(1), 21-40.

Huang, D. (2015). Exploring and Assessing Effectiveness of English Medium Instruction Courses: The Students' Perspectives. *Procedia, Social and Behavioral Sciences*, 173(C), 71-78.

Huang, H., & Curle, S. (2021). Higher education medium of instruction and career prospects: an exploration of current and graduated Chinese students' perceptions. *Journal of Education and Work*.

Hultgren, A. K., Jensen, C., & Dimova, S. (2015) Introduction: English-medium instruction in European higher education: from the North to the South, in Dimova, S, Hultgren, AK & Jensen, C (eds) English-Medium Instruction in European Higher Education. De Gruyter Mouton. pp. 1–16.

Hyland, K. (2016). Academic publishing and the myth of linguistic injustice. *Journal of Second Language Writing*, 31(C), 58-69.

Hynninen, N. (2012). ICL at the micro level: L2 speakers taking on the role of language experts. *AILA Review*, 25, 13–29.

Jenkins, J. (2019). English medium instruction in higher education: The role of English as lingua franca. Second Handbook of English Language Teaching: pp. 91–108.

Jenkins, J., & Mauranen, A. (Eds.) (2019). *Linguistic Diversity on the EMI Campus: Insider accounts of the use of English and other languages in universities within Asia, Australasia and Europe.* Abingdon: Routledge.

Jiang, L., Zhang, L. J. & May, S. (2019). Implementing English medium instruction (EMI) in China: Teachers' practices and perceptions, and students' learning motivation and needs. *International Journal of Bilingual Education and Bilingualism*, 22(2), 107–119.

Joe, Y. J., & Lee, H. K. (2013). Does English-Medium Instruction Benefit Students in EFL Contexts? A Case Study of Medical Students in Korea. *Asia-Pacific Education Researcher*, 22(2), 201–207.

Kamasak, S., Sahan, K., & Rose, H. (Under review 2020). Academic language-related challenges at an English-medium university. *Journal of English for Academic Purposes*.

Kang, S., & Park, H. (2005). English as the medium of instruction in Korean engineering education. *Korean Journal of Applied Linguistics*, 21(1), 155–174.

Keenan-Takagi, K. (2000). Embedding assessment in choral teaching. *Music Educators Journal*, 86(4), 42–49.

Kırkgöz, Y. (2016). Students' Approaches to Learning in an English-Medium Higher Education. *The Journal of Language Learning and Teaching*, 3 (2), 30-39.

Kim, E. (2002). Voices of learners from English-medium lecture. *Korean Journal of Applied Linguistics*, 18(2), 145–163.

Kim, S. (2020). 'Nabewa nabedayo!': Foreign faculty members' perceived need to learn the students' L1, *International Journal of Bilingual Education and Bilingualism*.

Kim, E., Kweon, S., & Kim, J. (2017). Korean engineering students' perceptions of English-medium instruction (EMI) and L1 use in EMI classes. *Journal of Multilingual and Multicultural Development*, 38(2), 130-145.

Kim, E. G., & Yoon, Jeong-Ro. (2018). Korean Science and Engineering Students' Perceptions of English-Medium Instruction and Korean-Medium Instruction, *Journal of Language, Identity & Education*, 17(3), 182-197.

Klaassen, R. G., & Graaff, E. (2001). Facing innovation: Preparing lectures for English-medium instruction in a non-native context. *European Journal of Engineering Education*, 26(3), 281–289.

Kling, J. (2019). *TIRF language education in review: English as a medium of instruction*. Monterey, CA & Baltimore, MD: TIRF & Laureate International Universities.

Kline, R. B. (2016). *Principles and Practice of Structural Equation Modeling*. New York: The Guilford Press.

Knapp, T. (2016). Why Is the One-Group Pretest–Posttest Design Still Used? *Clinical Nursing Research*, 25(5), 467-472.

Knapp, T. R., & Schafer, W. D. (2009). From Gain Score t to ANCOVA F (and Vice Versa) - Practical Assessment*, Research & Evaluation* (Vol. 14).

Krausz, J., Schiff, A., Schiff, J., & Hise, J. V. (2005). The impact of TOEFL scores on placement and performance of international students in the initial graduate accounting class. *Accounting Education: An International Journal*, 4(1): 103-111.

Kuckartz, U., & McWhertor, A. (2014). *Qualitative text analysis : a guide to methods, practice & using software*. London: SAGE Publications Ltd.

Kung, F. W. (2013). The more the merrier? Bilingualism in an academic perspective: exploring the implementation of English-medium instruction in Taiwanese tertiary education. *Asian EFL Journal Research Articles*, 4, 8–36.

Kym, I., & Kym, M. H. (2014). Students' perceptions of EMI in higher education in Korea. *Journal of Asia TEFL*, 11(2), 35–61.

Lavrakas, P. J. (2008). *Encyclopedia of survey research methods* Thousand Oaks, CA: SAGE Publications Ltd doi: 10.4135/9781412963947

Lei, J., & Hu, G. (2014). Is English-medium instruction effective in improving Chinese undergraduate students' English competence? *IRAL - International Review of Applied Linguistics in Language Teaching*, 52(2): 99–126.

Lightbown, P. (2014). *Focus on content-based language teaching* (Oxford key concepts for the language classroom series). Oxford.

Lin, H-Y. (2019). Perceptions of the Englishization of higher education in Taiwan: implementation and implications, *International Journal of Bilingual Education and Bilingualism*.

Lin, L. H. F., & Morrison, B. (2010). The Impact of the Medium of Instruction in Hong Kong Secondary Schools on Tertiary Students' Vocabulary. *Journal of English for Academic Purposes*, 9(4): 255–266.

Lo, Y. Y., & Macaro, E. (2012). The medium of instruction and classroom interaction : evidence from Hong Kong secondary schools. *International Journal of Bilingual Education and Bilingualism*, 15(1), 29e52.

Lohr, S., & Raghunathan, T. (2017). Combining survey data with other sources. *Statistical Science*, 32(2), 293–312.

Macaro, E (2015) English medium instruction: Time to start asking some difficult questions. Modern English Teacher 24/2: pp. 4–7.

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

Macaro, E. (2020). Exploring the role of language in English medium instruction. *International Journal of Bilingual Education and Bilingualism: Special Issue on The Role of Languages in English-Medium Instruction (EMI) at University*, 23(3), 263-276.

Macaro, E., Curle, S., Pun, J., An, J., & Dearden, J. (2018). A systematic review of English medium instruction in higher education. *Language Teaching*, 51, 36–76.

Macaro, E., & Han, S. (2019). English Medium Instruction in China's higher education: Teachers' perspectives of competencies, certification and professional development. *Journal of Multilingual and Multicultural Development*, 41, 219-231.

MacGregor, L. (2016). *CLIL in Japan: University teachers' viewpoints*. In P. Clements, A. Krause, & H. Brown (Eds.), Focus on the learner (pp. 426-432). Tokyo: JALT.

Marais, P., & Jordaan, F. (2000). Are we taking symbolic language for granted? J. Chemistry Educcation, 77(10), 1355–1357.

Martell, K., & Calderon, T. (2005). *Assessment of student learning in business schools: Best practice each step of the way*. Vol. 1, No. 1. Tallahassee, FL: Association for Institutional Research.

Martirosyan, N. M., Hwang, E., & Wanjohi, R. (2015). Impact of English Proficiency on Academic Performance of International Students. *Journal of International Students*, 5(1): 60–71.

Marshall, S., Gilmour, M., & Lewis, D. (1991). Words that Matter in Science and Technology, *Research in Science & Technological Education*, 9(1), 5-16.

Marshall, C., & Rossman, G. B. (2006). *Designing qualitative research* (4th ed.). Thousand Oaks, CA: Sage.

Mayring, P. (2019). *Qualitative content analysis: Demarcation, varieties, developments. Forum Qualitative Sozialforschung*, 20(3), Forum Qualitative Sozialforschung, September 2019, Vol.20(3).

McKinley, J. (2018). *Making the EFL to ELF transition at a Global Traction University*. In Bradford, A. & Brown, H. (Eds.), English-Medium Instruction at

Universities in Japan: policy, challenges and outcomes (pp. 238-249). Bristol: Multilingual Matters.

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

Mery, Y., Newby, J., & Peng, K. (2011). Assessing the reliability and validity of locally developed information literacy test items. *Reference Services Review*, 39(1), 98–122.

MEXT. (2014). Selection for the FY2014 Top Global University Project. Retrieved from http://docplayer.net/16379288-Selection-for-the-fy-2014-top-global-university-project-we-hereby-announce- the-selection-of-universities-for-the-top- global-university-project.html.

MEXT. (2018). 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]. Retrieved from https://www.mext.go.jp/content/20201005-mxt_daigakuc03-000010276_1.pdf.

Mills, A., Durepos, G., & Wiebe, E. (2010). *Encyclopedia of Case Study Research*. Thousand Oaks: SAGE Publications.

Moore, P. J. (2017). Unwritten Rules: Code Choice in Task-Based Learner Discourse in an EMI Context in Japan. In B. Fenton-Smith, P. Humphreys, & I. Walkinshaw (Eds.), English Medium Instruction in Higher Education in Asia- Pacific: From Policy to Pedagogy (pp. 299–320). Cham: Springer International Publishing.

Murphy, V.A., Arndt, H., Briggs Baffoe-Djan, J., 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. Education Endowment Foundation.

Neuman, W. Lawrence. (1989). Which Students Learn the Most, and Why? A Replication and Extension of the Szafran Pretest Study. *Teaching Sociology*, 17(1), 19-27.

Norris, S. P., & Phillips, L. M. (2003). How literacy in its fundamental sense is central to scientific literacy. *Science Education*, 87(2), 224-240.

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

Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.

Perkins, D. V. & Saris, R. N. (2001). A 'Jigsaw Classroom' Technique for Undergraduate Statistics Courses. *Teaching of Psychology*. 28(2):111–13.

Pecorari, D. (2020) *English medium instruction: Disintegrating language and content*. In: Dimova S and Kling J (2020) Integrating Content and Language in Multilingual Universities. Cham: Springer, 15–36.

Pecorari, D., & Malmström, H. (2018). At the crossroads of TESOL and English medium instruction. *TESOL Quarterly*, 52(3), 497-515.

Pessoa, S., Miller, R., & Kaufer, D. (2014). Students' challenges and development in the transition to academic writing at an English-medium university in Qatar. IRAL - *International Review of Applied Linguistics in Language Teaching*, 52(2), 127-156.

Phillipson, R. (2009). Linguistic Imperialism Continued. London: Routledge.

Phuong, Y. H., & Nguyen, T.T. (2019). Students' perceptions towards the benefits and drawbacks of EMI classes. *English Language Teaching*, 12(5), 88-100.

Price, B., & Randall, C. (2008). Assessing Learning Outcomes in Quantitative Courses: Using Embedded Questions for Direct Assessment. *Journal of Education for Business*, 83(5), 288-294.

Pukelsheim, F. (2012). The Three Sigma Rule. *The American Statistician*, 48(2), 88-91.

Pun J, Jin X (2021) Student challenges and learning strategies at Hong Kong EMI universities. *PLoS ONE*, 16(5): e0251564.

Querol-Julián, M., & Camiciottoli, B. C. (2019). The impact of online technologies and English medium instruction on university lectures in an international learning context: A systematic review. *ESP Today: Journal of English for Specific Purposes in Tertiary Level*, 7(1), 2–23.

Radjai, L., & Hammond, C. (2021). Internationalization of the curriculum in Japanese higher education: Strategies, pedagogies and practices in 'international' EMI classrooms. *Journal of Comparative & International Higher Education*, 11(Winter), 172–175.

Rashed., N. F., Hashimoto. M., & Oyabu. K. (2018). Effects of English Medium Instruction on Students' Learning Outcomes in Science: A Case Study at a College of Technology. *Forum of Language Instructors*, (12), 5-14, 2018-03.

Renssen, R. H. and Nieuwenbroek, N. J. (1997). Aligning estimates for common variables in two or more sample surveys. J. Amer. Statist. Assoc. 92 368–374. MR1436124

Reynolds, K. D., & West. S, G. (1987). A multiplist strategy for strengthening non-equivalent control group designs. Eval Rev. 11:691–714.

Richards, J. C., & Pun, J. (2021). A Typology of English-Medium Instruction. *RELC Journal*.

Ritchie, J., & Lewis, J. (2003). *Qualitative research practice : A guide for social science students and researchers*. Thousand Oaks, Calif. ; London: Sage.

Roberts, G., & Binder, D. (2009). Analyses Based on Combining Similar Information from Multiple Surveys. JSM Proceedings, *Survey Research Methods Section*, (2007), 2138–2147.

Rogier, D. (2012). The effects of English-medium instruction on language proficiency of students enrolled in higher education in the UAE (Doctoral dissertation). University of Exeter, United Kingdom.

Rose, H. (2019). Dismantling the Ivory Tower in TESOL: A Renewed Call for Teaching-Informed Research. *TESOL Quarterly*, 53(3), 895–905.

Rose, H., Curle, S., Aizawa, I., & Thompson, G. (2019). What drives success in English medium taught courses? The interplay between language proficiency, academic skills, and motivation. *Studies in Higher Education*, 1–13.

Rose, H. & Galloway, N. (2019). Global Englishes for Language Teaching. Cambridge University Press.

Rose, H., & McKinley, J. (2018). Japan's English-medium instruction initiatives and the globalization of higher education. *Higher Education*, 75(1), 111–129.

Rose, H., McKinley, J., Xu, X., & Zhou, S. (2020). *Investigating policy and implementation of English medium instruction in higher education institutions in China.* British Council.

Ross, C., Young, S., & Sturts, J. (2012). Evaluating the Student Learning Outcomes Assessment Process in Undergraduate Parks and Recreation Academic

Programs. SCHOLE: *A Journal of Leisure Studies and Recreation Education*, 27(1), 25-38.

Ryan, GW. & Bernard, HR. (2000). *Data management and analysis methods*. In NK Denzin & YS Lincoln (Eds.) Handbook of qualitative research (2nd Edition). pp. 769-802.

Salkind, N. J. (2010). Encyclopedia of research design (Vols. 1-0). Thousand Oaks, CA: *SAGE Publications*, Inc. doi: 10.4135/9781412961288

Samoff, J. (1999). No Teacher Guide, No Textbooks, No Chairs. Contending with Crisis in African Education. In Robert Arnove & Carlos Torres (Eds.) , Comparative Education. The Dialect of the Global and the Local (pp.393-431). Lanham: Rowman & Littlefield.

Sánchez-Pérez, M. (2021). Predicting content proficiency through disciplinary-literacy variables in English-medium writing, *System*, 97.

Sandhu, A. (2015). Internationalization of University Education in Japan. The Science/AAAS Custom Publishing Office. The American Association for the Advancement of Science. Retrieved July 27, 2020, from http://www.sciencemag.org/sites/default/files/Global_Japanese_Univ_19mar15.pdf.

Schenker, N. and T.E. Raghunathan, (2007), Combining Information from Multiple Surveys to Enhance Estimation of Measures of Health, *Statistics in Medicine*, 26, 1802-1811.

Schleppegrell, M. J. (2007) The Linguistic Challenges of Mathematics Teaching and Learning: A Research Review, *Reading & Writing Quarterly*, 23(2), 139-159

Schmidt-Unterberger, B. (2018). The English-medium paradigm: A conceptualisation of English-medium teaching in higher education. International Journal Of Bilingual Education And Bilingualism, 21(5), 527-539.

Sert, N. (2008). The language of instruction dilemma in the Turkish context. *System* (Linköping), 36(2), 156-171.

Siegel, J. (2020). Comprehension in English Medium Instruction (EMI) Lectures: On the Impact of Lecturer L2 English Use. *Language Learning in Higher Education (Berlin, Germany)*, 10(1), 73-91.

Spyros, K. & Wei. L. (2012). Are there additional benefits from being in small classes for more than one year? *Educational Research and Evaluation*, 18(7), 671-685.

Stevens, J. P. (2002). Applied multivariate statistics for the social sciences. Mahwah, NJ: Lawrence Erlbaum.

Smit, U. (2010). CLIL in an English as a lingua franca (ELF) classroom: On explaining terms and expressions interactively. In C. Dalton-Puffer; T. Nikula & U. Smit (eds.). Language use and language learning in CLIL classrooms, Amsterdam, Philadelphia: John Benjamins Pub. Co., 259-277.

Snow, M., & Brinton, D. (1988). Content-Based Language Instruction: Investigating the Effectiveness of the Adjunct Model. *TESOL Quarterly*, 22(4), 553-574.

Song, Y. (2019). English language ideologies and students' perception of international English-Medium-Instruction (EMI) Master's programmes: A Chinese case study. *English Today*, 35(3), 22-28.

Song Y., & Carheden, S. (2014). Dual meaning vocabulary (DMV) words in learning chemistry, *Chemistry Education Research and Practice*, 15(2), 128–141.

Sullivan, G. M., & Artino, A. R., Jr (2013). Analyzing and interpreting data from likert-type scales. *Journal of graduate medical education*, 5(4), 541–542.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Allyn & Bacon.

Taguchi, N. (2014). English-medium education in the global society: An Introduction. In Taguchi, N. (ed.), English-medium education in the global society (pp. 157–181). Special issue with *International Review of Applied Linguistics*, 52(2).

Thompson, G., Aizawa, I., Curle, S., & Rose, H. (2019). Exploring the role of self-efficacy beliefs and learner success in English medium instruction. *International Journal of Bilingual Education and Bilingualism*, 2019, 1–14.

Thompson, G., & McKinley, J. (2018). Integration of content and language learning. In the *TESOL Encyclopedia of English Language Teaching* (1st ed.). Edited by John I. Liontas (Project Editor: Margo DelliCarpini; Volume Editor: Shahid Abrar-ul-Hassan). Hoboken, NJ: Wiley.

Thompson, G., Takezawa, N., & Rose., H. (2021). Investigating self- beliefs and success for English medium instruction learners studying finance, *Journal of Education for Business*.

Trang, M. N., & Trinant, K. (2017) Challenges Encountered by Thai University Students in Understanding Scientific Content Courses Taught by Native Speakers of English. *Korea TESOL Journal*, 13(1), 177-203.

Trenkic, D., & Warmington, M. (2019). Language and literacy skills of home and international university students: How different are they, and does it matter? Bilingualism: *Language and Cognition*, 22(2): 349-365.

Trochim, W. M., Donnelly, J. P., & Arora, K. (2015). *Research Methods: The essential knowledge base*. Boston, MA: Cengage Learning.

Tseng, P., Richards, K., & Pilcher, N. (2018). Constructing English-medium instruction indicators in the shipping courses of Taiwan's higher education. *Maritime Business Review*, 3(1), 20-35.

Turhan, B., & Kırkgöz, Y. (2018). Motivation of engineering students and lecturers toward English medium instruction in Turkey. *Journal of Language and Linguistic Studies*, 14(1), 261-277.

Unterberger, B., & Wilhelmer, N. (2011). English-medium education in economics and business studies: capturing the status quo at Austrian universities. *International Journal of Applied Linguistics*, 161, 90-110.

Ute, S., & Dafouz, E. (2012). Integrating content and language in higher education: An introduction to English-medium policies, conceptual issues and research practices across Europe. *AILA Review*, 25(1), 1-12.

Vavelyuk, O. (2015). To Integrate Successfully: Language and Subject Studies in ESP Teaching. *Procedia - Social and Behavioral Sciences*, 199, 44–49.

Wagenaar, T. C. (2002). Outcomes Assessment in Sociology: Prevalence and Impact. *Teaching Sociology*. 30(4):403–13.

Wanders, F., Van Der Veen, I., Dijkstra, A., & Maslowski, R. (2020). The influence of teacher-student and student-student relationships on societal involvement in Dutch primary and secondary schools. *Theory & Research in Social Education*, 48(1), 101-119.

Weber, R. P. (1990). *Basic content analysis*. Newbury Park, CA: Sage.

Weiler, B, Moyle, BD, Torland, M, Wolf, ID & Witsel, M. (2014). Lessons learned from pilot testing an experimental communication intervention: generation Y and park benefits, 2014 CAUTHE 2014: Tourism and Hospitality in the Contemporary World: trends, change and complexity 1214-1217.

West, S. G. (1985). *Beyond the laboratory experiment: experimental and quasi-experimental designs for interventions in naturalistic settings*, in P. Karoly (ed.) Measurement Strategies in Health Psychology. New York: John Wiley.

Wilson, M., & Sloane, K. (2000). From principles to practice: An embedded assessment system. *Applied Measurement in Education*, 13(2), 181–208.

Woodrow, L. (2006). Academic success of international postgraduate education students and the role of English proficiency. *University of Sydney Papers in TESOL*, 1, 51-70.

Xie, W., & Curle, S. (2020). Success in English Medium Instruction in China: Significant indicators and implications. *International Journal of Bilingual Education and Bilingualism*, 1–13.

Yamarik, S. (2007). Does Cooperative Learning Improve Student Learning Outcomes? *Journal of Economic Education*. 38 (Summer):259–77.

Yang, W. (2015). Content and language integrated learning next in Asia: evidence of learners' achievement in CLIL education from a Taiwan tertiary degree programme, *International Journal of Bilingual Education and Bilingualism*, 18(4), 361-382.

Yan, W., & Cheng, L. (2015). How Language Proficiency Contributes to Chinese Students' Academic Success in Korean Universities. *Language Testing in Asia*, 5(1), 1-19.

Yang, M., O'Sullivan, P. S., Irby, D. M., Chen, Z., Lin, C., & Lin, C. (2019). Challenges and adaptations in implementing an English-medium medical program: a case study in China. *BMC Medical Education*, 19(1), 15.

Yıldız, M., Soruç, A., & Griffiths, C. (2017). Challenges and needs of students in the EMI (English as a medium of instruction) classroom. *Konińskie Studia Językowe*, 5(4), 387-402.

Yin, R. K. (2013). *Case study research: Design and methods*. Sage publications.

Yip, D. Y., Tsang, W. K., & Cheung, S. P. (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.

Yore, L. D. (2008). Science literacy for all students: Language, culture, and knowledge about nature and naturally occurring events. *Educational Studies in Language and Literature*, 8(1), 5-21.

Yoshida, T. (2015). Foreign Language Education in Japan: Exploring Qualitative Approaches, Sachiko Horiguchi, Yuki Imoto, & Gregory S. Poole (Eds.) (2015). Rotterdam: Sense Publishers, 191 pages.

Yuksel, D., Soruç, A., Altay, M., & Curle, S. (2021). A longitudinal study at an English Medium Instruction university in Turkey: the interplay between English language improvement and academic success. *Applied Linguistics Review*.

West, R., Guven, A., & Parry, J. (2015). *The State of English in Higher Education in Turkey*. Manchester: British Council Global English.

Wei, R., Feng, J., & Ma, Q. (2017). College students' perspectives on English-medium instruction and their English learning motivational intensity. In J. Zhao, & L. Q. Dixon (Eds.), *English-medium instruction in Chinese universities* (pp. 45-58). New York: Routledge.

Wilkinson, D. (2015). Educational Reforms and Development in Japan: Language and Culture Education for Global Competitiveness. *International Journal of Higher Education Management*, 1(2).

Zaif, F., Karapınar, A., & Yangın Eksi, G. (2017). A comparative study on the effectiveness of English-medium and Turkish-medium accounting education: Gazi University case. *Journal of Education for Business*, 92(2), 73-80.

Zhang, Z. (2018) English-medium instruction policies in China: internationalisation of higher education, *Journal of Multilingual and Multicultural Development*, 39:6, 542-555.

Zhang, Y., & Wildemuth, B. M. (2009). Qualitative analysis of content. In B. Wildemuth (Ed.), Applications of social research methods to questions in information and library (pp. 308–319). Westport, CT: Libraries Unlimited.

Zhao, J. (Ed.), Dixon, L. (Ed.). (2017). English-Medium Instruction in Chinese Universities. London: Routledge.

Zhiguo, Z. (2018) English-medium instruction policies in China: internationalisation of higher education, *Journal of Multilingual and Multicultural Development*, 39:6, 542-555.