CHALMERS
UNIVERSITY OF TECHNOLOGY

English-medium instruction and impact on academic performance: a randomized control study

Downloaded from: https://research.chalmers.se, 2026-09-07 18:25 UTC

Citation for the original published paper (version of record):
Bälter, O., Kann, V., Mutimukwe, C. et al (2024). English-medium instruction and impact on academic performance: a randomized control study. Applied Linguistics Review, 15(6): 2373-2396. http://dx.doi.org/10.1515/applirev-2022-0093

N.B. When citing this work, cite the original published paper.

research.chalmers.se offers the possibility of retrieving research publications produced at Chalmers University of Technology. It covers all kind of research output: articles, dissertations, conference papers, reports etc. since 2004. research.chalmers.se is administrated and maintained by Chalmers Library

(article starts on next page)

Olle Bälter, Viggo Kann, Chantal Mutimukwe and Hans Malmström*

English-medium instruction and impact on academic performance: a randomized control study

https://doi.org/10.1515/applirev-2022-0093
Received July 22, 2022; accepted August 8, 2023; published online September 11, 2023

Abstract: Stakeholders and researchers in higher education have long debated the consequences of English-medium instruction (EMI); a key assumption of EMI is that students' academic learning through English should be at least as good as learning through their first language (usually the national language). This study addressed the following question: "What is the impact from English-medium instruction on students' academic performance in an online learning environment?" "Academic performance" was measured in two ways: number of correctly answered test questions and through-put/drop-out rate. The study adopted an experimental design involving a large group (n = 2,263) randomized control study in a programming course. Student participants were randomly allocated to an English-medium version of the course (the intervention group) or a Swedish-medium version of the course (the control group). The findings were that students enrolled on the English-medium version of the course answered statistically significantly fewer test questions correctly; the EMI students also dropped out from the course to a statistically significantly higher degree compared to students enrolled on the Swedish version of the course. The conclusion of this study is thus that EMI may, under certain circumstances, have negative consequences for students' academic performance.

Keywords: English-medium instruction; impact; learning; online; experiment; programming

***Corresponding author: Hans Malmström**, Department of Communication and Learning in Science, Chalmers University of Technology, Gothenburg, Sweden, E-mail: mahans@chalmers.se. https://orcid.org/0000-0003-2450-7650
Olle Bälter and Chantal Mutimukwe, Department of Media Technology and Interaction Design, Royal Institute of Technology, Stockholm, Sweden, E-mail: ob1@kth.se (O. Bälter), chantal.mutimukwe@dsv.su.se (C. Mutimukwe). https://orcid.org/0000-0001-5626-1187 (O. Bälter). https://orcid.org/0000-0002-5966-7649 (C. Mutimukwe)
Viggo Kann, Department of Computer Science, Royal Institute of Technology, Stockholm, Sweden, E-mail: viggo@kth.se. https://orcid.org/0000-0003-3199-8953

Open Access. © 2023 the author(s), published by De Gruyter. This work is licensed under the Creative Commons Attribution 4.0 International License.

1 Introduction

In many higher education contexts around the world, including in Sweden (the site of this study), the use of English-medium instruction (EMI) is now a very common phenomenon (Malmström and Pecorari 2022; Curle et al. 2020). Typically, EMI involves tertiary-level disciplinary learning contexts (e.g., in physics, political science, or musicology) "in which at least some participants have a first language other than English, but in which all are expected to use English for some instructional purposes, and in which English is not taught but is nonetheless expected to learned" (Pecorari and Malmström 2018: 511). EMI is often adopted under the assumption that (i) learning an academic subject through second language English should be at least as good as learning it through the first language, and (ii) EMI should bring added value to students, most notably by improving students' English proficiency, but also, e.g., by furthering their global and intercultural competence (Pecorari and Malmström 2018; Curle et al. 2020; Macaro 2018). However, despite its popularity, the broader implications of EMI are still poorly understood (Macaro et al. 2018). Specifically, the effects on learning from studying academic content in English are under-researched (Macaro 2018).

To date, researchers have explored multiple dimensions relating to the advantages and disadvantages of EMI and used a range of different research designs. It is unsurprising, therefore, that these different research approaches have produced different outcomes. Thus, some researchers report no negative impact on students' learning of academic content because of EMI whereas others do identify such negative consequences.

The absence of conclusive answers to questions fundamental to EMI has led to calls for increased rigor in EMI research. Speaking to this point in a recent paper, Rubio-Alcalá et al. (2019: 199) argued that "it cannot be stated that the majority of the evidence found is reliable from a purely scientific point of view. On the contrary, only a small percentage meets the technical requirements for evidence."

The present study seeks to remedy this perceived lack of valid and reliable research concerning learning in EMI by adopting a large group randomized control study (experimental) design. Set in the context of a preparatory and introductory computer programming online course in Sweden, the study addresses this research question:

> *What is the impact from English-medium instruction on students' academic performance in an online learning environment?*

This study will bring important and novel insights to the field of EMI research and add nuance to stakeholder discussions in internationalized higher education.

2 Literature review

The question whether EMI results in academic learning advantages or deficits has hitherto been addressed by researchers in different ways. Most often, researchers have used questionnaires and interviews to explore students' and teachers attitudes about learning through English, and what the consequences might be. A small number of investigations have also explored the potential impact on learning from a more objective perspective, e.g., by measuring positive or negative impact through various tests designs, or by using data from objective sources, e.g., students' grades. In this section, we briefly illustrate some of the divergent findings reported by previous research.

2.1 Students' attitudes about learning in EMI

Cho (2012) surveyed students taking EMI courses and Korean-medium courses (using questionnaires and interviews) and found that EMI resulted in lower participation and reduced student attention compared to classes taught in Korean; of course, if students avoid the learning environment or if they are unfocused when they are in class, this could impact learning outcomes. Reports from, for example, Iceland, Hong Kong, Germany, and Norway also indicate that EMI students face learning challenges (Evans and Morrison 2011; Hellekjaer 2010; Ingvarsdottir and Ambjörnsdottir 2015). Many students in these contexts claim that they struggle to understand course content/disciplinary lectures in English. Hellekjaer (2010: 11), writing about students' self-reported experiences from EMI lectures in Norway and Germany, noted some of the main problems for students in this regard, e.g., "difficulties distinguishing the meaning of words [and] unfamiliar vocabulary", challenges which could directly impact learning. Similarly, EMI students surveyed in both Italy and Oman emphasize that being taught in their first language, compared to EMI, enhances their learning (Ellili-Cherif and Alkhateeb 2015; Guarda 2021).

However, some students are more positive. For example, a proportion of the students in Guarda's (2021) study said they welcome the added challenge from EMI and that the slower and more cognitively demanding process required in EMI is beneficial to their learning. Furthermore – and interestingly – Guarda reported that many students did not care whether they understood all aspects of the EMI teaching – to them, the many perceived advantages resulting from EMI (e.g., improved English proficiency) outweighed any disadvantages, including comprehension difficulties (for other positive accounts in this regard, see e.g., Ackerley (2017) and Guarda 2018).

2.2 Teachers' attitudes about students' learning in EMI

Teachers' attitudes are divided concerning a potential impact on students' learning from EMI. However, many teachers agree that, at least for a substantial proportion of students, learning through L2 English can be negatively affected (cf. Briggs et al. 2018; Macaro 2018; Macaro et al. 2018) (however, attitudes vary depending on factors such as teachers' age, years of teaching, and discipline).

Teacher research in both Turkey (Başıbek et al. 2014) and Sweden (Airey 2011) report concerns about teachers' inability to use English with enough sophistication and precision; this could cause academic content to be treated superficially. Airey (2011) also noted how many teachers fear that they are not able to contextualize what they are saying to the extent necessary, nor are they able to introduce important digressions which add certain "flavor" to the learning experience. Teaching also tends to become slower when teachers teach in foreign or second language English, meaning that the same amount of content cannot be covered in class (Cho 2012; Thogersen and Airey 2011).

Across many EMI contexts, teachers have expressed concern that EMI teaching is teacher-centered rather than student-centered; reduced classroom interaction and use of monologic teaching instead of engaging, dialogic and active learning could potentially also have negative effects on students' learning (Lasagabaster 2022; Lee and Prinsloo 2018).

2.3 Objective measures of effects on students' learning in EMI

The few existing studies adopting objectively grounded research designs can be broadly divided into two categories: studies that found no negative effect or even a small positive effect on content understanding from EMI, and studies that report a negative effect on content understanding.

Neither Park (2007) nor Joe and Lee (2013) reported any negative effects when they studied two relatively small cohorts of students enrolled in EMI courses in introductory linguistics (Park, n = 51) and medicine (Joe and Lee, n = 61) in South Korea. On the contrary, based on two pre/post-test design studies, they found that the medium of instruction seemed to have no effect on the understanding of lectures.

Tatzl and Messnarz (2013) investigated the influence of English as the examination language on the solution of physics and science problems by 96 Austrian engineering students in four groups, from first year to senior level. Half of each test group were given a set of 12 physics problems described in German (the official language in Austria); the other half received the same set of problems described in English. The hypothesis that the use of English would act as an additional barrier to

the comprehension and solution of physics problems was disconfirmed, i.e., English did not appear to constitute an obstacle in this receptive and productive task.

Spanish EMI scholars (Dafouz and Camacho-Minano 2016; Dafouz et al. 2014) also did not find any statistically significant differences when they compared Spanish EMI-students with non-EMI students based on coursework and students' final GPA. After comparing the grades of the two sets of students, the findings indicated that both groups obtained very similar results.

Grades were used also in the study by Costa and Mariotti (2017); Italian conomics students' exam grades were analyzed to compare the performance of students taught in English and students taught in Italian by the same lecturer, lecturing in two different courses but using exactly the same exam: when the exam grades of the students taking the Italian-medium course were compared to the grades of the students taking the English-medium course, no statistically significant differences were found.

A similar design was adopted by Reus (2020), with a similar outcome. Reus studied the performance of students' taking two economics courses; both courses were offered in both Spanish and English. No difference in the performance of the students taught in Spanish and the students taught in English was found when Reus analyzed their GPA and test scores obtained during the course.

Other researchers have, however, presented contradictory findings. Vinke (1995) conducted an experiment with a group of second-year engineering students in the Netherlands; the group was split in half and one group was given a lecture in English and the other group was given the same lecture in Dutch, by the same teacher. According to Vinke, the lectures were "highly similar but not identical". When the two groups sat the 40-item (Dutch) True/False-test after the lecture, there was a statistically significant difference between the groups: students in the English-medium group performed worse on the test, indicating that EMI affected learning outcomes. Twenty-five years later, de Vos et al. (2020) conducted another study set in the Netherlands with the same conclusion. By comparing close to 600 psychology students studying towards the same degree, either in Dutch or in English, the researchers found that students studying in their first language outperformed the EMI students when their grades were analyzed.

Grade comparison was also used in EMI research in Spain and Turkey. Arco-Tirado et al. (2018) based their investigation of Spanish teacher training students on GPA/grades and used counterfactual impact evaluation. Their finding was that EMI-students have a higher likelihood of obtaining a lower grade compared to non-EMI students. Civan and Coskun (2016) conducted a study across nine different departments at the University of Istanbul, all of which offer the same degree program in both Turkish and English. When the end of term-grades for students on the Turkish programs were compared with the grades from students on the equivalent

English-medium program, the grades of the Turkish students were statistically significantly better.

Finally, in a study of French students of law and engineering, Roussel et al. (2017) gave students a reading task in French and another reading task in English; the reading was followed by a content test consisting of questions asked in French. The analysis of the test scores was unequivocal: content uptake was negatively impacted when the learning happened in a foreign language.

Clearly, the vastly different outcomes concerning content learning effects reported in the studies reviewed in this section are an indication that more research is required. To this end, the present study adopts an experimental design and asks what the impact from English-medium instruction on students' academic performance is in an online learning environment.

3 Context, method and data

In this section we present the study context, the experimental design of the study, the randomized allocation and subsequent classification of the participants. Finally, we describe the analytical procedures.

3.1 Study context

The learning context targeted by this study was an online introductory course in programming.[1] Since it was launched, the objective of the course has been to level the (knowledge) playing field in preparation for university programming courses, which typically attract students of vastly varying knowledge of programming. The curriculum covers, e.g., image manipulation to learn about programming with variables, conditionals, loops, the use of predefined library functions, the relationship between hardware and software, the difference, and transfer, between analog and digital, networking and security. Colleagues of ours (professors of computer science) working with the Open Learning Initiative at Stanford allowed us to use their newly developed *Principles of Computing* course, which we translated to Swedish (both the Swedish and English version of the course was, however, culturally adapted, e.g., by replacing typical American images with images representative of the Swedish context).

1 The course website (revised since this study was conducted) can be accessed through https://www.sommarprogrammering.se/ (available in Swedish only).

In 2020, the course was run as a self-paced eight-module Massive Open Online Course (MOOC), offered in an English (EMI) as well as a Swedish version (SMI). Save for the language of instruction, the EMI and SMI versions of the course were identical.

The course comprised instructions and questions with direct (automated) feedback, packaged into eight modules. Central to this methodology was the use of formative questions throughout the learning process, with constructive feedback pointing students in the right direction and reinforcing correct answers (see Figure 1 for an example question).

Each of the eight course modules concluded with a module test with summative questions (similar in appearance to the example in Figure 1, but without the option to request hints). The course was designed for self-study, with no planned-for in-person teacher-student interaction, but students had the option to contact teachers via e-mail if they experienced problems.

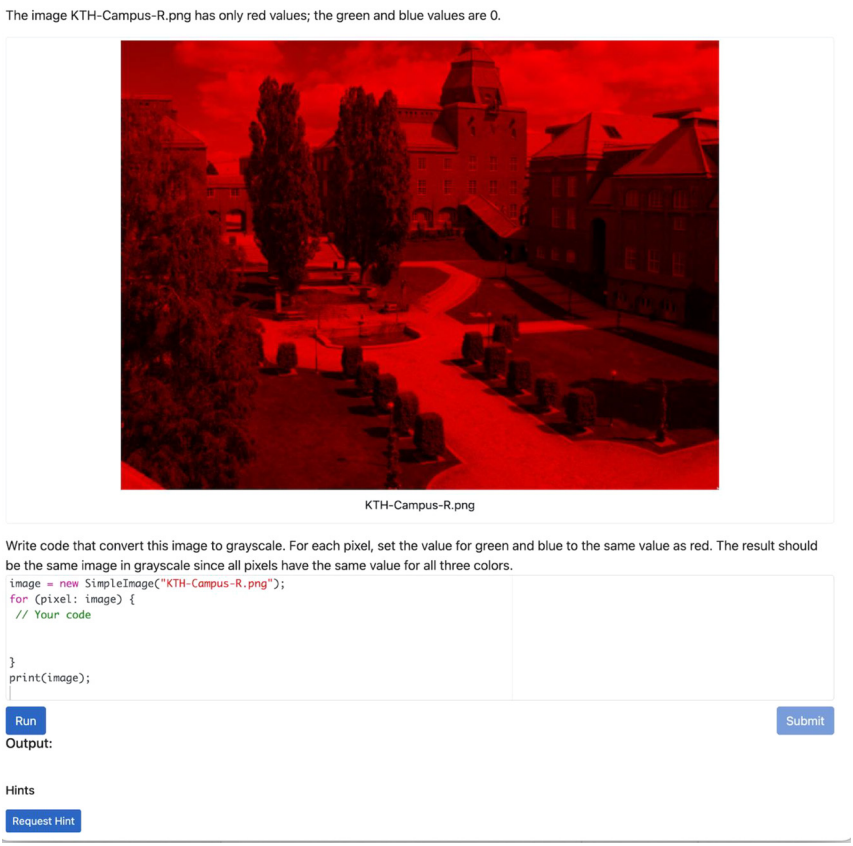

Figure 1: Example question from the course.

3.2 Study design

When EMI research is designed, attention should be paid to the methodological criticism directed against some of the EMI research to date (e.g., Macaro 2018). According to Rubio-Alcalá et al. (2019: 199), the field needs "to increase the quality of the research designs" [by adopting] "e.g., large-scale, randomized, longitudinal evaluations, experimental-control comparison, with evidence of no pre-test differences…". In our case, we adopted an experimental design involving a large parallel group randomized control study. Since the course was offered in an English (EMI) as well as a Swedish (SMI) version, and since what we wanted to investigate was the impact from using EMI, the two versions of the course were assigned 'intervention group' (EMI) and 'control group' (SMI) status respectively. For the avoidance of doubt, the EMI and the SMI versions of the course were identical, the Swedish version being a direct translation of the English original, making the study of the independent variable in this design – the medium of instruction – free from 'noise'.

3.3 Participants

A central distinguishing feature of a randomized control study is the random assignment of 'units', in this case students, to the intervention and control groups following certain inclusion and exclusion criteria. The students included in this study were identified during an enrollment process involving several steps. Prospective students (tens of thousands of students identified through the Swedish national university application system) were sent an invitation including a link to a webpage with general information about the course. The webpage had an application form where all prospective applicants were required to (i) self-assess their Swedish and English language proficiency using the CEFR self-assessment scale (Council of Europe 2020), and (ii) indicate their willingness to participate in the research.[2,3]

Three thousand three hundred and ninety-nine applicants (a self-selecting convenience sample) started the language self-assessment and 3,274 completed it. Of

2 While not without its detractors (see e.g., Hulstijn 2007), the CEFR framework has several benefits outweighing possible disadvantages in this study. The self-assessment grid has proven to be an intuitive and quick tool to use, and administering two validated language proficiency tests, one in English and one in Swedish, was not feasible in our context. We do, however, acknowledge the use of self-assessment as a potential limitation in the study design and a factor to be considered when the findings are interpreted.

3 Prospective students could sign up from June 11, 2020, until July 30, 2020.

these, 3,022 met the inclusion criteria set for the study: at least a B1-level for reading in Swedish as well as English; the B1-level for understanding written text was considered appropriate in view of the nature and design of the course and to enable students to engage properly with the course material.

At this stage of the enrollment process the applicants were randomly assigned to the EMI or the SMI version of the course (the enrollment cohort was split into two groups using a 50/50 allocation ratio). Use was made of a randomization feature built into the course management software; no allocation concealment mechanism was used. A total of 2,475 students completed the process of signing up for an account, resulting in a distribution of 1,265 students in the Swedish version of the course and 1,210 in the English version. As the students volunteered to participate in the study without knowing which version of the course they would eventually be assigned to, they were offered the possibility to swap language versions of the course. Two hundred and twelve students chose to do this, 81 assigned to the English version and 131 assigned to the Swedish version. After swapping, these students no longer qualified as randomly assigned and were therefore excluded from further analysis. This left us with a total of 2,263 participants included in the study, see Figure 2 and Table 1.

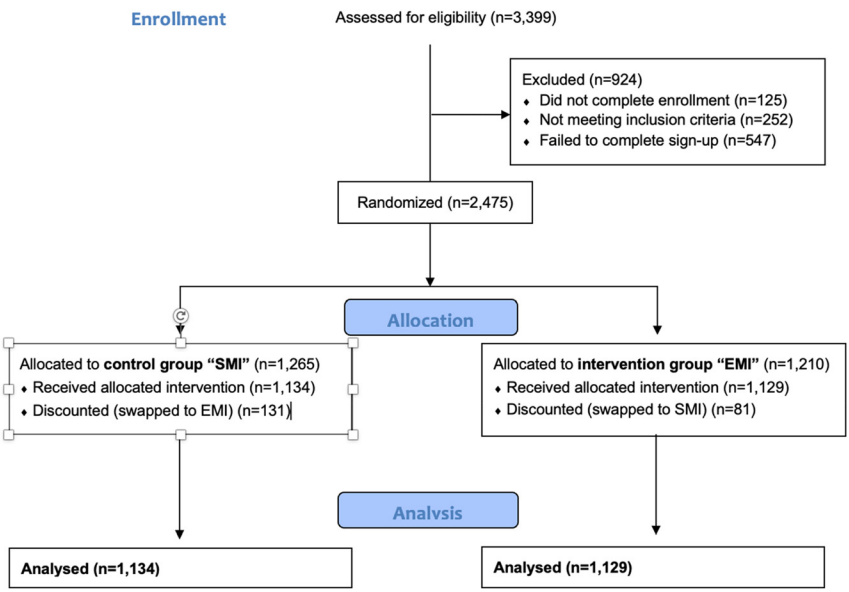

Figure 2: Participant flow diagram.

Table 1: Baseline demographic data for study participants in the Swedish and English version of the course.

		Swedish version	English version
Gender:			
	No answer	27 (2.4 %)	20 (1.8 %)
	Female	508 (44.8 %)	519 (46.0 %)
	Male	577 (50.9 %)	569 (50.4 %)
	"Prefer not to say"	22 (1.9 %)	21 (1.9 %)
	Total	**1,134 (100 %)**	**1,129 (100 %)**
Age:	<20	278 (24.5 %)	273 (24.2 %)
	21–30	556 (49.0 %)	644 (57.0 %)
	31–40	197 (17.4 %)	136 (12.1 %)
	41<	103 (9.1)	76 (6.7 %)
	Total	**1,134 (100 %)**	**1,129 (100 %)**
Years of schooling at time of enrollment:	(mean)	15.6	15.8

The overall drop-out rate for this course was substantial. The number of active participants (those who made an attempt to answer at least one of the module test questions) amounted to 815 students. The remainder (n = 1,448) were considered drop-out students (649 in the SMI course and 799 in the EMI course). Following the principle of 'intention to treat' (Hollis and Campbell 1999), which is fundamental to randomized control studies (by preserving the integrity of the randomization and avoiding a bias), both active students and drop-out students were included in the analysis.

Thus, for the purposes of the analysis in this study (see below), the participants for each version of the course were categorized into two main groups: (i) ALL STUDENTS (n = 2,263, of which 1,134 were in the SMI version, and 1,129 were in the EMI version: this group thus included all active students as well as all drop-out students; (ii) ACTIVE STUDENTS (n = 815, of which 485 were in the SMI version, and 330 were in the EMI version): this group only included the students who attempted to answer at least one module test question.

Each of the two main groups – ALL STUDENTS as well as ACTIVE STUDENTS – were then further sub-categorized as follows for analytical purposes: students who had self-assessed their language abilities as being (a) equally proficient in Swedish and English; (b) equally proficient or more proficient in Swedish compared to English (this would be the 'typical' first-cycle Swedish student on the course); (c) more proficient in English. Note that category (b) is an extension of category (a), i.e., all participants in (a) are also present in (b). This categorization is illustrated in Figure 3.

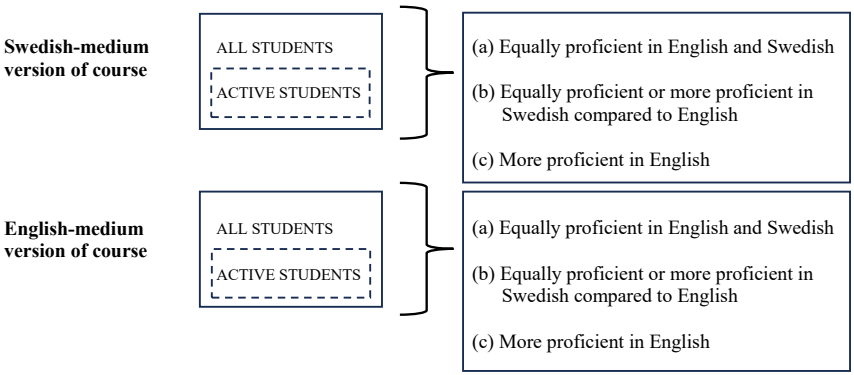

Figure 3: Categorization of students in the SMI and EMI versions of the course (ALL STUDENTS VS. ACTIVE STUDENTS) based on students' self-rated language proficiency.

3.4 Analytical procedures

'Academic performance' (cf. research question above) was analyzed according to two different measures. One measure of academic performance was students' *academic knowledge* (this included, to varying degrees, content knowledge, procedural knowledge, and conditional knowledge of programming) as indicated by correct versus incorrect answers on the summative test questions in the modules. The number of questions in each module test was between two and eight. The total number of summative questions was 42. We used the number of correctly answered questions, an integer between 0 and 42 for each student, as our first indication of academic performance. All scoring of the test questions was automated, the assessment of outcomes was thus blinded to group allocation.

A second kind of academic performance was *retention/drop-out*, i.e., the proportion of students who were active versus those dropping out (for whatever reason). To this end, we recorded the number of students who did/did not make an attempt to answer at least one of the module test questions.

For all statistical calculations we used STATA 17 (https://www.stata.com/features/documentation/).

We used descriptive analysis to calculate the means, standard deviation, and frequency distributions for both academic knowledge and retention. We used the Wilcoxon rank-sum (Mann–Whitney) test to compare differences in academic knowledge between groups in the study (within and across the SMI and EMI versions of the course). The Wilcoxon rank-sum test is usually used to compare outcomes between two independent groups of non-parametric data (Gad Consulting Services

2018; Rosner et al. 2003). It is generally used when there is reason to believe that the data are not normally distributed, when the sample sizes are small, or when the variances are heterogeneous (Leon 1998). We chose the Wilcoxon rank-sum test as we could not confirm the distributional assumptions for the different groups. For the rank-sum test, analysis is not conducted on actual raw data (numbers), but on the rank-transformed data. The data in two groups being compared are initially sorted in ascending order (Gad Consulting Services 2018). Each number in the two groups must receive a rank value. Beginning with the smallest number in either group, which would receive a rank of 1, each number is assigned a rank. If there are duplicate numbers, then each value of equal size will receive the median rank for the entire identically sized group. If the lowest number appears twice, both numbers receive the rank of 1.5. This means that the ranks of 1 and 2 have been used and that the next highest number has a rank of 3. This process continues until all the numbers are ranked. Then the sum of the ranks of each of the two groups is computed and compared to the expected sum of ranks of a random group of the same size. For this study we report both the rank sum and expected sum values.

Because the test score data was ordinal and not normally distributed, we used Cliff's delta (Cliff 1993) for estimating the effect size of the difference between the two versions of the course.

Finally, we used chi-square testing (Franke et al. 2012) to compare the relationship with respect to retention/dropout in the EMI and SMI versions of the course. In this case, effect size was estimated using Phi φ.

4 Findings

The findings of this study can be summarized thus: Students in the SMI version of the course appeared to outperform students in the EMI version of the course in terms of academic knowledge (as indicated by the mean test scores). Similarly, the student retention rate in the SMI course was higher than in the EMI course. When self-reported (CEFR) language skills were considered, notable differences between the SMI and the EMI versions of the course were observed with respect to academic knowledge and retention/drop-out; such differences were sometimes expected and sometimes unexpected.

4.1 Students' academic knowledge (test mean scores)

The academic knowledge of the group of ALL STUDENTS differed when the SMI and EMI versions of the course were compared. The mean test score in the SMI version of the

Table 2: Mean score and rank-sum test results for ALL STUDENTS (maximum score on the test was 42).

Course	Mean	Standard deviation	Rank sum	Expected	# students
SMI	7.24[a]	13.2	1,378,497[a]	1,283,688	1,134
EMI	4.18[a]	10.2	1,183,219[a]	1,278,028	1,129
Both	5.71	11.9	2,561,716	2,561,716	2,263

[a]Differences are statistically significant at 5 % according to the Wilcoxon rank-sum test ($p < 0.0001$). Cliff's delta is 0.15.

course was 7.24 compared to 4.18 in the EMI course, see Table 2, indicating that students in the Swedish course answered more questions correctly than students in the English course. On average, students in the Swedish course answered 73 % more questions correctly. Additionally, the students assigned to the Swedish course had a larger rank sum (of the scores) than expected. The Wilcoxon rank-sum (Mann–Whitney) test showed that the difference between SMI and EMI in this regard is statistically significant. The effect size (Cliff's delta) was 0.15.

Looking only at the ACTIVE STUDENTS (the students who attempted to answer at least one module question), the mean scores in the SMI and EMI versions of the course again differed; as can be seen in Table 3, the mean score of the SMI course was 16.9 compared to 14.3 in the EMI course and, according to the Wilcoxon rank-sum test, the difference was statistically significant. On average, the SMI students answered 18 % more questions correctly. The effect size (Cliff's delta) was 0.085.

Histograms over the score distribution in the courses are presented in Figure 4 (the SMI course) and Figure 5 (the EMI course). The histograms are similar with spikes at 2, 9 and 42 points. The first two spikes come from students finishing only the first test (2 points) or both the first and the second test (7 more points), with all answers correct. The last spike is for the maximum score (42), i.e., finishing all the test questions correctly.

Differences (in academic knowledge) between the SMI and EMI versions of the course were also observed when consideration was given to students' self-proclaimed level of (Swedish and English) language proficiency, sometimes in unexpected ways, see Table 4. Students in both the SMI and the EMI versions of the course had rated themselves using the CEFR-proficiency matrix, and we subsequently grouped them

Table 3: Mean score and rank-sum test results for ACTIVE STUDENTS (maximum score on the test was 42).

Course	Mean	Standard deviation	Rank sum	Expected	# students
SMI	16.9[a]	15.6	204,718[a]	197,880	485
EMI	14.3[a]	14.5	127,802[a]	134,640	330
Both	15.9	15.2	332,520	332,520	815

[a]Differences are statistically significant at 5 % according to the Wilcoxon rank-sum test ($p = 0.034$). Cliff's delta is 0.085.

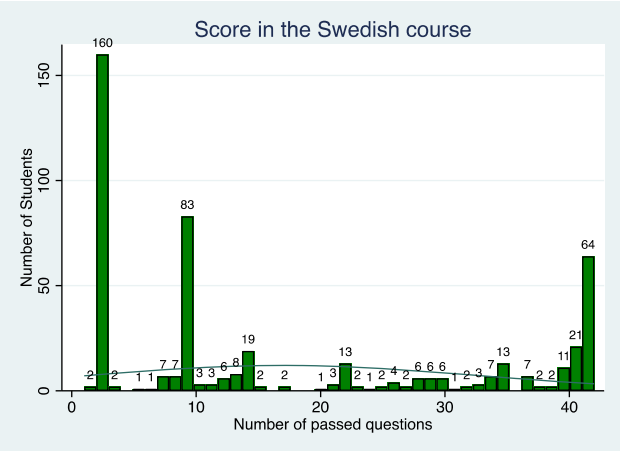

Figure 4: Histogram over scores in the SMI course.

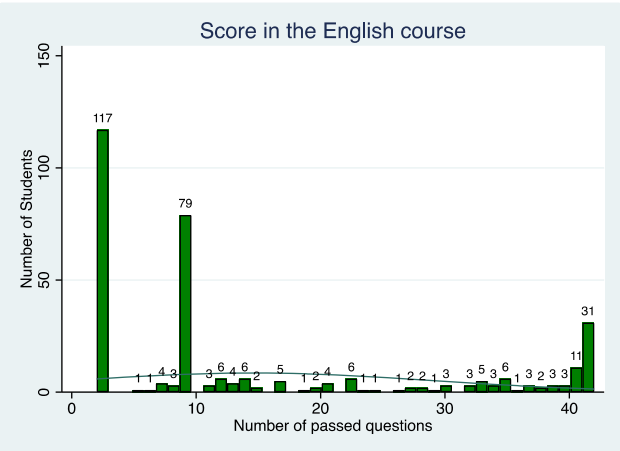

Figure 5: Histogram over scores in the EMI course.

as (a) equally proficient in Swedish and English; (b) equally proficient or better in Swedish compared to English; or (c) more proficient in English compared to Swedish.

Assuming that Swedish or English can be seen as a potential obstacle to students' learning the course content, students who claim to be as good in Swedish as in English *should* score equally well on the tests, regardless of which version of the course they were assigned to; they did not (the difference was statistically significant when ALL STUDENTS were included in the calculation, but not significant in the group of

Table 4: Mean score relative to self-reported language proficiency level (maximum score on the test was 42).

	SMI (mean scores of ALL STUDENTS/ACTIVE STUDENTS)	EMI (mean scores of ALL STUDENTS/ACTIVE STUDENTS)	
(a) Equally proficient in English and Swedish	6.2[a]/16.6	4.4[a]/15.0	Unexpected
(b) Equally proficient or more proficient in Swedish compared to English	7.3[b]/16.8	4.3[b]/14.5	Expected
(c) More proficient in English	6.6/19.5	2.8/11.3	Unexpected

[a,b]Differences are statistically significant at 5 % according to the Wilcoxon rank-sum test ([a]$p = 0.005$, [b]$p < 0.00001$).

ACTIVE STUDENTS). The SMI students in this proficiency category answered 41 % more questions correctly compared to the EMI students. Students who claimed to be equally good or better in Swedish compared to English *should* score better when tested in the Swedish version of the course; they did – students in this category taking the SMI version of the course answered 72 % more questions correctly (again, the difference was statistically significant when ALL STUDENTS were included, but not in the group of ACTIVE STUDENTS). Finally, students who claim to be better in English compared to Swedish *should* score better in the English version of the course; this could not be established since the difference between the SMI and EMI versions of the course was not statistically significant. However, looking purely at the descriptive statistics, it would appear that being more proficient in English is no guarantee for learning more academic content when studying in English; the opposite could be the case. Taken together, these findings relating to language proficiency and academic performance in EMI/SMI highlight something potentially interesting: (self-assessed) language proficiency may not necessarily in by itself be a critical factor for academic performance.

4.2 Retention/dropout in the SMI and EMI versions of the course

Six hundred and forty nine out of 1,134 (57 %) students dropped out of the Swedish-medium course, and 799 out of 1,129 (71 %) dropped out of the English-medium course. A chi-square test showed that the difference in drop-out was statistically significant, Chi2 (1) = 45; $p < 0.00001$, thus indicating that students in the English course were 25 % more likely to drop out. The effect size (φ) was 0.2.

Table 5: Retention rate relative to self-reported language proficiency level.

	SMI (retention/ drop-out)	EMI (retention/ drop-out)	Notes
(a) Equally proficient in English and Swedish	38 %/62 %[a]	30 %/70 %[a]	Unexpected
(b) Equally proficient or more proficient in Swedish compared to English	44 %/56 %[b]	30 %/70 %[b]	Expected
(c) More proficient in English	34 %/66 %	25 %/75 %	Unexpected

[a,b]Differences are statistically significant at 5 % according to Chi2-tests ([a]$p < 0.0004$, [b]$p < 0.000001$).

Some interesting differences relating to retention/drop-out were also observed when consideration was given to students' self-proclaimed level of (Swedish and English) language proficiency, see Table 5.

Assuming that language (Swedish *or* English) can be seen as a potential obstacle to students' remaining in the course and engaging with the course content, students who claim to be as good in Swedish as in English *should* remain/drop out to the same degree, regardless of which version of the course they are in; they did not, instead drop-out was more pronounced for this proficiency group in the EMI version of the course. Students who claim to be equally good or better in Swedish compared to English *should* drop out to a higher degree in the EMI course; they did. Finally, students who claim to be better in English compared to Swedish *should* remain to a higher degree in the EMI course; they did not, instead drop-out was more pronounced for this proficiency group in the EMI version of the course (but the difference between SMI and EMI was non statistically significant in this case).

The same trend – a seemingly higher tendency to drop out from the EMI course compared to the SMI course – was observed for *every CEFR-category* when a one-to-one comparison between the SMI and EMI course was made: whether students self-reported a proficiency level of B1-C2 in Swedish, or B1-C2 in English, made no difference: the drop-out rate was constantly higher in the EMI course (and more often than not the difference was statistically significant).

5 Discussion

This research explored the impact from EMI on students' academic performance in a self-paced programming MOOC. The study was set up as a randomized control study involving >2,000 students taking either a Swedish-medium (SMI) or English-medium (EMI) version of the course, and academic performance was measured along two dimensions: academic knowledge (assessed by test scores) and student retention. Based on the findings, and despite effect sizes that are, by certain standards, limited,

there is a case for saying that EMI could have a negative impact on students' academic performance; the EMI students (intervention cohort) in this study performed worse when tested on their academic knowledge, and they dropped out of the course to a higher degree than students in the SMI course (control group).

5.1 Students ostensibly learn less in EMI

The finding that SMI students answered statistically significantly more questions correctly when they were assessed in the course lends support to earlier research (using similar designs) on the implications of EMI for students' academic performance (cf. Roussel et al. 2017; Vinke 1995). This begs the question why studying in (L2) English in contexts like ours (self-paced learning in an online environment) may result in a learning deficit.

It has been argued by earlier EMI research that studying academic content in English may require more cognitive effort from the students, such that the simultaneous processing of content and a second/foreign language could lead to cognitive overloading (e.g., Guarda 2021; Roussel et al. 2017). To this end, the cognitive effort would be greater the less linguistically proficient the student is (in English). It is possible that at least some of the students in our study suffered from too limited English proficiency to be able to learn successfully in the EMI environment – it should be noted that no internationally recognized benchmark exists for what counts as sufficient English proficiency for studying through EMI. We know from earlier research that, for example, understanding the meaning of words in English can present a considerable challenge to students in EMI (e.g., Evans and Morrison 2011; Hellekjaer 2010). Our use of the CEFR-instrument did not allow us to tap into details of the students' language proficiency, but it is certainly possible that some of the vocabulary encountered in the course, e.g., academic vocabulary, was beyond the students' capacity. This could have set the students up for a challenging task when working through the modules of the course and attempting to answer the test questions. In this regard, the absence of a teacher in the self-study environment provided by the online course also meant that the students had zero access to any kind of 'language support' which could be supplied by a teacher (support which is, in theory, readily available in many EMI classrooms involving both students and a teacher). Despite EMI teachers' reported unwillingness to act as 'language teachers' (Airey 2012), there is ample evidence that content teachers can provide both explicit and incidental language support during EMI classes (Lasagabaster and Doiz 2021), thereby furthering students' understanding of subject content and their development of disciplinary literacy (Malmström and Pecorari 2021). To this end, it is thus possible that the students in the EMI version of our course suffered since "learning in

a foreign language without any language instructional support provides no advantage to content learning" (Roussel et al. 2017: 77). Future research replicating the current study could introduce elements of language support as a variable in the research design to investigate this further.

However, our findings indicate that insufficient language proficiency cannot be the sole explanation for why students in the SMI version of the course outperformed their colleagues in the EMI version: even students who claimed to be as proficient in English as in Swedish answered fewer questions correctly when tested in the EMI version of the course (and, at face value at least, EMI students who claimed to be more proficient in English than in Swedish also performed worse than their colleagues in the SMI version of the course). The finding that students' level of English proficiency did not seem to be related to academic performance runs counter to recently published research on EMI in Turkey. In their (four-year) longitudinal study of business and engineering students studying in an EMI environment, Yuksel et al. (2023) reported a correlation between (modestly) increasing English proficiency and (modestly) increasing academic achievement (measured by test scores and GPA).

5.2 Attrition greater in EMI

We adopted the basic assumption that dropping out from the course leads to reduced learning (as a direct consequence of 'non-engagement' with course content). Little research on EMI has studied the student retention/drop-out phenomenon in a systematic way. However, many EMI scholars, as well as stakeholders in higher education, have commented that attrition seems to happen to a high degree in EMI contexts (e.g., de Vos et al. 2020; Galloway et al. 2017; Kojima 2021; Staub 2022), thus lending support to our finding a particularly high level of attrition in the EMI version of the programming course.

Several studies concerned with EMI have suggested that drop-out rates in EMI are attributable to the added burden of studying through a second/foreign language, for example Lueg and Lueg's (2015) investigation of EMI in a Danish business school, where the authors claimed that "for EMI, this implies that the expectation of barriers (low English proficiency, inferior performance) leads to strategies of 'self-elimination' or 'self-exclusion'" (2015: 12). On this assumption, students who struggle with English ought to be more prone to drop out. Interestingly, our findings do not support this assumption; in the group of students who claimed to be equally proficient in Swedish and English, attrition was still much higher in the EMI version of the course. Similarly, when the drop-out rate was investigated from the point of view of CEFR-ratings (rather than with reference to our purposive categorization), attrition

was *always* higher in the EMI course, whether students self-rated as English B1 or English C2 (or anywhere in between).

On the back of our findings with respect to retention/drop-out, we suggest that it is important (i) to closely monitor attrition in EMI (cf. de Vos et al. 2020), particularly in those rare cases like ours when it is possible to make a direct comparison with L1 education and, when doing this, (ii) to explore not only language proficiency but also other factors potentially contributing to attrition in EMI (cf., Sinha et al. 2018 who reference different kinds of "push-", "pull-" and "fallout factors" responsible for students' dropping out of EMI). In this regard, the fact that the course investigated here was a self-paced MOOC could have greatly affected the overall propensity to drop out, since drop-out rates from MOOCs tend to be high in general (Jordan 2015). Our data do not allow speculation about the relative importance of different factors contributing to attrition; thus, we cannot say whether, e.g., the higher attrition in the EMI course was exacerbated because of factors relating to the course design (online, limited interaction etc.). However, there is no obvious reason to believe that the EMI course should disincentivize students in this regard more so than the SMI course (the course design was identical – the only differentiating factor was the medium of instruction).

5.3 Methodological reflections and limitations

This study heeded calls from the community of EMI scholars for increased quality in the field's research designs (e.g., Macaro 2018; Rubio-Alcalá et al. 2019), particularly those arguing for the need to strike a balance between conducting research under 'natural' conditions and research validity (Roussel et al. 2017: 78):

> Our view is that there has been far too little emphasis on studies that strictly control variables and the omission of such studies can run the risk of unbalanced conclusions and inappropriate recommendations. An appropriate balance between ecologically valid and experimentally valid studies is required.

To the best of our knowledge, only three earlier studies of EMI have adopted randomized allocation of students and used control groups as part of their research design to ensure comparison groups are not systematically different at baseline: Roussel et al. (2017), Tatzl and Messnarz (2013), and Vinke (1995), but only two (Roussel et al. and Vinke) involved actual elements of 'instruction'. Both these studies recall the tentative conclusion from the present study: education based on EMI could have negative consequences for academic learning. The current design (just like the designs adopted by Roussel et al. and Vinke) enabled us to control for interfering

variables that plague much research in the field of learning, e.g., student background factors, motivation, learning strategies, and emotions towards the education.[4]

Despite observing statistically significant differences between the SMI group and the EMI group (meaning that the differences are real and not attributable to chance), the effect sizes recorded were limited, indicating that the *actual* difference between the groups (for practical purposes) may be limited. However, it is well-known that small effect sizes are common in educational research, particularly in studies of student achievement, and both design and sample size are important factors to consider. Cheung and Slavin (2016) conducted a meta-study of effect sizes in education and found that randomized control studies have much lower effect sizes than other types of designs: the mean effect size for the 196 randomized control studies included in their analysis was 0.16. In another meta-analysis of educational interventions, Slavin and Smith (2009) reported a median effect size for studies with >2,000 participants of 0.07. It is difficult to speculate about the practical significance of the differences reported in here; no studies of a similar scope and design have been conducted in EMI, and Bakker et al. (2019: 6) remind us that effect sizes should only ever be interpreted in relation to effect sizes from "comparable studies with similar characteristics (research design, sample size, type of measurement, type of variable influenced, etc.)." Until such comparable research is presented, our findings should be interpreted with a level of caution.

The present study has several limitations. First, it must be acknowledged that this is a short-term study, providing a snapshot of EMI as realized in a single course; future work is required to establish whether the results reported here are stable over time and extend to academic programs (though, admittedly, this would involve a number of practical challenges relating to the experimental design). In this regard, earlier (non-experimental) research in EMI has produced mixed results; some studies have reported that learning challenges associated with EMI subside over time (e.g., Evans and Morrison 2011) whereas other have indicated that they persist (e.g., Civan and Coskin 2016).

A second limitation is that this study covered a single subject/discipline—programming/computer science (broadly speaking)—and earlier research has indicated that there are considerable disciplinary differences associated with challenges and development in EMI (cf. Dafouz et al. 2014). The design of the present study lends itself to a replication in the same kind of online learning context, but in another discipline; such a replication would add to the external validity of the research findings.

4 By highlighting the advantages of this research design, we do not wish to disparage EMI research adopting other kinds of designs, nor do we suggest that the conclusions from such research are flawed; they just offer a different lens through which EMI is observed.

Third, the presence/absence of a teacher in the learning environment is both an advantage and a limitation when the impact of EMI is studied. All other things being equal, teachers can support the learning experience, scaffold students' engagement with the course content, answer questions or provide other forms of guidance. However, in the context of an experiment, the presence of a teacher could be viewed as a disadvantage since the teacher (their pedagogical approach, language proficiency, or level of disciplinary knowledge) could be seen to introduce confounding variables. When there is no teacher, none of this becomes a factor in the experiment. At the same time, since most teaching, including digital education formats like ours, does include a teacher, the absence of a teacher reduces the ecological validity of the study as far as teacher-led education is concerned.

6 Conclusions

This research indicates that the medium of instruction can have consequences for students' academic performance in EMI under certain circumstances. The finding that students subjected to English-medium instruction answer significantly fewer test questions correctly and drop out from the education to a much higher degree compared to the students accessing the education in the national language should give stakeholders pause and inform the continued discussion concerning the (i) advantages and disadvantages of adopting EMI, and (ii) what (pedagogical or linguistic) support might be needed to scaffold students' learning experience in EMI. The limited effect sizes observed when the academic performance of students in the EMI version of the course was compared with the SMI version are acknowledged; this could be an indication that the differences, even if statistically significant, deserve to be interpreted with caution. Importantly, the outcome of a single study should not result in *ad-hoc* changes to language or education policies in higher education, whether at local or national levels, and the results from the present study need to be corroborated by future research adopting similar rigorous designs.

References

Ackerley, Katherine. 2017. What the students can teach us about EMI and language issues. In Katherine Ackerley, Marta Guarda & Francesca Helm (eds.), *Sharing perspectives on English-medium instruction*, 257–284. Bern: Peter Lang.

Airey, John. 2011. Talking about teaching in English: Swedish university lecturers' experiences of changing teaching language. *Iberica: Revista de la Asociacion Europea de Lenguas para Fines Especificos* 22(22). 35–54.

Airey, John. 2012. I don't teach language. The linguistic attitudes of physics lecturers in Sweden. *AILA Review* 25. 64–79.

Arco-Tirado, José Luis, Fransisco Fernández-Martín, Ana M. Ramos-García, Levente Littvay, Javier Villoria & Jose A. Naranjo. 2018. A counterfactual impact evaluation of a bilingual program on students' grade point average at a Spanish university. *Evaluation and Program Planning* 68. 81–89.

Bakker, Arthur, Jinfa Cai, Lyn English, Gabriele Kaiser, Vilma Mesa & Wim Van Dooren. 2019. Beyond small, medium, or large: Points of consideration when interpreting effect sizes. *Educational Studies in Mathematics* 102. 1–8.

Başıbek, Nurcihan, Mustafa Dolmacı, Behice Ceyda Cengiz, Burcu Bür, Yeşim Dilek & Bayram Kara. 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 Behavioral Sciences* 116. 1819–1825.

Briggs, Jessica G., Julie Dearden & Ernesto Macaro. 2018. English medium instruction: Comparing teacher beliefs in secondary and tertiary education. *Studies in Second Language Learning and Teaching* 8(3). 673–696.

Cheung, Alan CK & Robert E. Slavin. 2016. How methodological features affect effect sizes in education. *Educational Researcher* 45(5). 283–292.

Cho, Dong Wan. 2012. English-medium instruction in the university context of Korea: Tradeoff between teaching outcomes and media-initiated university ranking. *Journal of Asia TEFL* 9(4). 135–163.

Civan, Abdulkadir & Ali Coşkun. 2016. The effect of the medium of instruction language on the academic success of university students. *Educational Sciences: Theory and Practice* 16(6). 1981–2004.

Cliff, Norman. 1993. Dominance statistics: Ordinal analyses to answer ordinal questions. *Psychological Bulletin* 114(3). 494–509.

Costa, Francesca & Christina Mariotti. 2017. Differences in content presentation and learning outcomes in English-medium instruction (EMI) vs. Italian-medium instruction (IMI) contexts. In Jennifer Valcke & Robert Wilkinson (eds.), *Integrating content and language in higher education: Perspectives on professional practice*, 184–204. Frankfurt: Peter Lang.

Council of Europe. 2020. Common European framework of reference for languages self-assessment grid: Common reference levels. https://www.coe.int/en/web/common-european-framework-reference-languages/table-2-cefr-3.3-common-reference-levels-self-assessment-grid (Retrieved 20 March 2023).

Curle, Samantha, Reka Ratkaine Jablonkai, Jenna Mittelmeier, Kari Sahan, Ann Veitch & Nicola Galloway. 2020. English medium part 1: Literature review. In Nicola Galloway (ed.), English in higher education (Report No. 978-0-86355-977–8). British Council.

Dafouz, Emma & M. Mar Camacho-Miñano. 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, Emma, M. Mar Camacho & Elena 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.

de Vos, Johanna F., Herbert Schriefers & Kristin Lemhöfer. 2020. Does study language (Dutch versus English) influence study success of Dutch and German students in the Netherlands? *Dutch Journal of Applied Linguistics* 9(1–2). 60–78.

Ellili-Cherif, Maha & Haitham Alkhateeb. 2015. College students' attitude toward the medium of instruction: Arabic versus English dilemma. *Universal Journal of Educational Research* 3(3). 207–213.

Evans, Stephen & Bruce Morrison. 2011. Meeting the challenges of English-medium higher education: The first-year experience in Hong Kong. *English for Specific Purposes* 30(3). 198–208.

Franke, Todd Michael, Timothy Ho & Christina A. Christie. 2012. The chi-square test: Often used and more often misinterpreted. *American Journal of Evaluation* 33(3). 448–458.

Gad Consulting Services. 2018. Statistical methods in toxicology. *Comprehensive Toxicology* 3. 281–293.

Galloway, Nicola, Jaroslaw Kriukow & Takuya Numajiri. 2017. *Internationalisation, Higher education and the growing demand for English: An investigation into the English medium of instruction (EMI) movement in China and Japan*. London: British Council. https://www.research.ed.ac.uk/en/publications/internationalisation-higher-education-and-the-growing-demand-for- (Retrieved 23 January 2023).

Guarda, Marta. 2018. "I just sometimes forget that I'm actually studying in English": Exploring student perceptions on English-medium instruction at an Italian university. *Rassegna Italiana Di Linguistica Applicata* 2. 129–144.

Guarda, Marta. 2021. *Student Perspectives on English-medium instruction: Insights from an Italian university*. London: Routledge.

Hellekjær, Glen-Ole. 2010. Lecture comprehension in English-medium higher education. *HERMES-Journal of Language and Communication in Business* (45). 11–34. https://doi.org/10.7146/hjlcb.v23i45.97343.

Hollis, Sally & Fiona Campbell. 1999. What is meant by intention to treat analysis? Survey of published randomised controlled trials. *BMJ* 319(7211). 670–674.

Hulstijn, Jan H. 2007. The shaky ground beneath the CEFR: Quantitative and qualitative dimensions of language proficiency. *The Modern Language Journal* 91(4). 663–667.

Ingvarsdóttir, Hafdís & Birna Arnbjörnsdóttir. 2015. English in a new linguistic context: Implications for higher education. In Slobodanka Dimova, Anna Christina Hultgren & Christian Jensen (eds.), *English-medium instruction in European higher education*, 137–156. Boston, MA: De Gruyter.

Joe, Yoon Ju & Hee-Kyung Lee. 2013. Does English-medium instruction benefit students in EFL contexts? A case study of medical students in Korea. *The Asia-Pacific Education Researcher* 22(2). 201–207.

Jordan, Katy. 2015. Massive open online course completion rates revisited: Assessment, length and attrition. *International Review of Research in Open and Distributed Learning* 16(3). 341–358.

Kojima, Naoko. 2021. *Student motivation in English-medium instruction: Empirical studies in a Japanese university*. Abingdon: Routledge.

Lasagabaster, David. 2022. Teacher preparedness for English-medium instruction. *Journal of English-Medium Instruction* 1(1). 48–65.

Lasagabaster, David & Aintzane Doiz (eds.). 2021 *Language use in English-medium instruction at university: International perspectives on teacher practice*. London: Routledge.

Lee, Given & Christiaan Prinsloo. 2018. Pedagogic challenges of English-mediated instruction in the Korean EFL context. *Journal of Asia Tefl* 15(1). 217.

Leon, Andrew C. 1998. Descriptive and inferential statistics. *Comprehensive Clinical Psychology* 3. 243–285.

Lueg, Klarissa & Rainer Lueg. 2015. Why do students choose English as a medium of instruction? A Bourdieusian perspective on the study strategies of non-native English speakers. *Academy of Management Learning & Education* 14(1). 5–30.

Macaro, Ernesto. 2018. *English medium instruction*. Oxford: Oxford University Press.

Macaro, Ernesto, Samantha Curle, Jack Pun, Jiangshan An & Julie Dearden. 2018. A systematic review of English medium instruction in higher education. *Language Teaching* 51(1). 36–76.

Malmström, Hans & Diane Pecorari. 2021. Epilogue: Disciplinary literacies as a nexus for content and language teacher practice. In David Lasagabaster & Aintzane Doiz (eds.), *Language use in english-medium instruction at University*, 213–221. Routledge.

Malmström, Hans & Diane Pecorari. 2022. *Language choice and internationalisation: The roles of Swedish and English in research and higher education. (Report):* The Language Council of Sweden. http://sprakochfolkminnen.diva-portal.org/smash/get/diva2:1681930/FULLTEXT01.pdf (Retrieved 5 January 2023).

Park, H. 2007. English medium instruction and content learning. *English Language and Linguistics* 23. 257–274.

Pecorari, Diane & Hans Malmström. 2018. At the crossroads of TESOL and English medium instruction. *TESOL Quarterly* 52(3). 497–515.

Reus, Lorenzo. 2020. English as a medium of instruction at a Chilean engineering school: Experiences in finance and industrial organization courses. *Studies in Educational Evaluation* 67. https://doi.org/10.1016/j.stueduc.2020.100930.

Rosner, Bernard, Robert J. Glynn & Mei-Ling Ting Lee. 2003. Incorporation of clustering effects for the Wilcoxon rank sum test: A large-sample approach. *Biometrics* 59. 1089–1098.

Roussel, Stéphanie, Danielle Joulia, André Tricot & John Sweller. 2017. Learning subject content through a foreign language should not ignore human cognitive architecture: A cognitive load theory approach. *Learning and Instruction* 52. 69–79.

Rubio-Alcalá, Fernando D., José Luis Arco-Tirado, Francisco D. Fernández-Martín, Rocío López-Lechuga, Elvira Barrios & Víctor Pavón-Vázquez. 2019. A systematic review on evidences supporting quality indicators of bilingual, plurilingual and multilingual programs in higher education. *Educational Research Review* 27. 191–204.

Sinha, Yogesh, Thomas Roche & Manisha Sinha. 2018. Understanding higher education attrition in English-medium programs in the Arab Gulf states: Identifying push, pull and fallout factors at an Omani university. In Rahma Al-Mahrooqi & Christopher Denman (eds.), *English education in Oman: Current scenarios and future trajectories*, 195–229. Singapore: Springer.

Slavin, Robert & Dewi Smith. 2009. The relationship between sample sizes and effect sizes in systematic reviews in education. *Educational Evaluation and Policy Analysis* 31(4). 500–506.

Staub, Donald F. 2022. Why student retention matters for Turkish EMI universities? In Yasemin Kirkgöz & Ali Karakaş (eds.), *English as the medium of instruction in Turkish higher education*, 257–273. Cham: Springer.

Tatzl, Dietmar & Bernd Messnarz. 2013. Testing foreign language impact on engineering students' scientific problem-solving performance. *European Journal of Engineering Education* 38(6). 620–630.

Thøgersen, Jacob & John Airey. 2011. Lecturing undergraduate science in Danish and in English: A comparison of speaking rate and rhetorical style. *English for Specific Purposes* 30(3). 209–221.

Vinke, Adriana Anthonia. 1995. "English as the medium of instruction in Dutch engineering education." Doctoral Dissertation. The Netherlands: Delft University of Technology.

Yuksel, Doğan, Adem Soruç, Mehmet Altay & Samantha Curle. 2023. A longitudinal study at an English medium instruction university in Turkey: The interplay between English language improvement and academic success. *Applied Linguistics Review* 14(3). 533–552.