What Predicts Academic Achievement in EMI Courses? Focus on Vocabulary Knowledge and Self-Perceptions of L2 Skills

Ahmed Masrai, Dina Abdel Salam El-Dakhs, and Noorchaya Yahya

Abstract
This study examines the relationship between L2 vocabulary knowledge, self-rating of word knowledge, self-perceptions of four language skills (listening, reading, speaking, and writing), and students’ academic achievement. An objective measure of lexical knowledge and questionnaire on self-perception and self-rating of vocabulary knowledge were administered to 106 undergraduate students in an English as a medium of instruction (EMI) program. The students’ academic achievement was measured through their EMI course grades. Results showed positive significant correlations between learners’ vocabulary knowledge, whether measured objectively or subjectively, and self-perceptions of the four skills. Results also showed that vocabulary knowledge was the strongest correlate with academic achievement, followed by self-perception. Interestingly, vocabulary knowledge explained the largest unique variance in the learners’ academic achievement. The contribution of students’ self-perceptions of L2 use, although predicted as an additive value to the model of academic achievement, was minimal compared to that explained by vocabulary knowledge. Findings of the study address implications directed to teaching and assessing lexical knowledge in the EMI context.

Keywords
vocabulary knowledge, English as a medium of instruction, academic achievement, self-perceptions, language skills

Introduction
The last three decades have witnessed a dramatic increase in the number of universities in non-English speaking countries that offer their programs through English as a medium of instruction (EMI). While some of these programs focus on the delivery of content in English without explicit language objectives (e.g., Airey, 2016), other programs aim for the simultaneous mastery of both content and the English language to varying degrees (Chapple, 2015; Taguchi, 2014). The EMI programs with a clear focus on language skills fall within the framework of content-based instruction (CBI) models of second language learning which provide content instruction while developing the learners’ second/foreign language (Brinton & Snow, 2017). Whether the EMI program is content oriented or content-and-language focused, learners’ lexis plays an extremely important role in these EMI programs because lexical knowledge is a prerequisite for other language skills, such as reading, listening, speaking, and writing (e.g., Alderson & Banerjee, 2001; Staehr, 2008) and for overall language proficiency (Lafer, 1989; Milton, 2009; Nation, 2001). In fact, there is a general agreement in the literature that the minimal threshold of passive vocabulary needed for enrolment in English medium universities is 5,000-word families (Meera et al., 1997; Nation, 1990; Schmitt, 2000). Passive vocabulary refers to the number of words a learner knows receptively (i.e., for the purpose of listening and reading; Milton, 2009). It is thus logical that vocabulary knowledge should be a key determinant of EMI students’ experience and academic achievement.

Due to the importance of vocabulary knowledge for EMI students, some studies have recently examined the development of vocabulary knowledge in the EMI context. For example, Malmström et al. (2016) examined the claim that EMI is beneficial for students’ development of lexical knowledge. The corpus data included texts produced by Swedish
university Master of Science students whose English language proficiency was assessed at an advanced level. Examining the students’ texts revealed that academic vocabulary items accounted for approximately 20% of the students’ writing. Additionally, no significant differences were found between home and international students in any of the vocabulary measures used in the study, such as lexical sophistication and diversity. Regarding the development of lexical knowledge, the findings showed significant but very modest learning gains across years of study.

Another interesting study was conducted by Coxhead and Boutorwick (2018) who investigated the development of students’ vocabulary knowledge at an international high school in Germany in which English was used as a medium of instruction. The vocabulary knowledge of 468 participants was tested on entry to their first year of secondary school (Grade 6) and annually thereafter for a period of 6 years. The Vocabulary Levels Test (VLT; Nation, 1983; Schmitt et al., 2001) was the assessment measure of vocabulary knowledge. The results showed that while native speakers had mastery of the high frequency (first 2,000 level) words of the VLT in Grade 6, non-native speakers had varied scores on the 2,000 level. Based on these results, the non-natives with low scores were given English as an additional language support, which enabled them to master the first 2,000 level of the VLT by Grade 9. Interestingly, the native and non-native speakers, including those with low and high scores on the 2,000 level in Grade 6, scored similarly at the 3,000 level and academic sections of the test by Grade 10. Hence, providing additional language support seemed to enhance language learners’ development of vocabulary knowledge, but the effect may require sustained exposure/practice over a long period.

In the same vein, Zhang and Pladevall-Ballester (2021) examined language learning outcomes data in two EMI programs, namely Film Production and International Trade, in China. The students took pre-post discipline-specific vocabulary and writing tasks as well as pre-post grammar proficiency tests over the course of one semester and after the same number of hours of exposure. The results revealed that the students’ general grammar proficiency showed almost no progress whereas notable gains were found in productive and receptive vocabulary tasks and in the writing task. The vocabulary learning gains were more notable in the International Trade Program than the Film Production Program while the two programs had almost the same learning gains in writing.

In addition to the development of vocabulary knowledge, other studies have recently investigated the relationship between vocabulary knowledge and academic achievement/success in the EMI context, which is the focus of the current study. For example, Uchihara and Harada (2018) investigated the relationship between the vocabulary knowledge of 35 Japanese-speaking university students and their self-perceptions of performance in the four language skills (reading, listening, writing, and speaking) and academic achievement in an EMI course at a private university in Tokyo, Japan. The students completed a version of Schmitt et al. (2001) Vocabulary Levels Test (VLT), a word matching task that measures meaning recognition of word knowledge in the written form, and McLean et al. (2015) Listening Vocabulary Levels Test (LVLT), a multiple-choice task that assesses aural receptive vocabulary size. In addition to these objective measures, students also self-rated their vocabulary knowledge on a 6-point Likert scale (1 = not confident and 6 = very confident). The students’ scores on the three vocabulary measures were compared with their self-perceptions of language skills on a self-perception questionnaire, and their scores on an EMI course whose primary goal was to develop students’ subject knowledge rather than improve English language skills. The results showed that learners with larger aural vocabulary sizes self-rated their spoken language use more favorably, and the learners who self-rated their vocabulary knowledge at a high level expressed more confidence in their productive language skills. Interestingly, learners with larger written vocabulary sizes were more likely to perceive themselves as less proficient in performing EMI tasks. Surprisingly, none of the vocabulary measures were significantly associated with the students’ academic achievement.

Another relevant study was conducted by Rahman (2020) at a Malaysian public university. The study attempted to investigate the associations among EMI students’ vocabulary size, English language proficiency, and cumulative grade point average (CGPA). The study also aimed to determine whether vocabulary size or language proficiency was a better predictor of students’ CGPA. The students’ vocabulary size was assessed with the use of the Vocabulary Size Test (VST; Nation & Beglar, 2007), their language proficiency was derived from their university English test scores, and their CGPAs were obtained from their academic transcripts. The results showed a moderate and significant positive association among all variables. Additionally, vocabulary size proved to be a better predictor of the students’ CGPA than their language proficiency as it contributed 25% to the overall CGPAs. Hence, contrary to the findings by Uchihara and Harada (2018) and Rahman’s (2020) results demonstrated that vocabulary knowledge is an important contribution to students’ academic achievement in EMI programs.

In the Arab World, in which the current study takes place, only few studies have attempted to examine the relationship between vocabulary knowledge and students’ academic achievement in EMI programs. One interesting study that was carried out in the UK with a focus on Arab students studying in a British university (English as a second language [ESL] context) was Alsaggar and Milton (2016). They attempted to examine whether the students’ lexical repository, intelligence, and foreign language learning aptitude could predict their academic performance. To this end, 36 undergraduate Arab students completed a vocabulary test (El-Dakhs et al., 2022), an intelligence test (Wechsler
Abbreviated Scale of Intelligence [WASI]) and a foreign language aptitude test (Modern Language Aptitude Test [MLAT]; Carroll & Sapon, 1959). Correlating the scores of these measures with the participants’ Grade Point Average (GPA) as a measure of academic achievement, it was clear that a vocabulary knowledge threshold of 5,000 and above is necessary for enrolment in EMI university programs. As for the other measures, no correlation was found between intelligence or foreign language aptitude and the participants’ academic performance, but a significant relationship was noted between intelligence and word knowledge.

Two other interesting studies in this regard were conducted in the Sultanate of Oman by Roche and Harrington (2013) and Harrington and Roche (2014). Carried out in a non-English speaking country, these two studies focused on EFL learners. In their first study, Roche and Harrington (2013) investigated whether lexical knowledge is a predictor of written academic English proficiency (AEP) and overall academic achievement in an English-medium College of Applied Sciences. To this end, 70 freshmen and senior students who were native speakers of Arabic completed a practice IELTS test to assess their academic English written proficiency and a computerized timed YES/NO response test to assess their lexical knowledge. Comparing the scores on these tests with the students’ Grade Point Average (GPA) as a measure of academic achievement, it was found that vocabulary size and speed correlated with both academic writing and GPA measures.

In their second study, Harrington and Roche (2014) examined whether students’ lexical knowledge as measured by the same timed Yes/No test used in Roche and Harrington (2013) can predict the students’ grade point averages (GPA) across different academic disciplines in the same EMI university. The results showed that word accuracy was a better predictor of academic performance than response time and it accounted for as much as 25% of variance in the students’ GPA in the Humanities, Computing, and Business majors. However, the Engineering majors displayed a different pattern of results with response time, not accuracy, being the significant predictor of the students’ GPAs, accounting for 40% of the variance in academic performance.

As it must be noted, the three above studies, which focused on Arab learners of English, used a Yes/No test format to assess vocabulary knowledge. Despite the wide use of this type of test format in vocabulary studies, it has its drawbacks which may cast doubt on its predictive power (e.g., Beeckmans et al., 2001; Harsch & Hartig, 2016), including its reliance on self-reporting, decontextualized presentation of words and high probability of guessing. Hence, it is important to use other measures of vocabulary knowledge to validate the conclusions drawn from this type of test format. These limitations of the vocabulary measure in these studies partially motivated the present study which employed a standard test with a multiple-choice format. The current study was also motivated by the contradiction between the results of the three studies including Arab learners and the findings in Uchihara and Harada’s (2018) study. It was not clear why vocabulary size was dissociated from academic achievement in the Japanese context while it had a significant effect in the Arab context.

Based on this contradiction, these questions were raised: Why did measures of lexical knowledge lose their predictive power of academic achievement in the EMI context in Japan? Is it a difference of test format (i.e., yes/no test vs. matching or multiple-choice test)? Is it the different measurement of academic achievement (GPA vs. a course grade)? Was the focus on the mastery of content instead of language skills in the target course of Uchihara and Harada’s (2018) study a contributing factor?

The current study, which partially replicates Uchihara and Harada’s (2018) work, thus aims to examine the relationship between vocabulary knowledge, as measured objectively through vocabulary tests and subjectively through self-rating, the students’ self-perceptions of L2 use (i.e., the students’ perception of their performance in reading, writing, speaking, and listening) and the students’ academic achievement in EMI programs. This study will examine the said relationship in a specific form of EMI courses; namely, content-based instruction (CBI) courses which aim to develop both the learners’ knowledge of the subject matter while enhancing their mastery of academic language use. The focus of this study is on the role of vocabulary knowledge in the success of university EMI students in the Saudi context. In this context, English language teaching has witnessed some changes through its history. For example, since its inclusion in the school curriculum, English was only introduced to the students at the age of 13. In 2004, the Saudi Ministry of Education included sixth graders (at the age of 12) in its English language teaching plan. Another change was introduced in 2011 to 2012, where the Ministry started teaching English from the fourth grade (at the age of 10). Interestingly, the Ministry has approved teaching English to the students from grade 1 (at the age of 6) from this academic year (2021). Furthermore, a number of Saudi universities have recently adopted EMI in their curriculum. This change has posed some challenges to the students, and lack of vocabulary knowledge was suggested to be one of the determinant factors of inadequate understanding of academic lectures (El-Dakhs et al., 2021).

The current study is sought to contribute to our understanding of the role of vocabulary knowledge in students’ academic achievements in the EMI context, particularly in Saudi Arabia.

More specifically, the current study aims to answer the following research questions:

1. Is there a relationship between vocabulary knowledge, self-rating of vocabulary knowledge, and students’ self-perceptions of L2 performance?
2. To what extent do the variables in RQ1 associate with students’ academic achievement?
3. To what extent can vocabulary knowledge, self-rating of vocabulary knowledge, and self-perception of L2 performance predict the variance in students’ academic achievements in the EMI context?

**Method**

**Participants**

A total of 106 female undergraduates at an English language and translation department at a public Saudi university took part in the study. The participants were all Saudi nationals and native speakers of Arabic. Their ages ranged between 20 and 24 ($M_{age} = 20.9, SD=1.8$). The participants belonged to four intact classes of the course “Advanced Readings in Culture.” The participants were all second-year students studying in the English language and translation department.

The “Advanced Readings in Culture” course is a 30-hour CBI course that aims to enhance the students’ cultural awareness while boosting their mastery of academic English. As Richards and Rodgers (2001) put it, “Content-Based Instruction refers to an approach to second language teaching in which teaching is organized around the content or information that students will acquire, rather than around a linguistic or other type of syllabus” (p. 204). The course is organized around cultural concepts, such as sub-cultures, culture and ideology, mass media and modern culture, and the issue of identity. Meanwhile, the course aims to enhance the students’ English language knowledge and skills through raising their awareness of the relation between language and culture (e.g., how language varies based on ethnicity, gender, nationality, social class, and identity) and engaging them in several oral and written academic tasks in English (e.g., oral presentations, academic essays, etc.) for which the feedback is provided on both the content and the language use. Hence, the course is a typical CBI-course that focuses on “the teaching of language through exposure to content that is interesting and relevant to learners” (Brinton, 2003, p. 201).

**Instruments**

**New Vocabulary Levels Test (NVLT)**

The NVLT (Webb et al., 2017) is a receptive vocabulary test. It employs a matching format through which the test-takers are presented with 150 questions across 5 frequency-levels (1,000, 2,000, 3,000, 4,000, and 5,000). Each frequency-level includes 30 questions divided into 10 clusters of 6 words (3 key-items and 3 distractors) and 3 definitions. The test-taker’s task is to choose the correct word corresponding to each definition. An example cluster of the test is illustrated in Figure 1.

As seen from this example cluster, each of the three definitions to the left has only one corresponding correct answer from the six options to the right. For example, the option ‘song’ is the correct match to “piece of music.”

**Self-Rating Vocabulary Measure**

A self-rating of vocabulary knowledge was used to assess the learners’ perceived lexical proficiency in English. The participants were asked to rate their knowledge of L2 English vocabulary on a 6-point Likert scale, ranging from 1 (not confident) to 6 (very confident).

**Self-Perceptions of L2 Use Across the Four Skills**

A self-perceptions questionnaire of L2 use in the EMI context developed by Uchihara and Harada (2018) was used in the present study. The questionnaire consists of four sections, each of which targeted listening, speaking, reading, and writing skills. The questionnaire items ($n=29$) were sub-categorized as follows: 6 items in the listening section, 10 items in the speaking section, 8 items in the reading section, and 5 items in the writing section (see Appendix A). The participants were required to respond to each item on a 6-point Likert scale, ranging from 1 (strongly disagree) to 6 (strongly agree).

**Academic Achievement Scores**

The learners’ academic achievement was assessed using the CBI final course grade. This grade was based on three types of course assessment; namely: (1) an oral presentation, (2) two written assignments, and (3) a number of quizzes and a final exam that mainly used closed-ended questions in the form of multiple-choice, matching, and gap-fill tasks. The final accumulated grade, on a 100-point scale, was considered for the data analysis.
Procedure

Four types of data were collected for the purpose of the study using the following procedure: In week 3 of the CBI course, the students completed Uchihara and Harada’s (2018) self-perceptions questionnaire of their L2 use of the four language skills and self-rated their vocabulary knowledge. In week 5 of the course, the students completed Webb et al. (2017) NVLT as an objective measure of their vocabulary knowledge. No time limit was assigned for either of the two tests, but the majority of the students completed the first two tasks together in 20 minutes while they took around 45 minutes to complete the NVLT. After the students have completed their final exam, their overall grade (out of 100) was collected. It should be noted that the students’ program was timely tight and full of prescheduled departmental exams. Therefore, the study instruments administration was only possible in the weeks identified above. However, since the study is not interventional in nature, we believe this should not affect the study results.

Data Analysis

The present study is quantitative in nature. Thus, to answer the research questions, the data were analyzed using various statistical measures, including descriptive statistics, correlational analyses (Spearman rho and Pearson’s correlation), and hierarchical regression analyses. Correlational analyses were employed to answer research questions 1 and 2, and hierarchical regression analyses were performed to answer research question 3. SPSS V. 25 was used to conduct the data analyses.

Results

Descriptive Statistics

Prior to responding to each specific research question, descriptive statistics (means, standard deviations, minimum, and maximum score) of the participants’ scores on NVLT, self-rating of vocabulary knowledge, self-perception of performance on the four skills, and academic achievement were first computed. The results are summarized in Table 1.

Scores on the NVLT were reported for each frequency level and also tallied by summing all sub-scores from each level (1,000, 2,000, 3,000, 4,000, and 5,000). Cronbach’s alpha analysis showed a good reliability index of the NVLT ($\alpha = .84$). The participants, on average, appeared to know about 80% (119.79 out of 150) of the items in NVLT. However, Webb et al. (2017) recommended cutting point of mastery (29/30 for 1,000–3,000 word-levels; and 24/30 for 4,000 and 5,000 word-levels), suggests that the learners, as a group, tend to only master knowledge of the 1,000-word level of NVLT. The level of mastery of vocabulary across 2,000 to 5,000 word-level appeared to decline in relation to word frequency. A profile of the learners’ knowledge of each frequency level, high on the left and tapering off to the right, is illustrated in Figure 2.

The learners’ self-perception of their vocabulary knowledge was 4.56 on a Likert-scale from 1 to 6. This shows that they appeared fairly confident of their vocabulary knowledge. In respect of the learners’ self-perceptions of L2 performance on the four skills, all items appeared to have a good reliability index ($\alpha = .86$). Further, the reliability indices were also acceptable for each sub-category: listening ($\alpha = .73$), speaking ($\alpha = .86$), reading ($\alpha = .83$), and writing ($\alpha = .84$). The learners’ average score on a 6-point scale suggest that they positively perceived their L2 performance on the four language skills ($M=4.82$, $SD=0.61$). However, sub-category results showed that the learners were relatively more confident in the receptive language skills (listening, $M=5$, $SD=0.69$; reading, $M=4.85$, $SD=0.70$) than in the productive skills (speaking, $M=4.64$, $SD=0.75$; writing, $M=4.56$, $SD=0.79$). Detailed descriptive statistics of the participants’ self-perceptions of their performance on each

Table 1. Descriptive Statistics of Scores on NVLT, Self-Rating of Vocabulary Knowledge (SR), Self-Perception of Performance in the Four Skills (SP), and Academic Achievement (AA).

| Measure          | MPS  | Minimum | Maximum | $M$ (%) | SD  |
|------------------|------|---------|---------|---------|-----|
| NVLT 1,000       | 30   | 25      | 30      | 29.62 (99) | 0.86 |
| NVLT 2,000       | 30   | 4       | 30      | 27.65 (92) | 3.77 |
| NVLT 3,000       | 30   | 5       | 30      | 23.81 (79) | 4.94 |
| NVLT 4,000       | 30   | 3       | 30      | 20.38 (68) | 6.38 |
| NVLT 5,000       | 30   | 1       | 30      | 18.33 (61) | 6.89 |
| NVLT Total       | 150  | 46      | 149     | 119.79 (80) | 19.63 |
| SR               | 6    | 3       | 6       | 4.56    | 0.95 |
| SP listening     | 6    | 3       | 6       | 5       | 0.69 |
| SP speaking      | 6    | 2.70    | 6       | 4.64    | 0.75 |
| SP reading       | 6    | 2.38    | 6       | 4.85    | 0.70 |
| SP writing       | 6    | 2.80    | 6       | 4.56    | 0.79 |
| SP average       | 6    | 3.60    | 5.90    | 4.82    | 0.61 |
| AA               | 100  | 62.60   | 98.60   | 84.73   | 7.95 |

Note. $N=106$; MPS = maximum possible score.
SAGE Open

item of the four skills measures are reported in Table B1 of Appendix B.

The results of academic achievement of learners enrolled in the EMI course, as measured by their grades on a 100-point scale, suggest that most of them passed the course with very good standing ($M = 84.73$, $SD = 7.95$).

RQ1: Is there a relationship between vocabulary knowledge, self-rating of vocabulary knowledge, and students’ self-perceptions of L2 performance?

In order to answer the first research question, correlation analyses were performed between two measures of vocabulary knowledge (NVLT—objective measure and self-rating—subjective measure) and learners’ self-perceptions of L2 use. As the scores from self-perceptions are not normally distributed, Spearman rho correlation analysis was operationalized. The inter-correlation results are shown in Table 2. Interestingly, both vocabulary measures were found to correlate significantly with the learners’ self-perceptions of performance in the four language skills. However, scores on the NVLT correlated more strongly with the receptive language skills ($r_s = .32$ with listening; $r_s = .32$ with reading) than the productive skills ($r_s = .28$ with speaking; $r_s = .24$ with writing). Although these correlations were significant, it should be noted that they are only moderate. This could be attributed to the fact that self-perceptions measures might not be sensitive enough to elicit more accurate performance in the language skills. Furthermore, the NVLT is a receptive measure of lexical knowledge, thus, it might not be sensitive to tap depth of knowledge associated with productive skills like speaking and writing. Self-rating, on the other hand, while correlated most strongly with the reading skill ($r_s = .55$), it correlated less strongly with listening ($r_s = .39$) than the productive skills ($r_s = .42$ with speaking; $r_s = .48$ with writing). In a general sense, the results of the correlation between self-rating and self-perceptions indicate that learners who are confident in their vocabulary knowledge view themselves as proficient users of English.

To further examine the relationship between vocabulary knowledge and the learners’ self-perceptions of L2 performance, correlation analysis was conducted between each frequency level of the NVLT, and self-perception of each language skill. Correlation results are summarized in Table 3. While the overall score on the NVLT significantly correlated with the learners’ self-perceptions of L2 use, the correlation analysis was conducted between each frequency level of the NVLT, and self-perception of each language skill. Correlation results are summarized in Table 3. While the overall score on the NVLT significantly correlated with the learners’ self-perceptions of their L2 use of each language skill, it was not the case with each frequency level of NVLT. The 1,000-frequency level did not correlate with any language skill, and 2,000-frequency level was only found to correlate with listening ($r_s = .22$). Scores on the 3,000, 4,000, and 5,000-frequency levels were significantly correlated with the four skills average scores. These results suggest that learners with vocabulary knowledge in levels beyond the 1,000 and 2,000-frequency levels most likely perceive themselves as competent users of L2 in the EMI context.

RQ2: To what extent do the variables in RQ1 associate with students’ academic achievement?

![Figure 2. Vocabulary knowledge profile across frequency levels (NVLT).](image)

Table 2. Correlation Between Vocabulary Measures and SP.

| NVLT (ES) | SR (ES) |
|----------|---------|
| SP listening | .32** (.10) | .39** (.15) |
| SP speaking | .28** (.08) | .42** (.18) |
| SP reading | .32** (.10) | .55** (.30) |
| SP writing | .24* (.06) | .48** (.23) |
| SP average | .33** (.11) | .53** (.28) |

Note. $N = 106$; ES = correlation effect size.
*Correlation is significant at the .05 level (two-tailed).
**Correlation is significant at the .01 level (two-tailed).
In response to the second research question, correlation analysis was performed. Results of the correlation analysis showed a significant positive association of scores on NVLT and self-perception performance in the four skills (\(r = .37, p < .01\); \(r = .31, p < .01\), respectively) with learners’ academic achievement. However, the results revealed no significant correlation between the students’ self-rating of vocabulary knowledge and their academic achievement. This absence of correlation between the two may suggest that self-rating of vocabulary knowledge is not a sensitive measure of academic success. Alternatively, a direct objective measure, such as the NVLT, may be a good predictor of academic performance in EMI situations.

To further explore the data, correlation analysis was also performed between each frequency level of the NVLT and the learners’ academic achievement in the EMI course. The results indicated no significant correlation between the 1,000-frequency level and the learners’ scores. However, vocabulary knowledge on the 2,000, 3,000, 4,000, and 5,000-frequency levels were found significantly correlated with the learners’ academic achievement (\(r = .27, p < .01\); \(r = .37, p < .01\); \(r = .30, p < .01\); \(r = .28, p < .01\), respectively). This finding suggests the importance of word knowledge beyond the most 1,000 frequent words to academic achievement. The results also showed a significant positive association between the learners’ academic success and their self-perceptions of listening (\(r_s = .30\)), speaking (\(r_s = .32\)), and writing (\(r_s = .35\)) abilities but not with reading.

### RQ3: To What Extent Can Vocabulary Knowledge, Self-Rating of Vocabulary Knowledge, and Self-Perception of L2 Performance Predict the Variance in Students’ Academic Achievements in the EMI Context?

In order to answer the third research question, multiple hierarchical regression analyses were computed. These analyses were performed to examine the predictive value of students’ academic achievement based on their vocabulary knowledge scores on NVLT, self-rating of vocabulary knowledge, and self-perceptions of performance on the four skills. A summary of the hierarchical multiple regression analyses is provided in Table 4. Results showed that a significant regression equation was found (\(F(3,102) = 7.87, p < .001\)), with an \(R^2\) of

### Table 3. Correlation Between Scores on VLT of Each Frequency Level and SP.

| VLT (ES) | 1,000  | 2,000  | 3,000  | 4,000  | 5,000  |
|----------|--------|--------|--------|--------|--------|
| SP listening | .13    | .22* (.05) | .37** (.14) | .32** (.10) | .24* (.06) |
| SP speaking  | .11    | .18    | .25* (.06) | .32** (.10) | .26** (.07) |
| SP reading   | .04    | .18    | .31** (.10) | .33** (.11) | .33** (.11) |
| SP writing   | .02    | .15    | .22* (.05) | .22* (.05) | .28** (.08) |
| SP average   | .06    | .20* (.04) | .34** (.12) | .33** (.11) | .31** (.10) |

Note: \(N = 106\); ES = correlation effect size.

*Correlation is significant at the .05 level (two-tailed).

**Correlation is significant at the .01 level (two-tailed).

### Table 4. Regression Models Addressing the Predictive Value of Vocabulary Measures and SP on Academic Achievement.

| Models | R 2 | \(\Delta R^2\) | \(\Delta R^2\) Unstandardized | Standardized |
|--------|-----|----------------|------------------------------|--------------|
| Step 1 | .34 | .12**          | 68.07                        | .34**        |
| Constant| 68.07 | 4.53         |                              |              |
| Step 2 | .40 | .16**          | 0.14                         | .28**        |
| Constant| 0.14 | 0.04         |                              |              |
| Step 3 | .43 | .19            | 57.57                        | .22          |
| Constant| 57.57 | 6.39         |                              |              |
| Step 3 | .43 | .19            | 2.80                         | .32**        |
| Constant| 2.80 | 1.23         |                              |              |
| Step 3 | .43 | .19            | -1.63                        | -0.20        |
| Constant| -1.63 | 0.88         |                              |              |

Note: \(N = 106\).

*p < .05 (two-tailed). **p < .01 (two-tailed).
Discussion

This study set out to explore the role of vocabulary knowledge in EMI success, and whether confident self-perceptions of L2 use predict variance in students' course grade, over and above vocabulary knowledge. Our results reveal a number of intriguing findings. Two of which are related to the descriptive statistics of the NVLT, self-rating of vocabulary knowledge, and the self-perceptions of language skills. First, although the students' scores on the NVLT suggest their mastery of only the first 1,000-word level, according to Webb et al. (2017) recommended cutting point of mastery, the students rated their vocabulary knowledge at 4.56 out of 6. This overstated self-rating of vocabulary knowledge does not accord with the learners' actual known vocabulary from the objective measure (NVLT). This result suggests that reliance on only self-reported competence of vocabulary knowledge may not accurately reflect students' academic performance. We thus recommend the use of documented objective measures of vocabulary knowledge to better establish this kind of relationship. Suggested measures include the NVLT (Webb et al., 2017) and the Vocabulary Size Test (Nation & Beglar, 2007). Second, the students seemed generally more confident in their receptive than productive language skills. This self-perception reflects a classic finding in the literature that productive language skills often lag behind receptive language skills for foreign language learners (e.g., Fan, 2000; Laufer, 1998; Laufer & Paribakht, 1998; Webb, 2008). There may thus be a need to engage students in a variety of productive language activities in order to further enhance their confidence in their productive skills.

The study also presents interesting findings in relation to the association between vocabulary knowledge and the students' self-perceptions of the four language skills. The results reveal a positive correlation between the NVLT scores, particularly those beyond the 2,000 word-level, and the students' self-perceptions. This means that the students with good vocabulary knowledge at the 3,000-, 4,000-, and 5,000-word levels are more likely to perceive themselves as proficient language users. This result comes in contrast to Uchihara and Harada's (2018) finding that learners with a large written vocabulary size were more likely to perceive themselves as less proficient in performing EMI tasks. A possible explanation for this difference could relate to the fact that Uchihara and Harada's (2018) study was related to a content-based course. Hence, the students whose general English vocabulary size was large could have still faced serious difficulties with completing the required tasks in the course which mainly required the use of technical vocabulary. What could have even made the tasks more challenging for the students is that the tasks were mainly productive and open-ended.

Similar to the NVLT, the students' self-rating of vocabulary knowledge correlated positively with the students' self-perceptions. The association was stronger between the students' self-rating of vocabulary knowledge and their self-perception of the reading skill than the skills of listening, speaking, and writing. This means that the students who were confident in their vocabulary knowledge were more likely to perceive themselves as proficient language users, and particularly good readers.

The results in relation to students' academic achievement in the EMI course are similarly revealing. A strong association was found between the vocabulary knowledge scores and the students' scores on the EMI course. Regression analyses suggest that vocabulary knowledge proves to be the strongest predictor of students' academic success. These findings come in line with the works of Alsager and Milton (2016), Roche and Harrington (2013), Harrington and Roche (2014), and Rahman (2020) which showed positive correlations between learners' vocabulary knowledge and their EMI academic achievement. The results, however, do not conform with the findings of Uchihara and Harada (2018) in which none of the vocabulary measures was significantly associated with the students' academic achievement. This difference may again be attributed to the fact that the current study relied on a CBI course for measuring students' academic achievement while Uchihara and Harada (2018) made use of a course that solely focused on the development of content knowledge regardless of language skills. Hence, it is highly likely that the learners' vocabulary knowledge, particularly general vocabulary knowledge in the case of NVLT, will show stronger association with a CBI course that partially aims at enhancing the learners' mastery of language skills. Perhaps, a measure of technical vocabulary knowledge in the case of Uchihara and Harada (2018) would have proved more relevant due to the focus of their EMI course on...
the mastery of the subject matter. Similarly, the present study has the limitation of not considering a measure of technical vocabulary knowledge or a measure of aural vocabulary knowledge. Operationalizing such measures could have revealed interesting findings.

It is also important to examine the association between the students’ self-rating of their vocabulary knowledge and their self-perceptions of language skills and their academic achievement. Interestingly, the students’ self-rating of vocabulary knowledge did not show a significant association with the students’ academic achievement. This comes in line with our earlier comment on the students’ overstated self-rating of their vocabulary knowledge. As for students’ self-perceptions, the results revealed a positive association between the students’ self-perceptions of the listening, speaking, and writing skills and their academic achievement. The predictive power of the students’ self-perceptions of language skills was, however, marginal in relation to academic achievement compared to that of vocabulary knowledge, as measured with the NVLT. Three comments are worth mentioning here. First, the predictive power of the NVLT of the students’ academic achievement is the strongest in the current study. Knowledge of receptive vocabulary knowledge could explain variation in the students’ academic scores relatively well. Second, the students with more confidence in their language skills seemed to perform better in their CBI course. Achieving high scores on courses can, after all, increase students’ confidence. It can simultaneously enhance their motivation to improve their language and, hence, maintain their high scores in CBI courses (Dörnyei, 2005). Third, the students’ self-perceptions of their reading skills did not correlate with their academic achievement. This result could be an artifact of the self-perception measure. It is likely that if an objective measure of the reading skill has been used in relation to academic achievement, a different result might emerge. This finding requires further investigation. A conclusion about whether the students in the current study expressed too high or too low confidence in their reading skills cannot yet be drawn.

Limitations

This study has a number of limitations that need to be noted. First, data were collected from only female participants. It is important to conduct future studies with gender-balanced samples. Second, the participants were English majors, who have special training in the English language. It is recommended to target students of other disciplines in future studies in order to get a better understanding of academic achievement in EMI programs. This limitation leads to considering vocabulary of technical nature, and not only focusing on general vocabulary, in future work to better understand its value in the EMI context. Third, it will be intriguing to expand our assessment of vocabulary knowledge. While the current study focused on measuring the students’ written receptive vocabulary knowledge, it will be interesting to measure the students’ aural/productive/technical vocabulary knowledge in future studies. Furthermore, it is evident from the participants’ vocabulary knowledge scores on the NVLT, the objective measure of vocabulary, that they overrated their vocabulary knowledge on the self-rating measure. This lends support to the limitation of self-rating vocabulary measures as indicators of academic achievement. Finally, we relied on the accumulative students’ grades to assess their academic achievement. It might be interesting to look at the relationship between different modalities of vocabulary knowledge and several constituents of academic achievement.

Implications and Concluding Remarks

The results of this study highlight important methodological implications for research in the field of EMI. Before conducting cross-study comparisons, it is important to ensure that constructs in the different studies have the same definitions. For example, is academic achievement defined in terms of the accumulative GPA or the final grade of a single course? It is also important to carefully consider the comparability of the target course, types of tasks, and the learning context. In EMI, courses can focus only on the mastery of the subject matter while other courses may aim at the mastery of the subject matter and the language skills simultaneously. Likewise, the types of tasks the students complete in each course vary greatly. They may require receptive, controlled productive, or free productive vocabulary knowledge, for instance.

The study also presents important pedagogical implications. As the focus of this study was on Arabic-speaking learners of EFL, the implications are of particular relevance to them, in hope they inform the learning and teaching in the EMI context. Most importantly, since vocabulary knowledge turned out to be the strongest predictor of variance in academic achievement, greater focus on vocabulary knowledge should be placed in EMI programs, especially the CBI courses. It is particularly important to train students on the use of the first 5,000 word-level in line with the general agreement that this level constitutes the threshold for enrollment in EMI programs (Meara et al., 1997; Nation, 1990; Schmitt, 2000). In this regard, activities on productive as well as receptive vocabulary knowledge should be balanced in order to boost students’ confidence in their receptive and productive language skills (e.g., Fan, 2000; Laufer, 1998; Laufer & Paribakht, 1998; Webb, 2008). The results have also showed that students somehow overestimate their vocabulary knowledge, and to a lesser extent, their language skills. Hence, it is recommended to engage students in self-evaluation activities and to provide them with constant feedback on their performance in order to help them assess their competence adequately, and, thus, invest sufficient time and effort to enhance their vocabulary knowledge and language skills.
In terms of assessment, it is recommended that teachers employ objective measures of the language skills than only using self-perceptions questionnaires to better monitor the performance of their learners. Although the findings of the study showed that students’ self-perceptions positively correlated with academic achievement, their contribution to the model of academic achievement was marginal compared with the explanatory power of vocabulary knowledge, as assessed objectively.

Appendix A

The Questionnaire for Measuring Self-Perceptions of Performance

This questionnaire is part of a research project investigating students’ perceptions of performance in English-medium instruction (EMI) courses. Your answers help us learn more about students’ learning experience in future EMI courses offered in the university.

Questions about your perceptions of performance in terms of four language skills in English-medium courses you are currently attending or you have attended before. If you are attending or have attended several courses, answer the questions on the ground of your general impression of these courses.

Strongly disagree 1 2 3 4 5 6 Strongly agree

Listening skill

1. I can understand the content of discussion.
2. I can take notes well.
3. I can understand technical terms used by the lecturer and other students.
4. I can understand the content of questions asked by other students.
5. I can understand course content in a general sense during lecture.
6. I can understand course content in detail.

Speaking skill

7. I can speak in a comprehensible manner.
8. I can state my opinions fluently.
9. I can speak with accurate grammar.
10. I can speak with contextually appropriate vocabulary.
11. I can give a presentation in a clear and logical manner without looking at my note.
12. I can engage in group discussion actively.
13. I can engage in classroom discussion actively.
14. I can answer the questions asked by the lecturer.
15. I can ask questions about unclear content.
16. I can explain the meanings of technical terms.

Reading skill

17. I can read quickly to find the information I need.
18. I can keep reading without being distracted by difficult vocabulary.
19. I can understand technical terms appearing in textbooks.
20. I can read the author’s intention and attitude.
21. I can understand the syllabus and assignment instructions written in English.
22. I can understand the main points.
23. I can understand the content in detail.
24. I can read critically on the basis of my prior knowledge and experience.
Self-Perceptions of Students' Performance Across Four Skills.

| Skill                  | M (Mdn) | SD  | Range     |
|------------------------|---------|-----|-----------|
| Self-perceptions overall scores | 4.8 (4.9) | 0.6 | 3.6–5.9   |
| **Listening (average)** |         |     |           |
| Q1. I can understand the content of discussion. | 5.0 (5.1) | 0.7 | 3–6     |
| Q2. I can take notes well. | 5.5 (6.0) | 0.8 | 1–6     |
| Q3. I can understand technical terms used by the lecturer and other students. | 4.5 (4.5) | 1.2 | 1–6     |
| Q4. I can understand the content of questions asked by other students. | 4.7 (5.0) | 1.2 | 2–6     |
| Q5. I can understand course content in a general sense during lecture. | 5.4 (6.0) | 0.9 | 1–6     |
| Q6. I can understand course content in detail. | 5.6 (5.0) | 1.0 | 2–6     |
| ** Speaking (average)** |         |     |           |
| Q7. I can speak in a comprehensible manner. | 4.7 (5.0) | 1.2 | 1–6     |
| Q8. I can state my opinions fluently. | 4.5 (4.5) | 1.1 | 1–6     |
| Q9. I can speak with accurate grammar. | 4.6 (5.0) | 1.1 | 1–6     |
| Q10. I can speak with contextually appropriate vocabulary. | 4.7 (5.0) | 1.0 | 2–6     |
| Q11. I can give a presentation in a clear and logical manner without looking at my notes. | 4.5 (5.0) | 1.2 | 1–6     |
| Q12. I can engage in group discussion actively. | 5.2 (5.0) | 0.9 | 3–6     |
| Q13. I can engage in classroom discussion actively. | 4.6 (5.0) | 1.2 | 1–6     |
| Q14. I can answer the questions asked by the lecturer. | 4.8 (5.0) | 1.0 | 1–6     |
| Q15. I can ask questions about unclear content. | 4.8 (5.0) | 1.3 | 1–6     |
| Q16. I can explain the meanings of technical terms. | 4.3 (4.0) | 1.2 | 1–6     |
| **Reading (average)** |         |     |           |
| Q17. I can read quickly to find the information I need. | 5.0 (5.0) | 1.0 | 2–6     |
| Q18. I can keep reading without being distracted by difficult vocabulary. | 4.4 (4.0) | 1.1 | 1–6     |
| Q19. I can understand technical terms appearing in textbooks. | 4.1 (4.0) | 1.1 | 1–6     |

Subject #
Thank you for your cooperation.

Appendix B

Self-Perceptions of Students’ Performance Across Four Skills.
## Acknowledgments

The researchers thank Prince Sultan University for funding this research project through the research lab [Applied Linguistics Research Lab-RL-CH-2019/9/1].

## Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This project through Prince Sultan University (RL-CH-2019/9/1) and the Deanship of Scientific Research, King Saud University. This research project was also supported by a grant from the Research Centre for the Humanities, Deanship of Scientific Research, King Saud University.

## Ethical Approval

A permission to conduct this study was sought from the College of Languages and Translation, King Saud University. Verbal consent was sought from all participants.

## ORCID iD

Ahmed Masrai https://orcid.org/0000-0002-6778-5952
Dina Abdel Salam El-Dakhs https://orcid.org/0000-0003-2943-987X
Noorchaya Yahya https://orcid.org/0000-0002-0769-4581

## References

Airey, J. (2016). EAP, EMI or CLIL? In K. Hyland & P. Shaw (Eds.), The Routledge handbook of English for academic purpose (pp. 71–83). Routledge.
Alderson, J. C., & Banerjee, J. (2001). Language testing and assessment (part 1) “state of the art review”. Language Testing, 18, 213–236.
Alsaafer, R., & Milton, J. (2016). Investigating the relationship between vocabulary knowledge and academic success of Arabic undergraduate learners in Swansea University. Language in Focus Journal, 2, 88–124. https://doi.org/10.1515/lifijsal-2016-0010
Beeckmans, R., Eyckmans, J., Janssens, V., Dufranne, M., & van de Velda, H. (2001). Examining the yes/no vocabulary test: Some methodological issues in theory and practice. Language Testing, 18(3), 235–274.
Brinton, D. M. (2003). Content-based instruction. In D. Nunan (Ed.), Practical English language teaching. McGraw-Hill Contemporary.
Brinton, D. M., & Snow, M. A. (2017). The evolving architecture of content-based instruction. In M. A. Snow & D. M. Brinton (Eds.), The content-based classroom: New perspectives on integrating language and content (pp. 2–20). University of Michigan Press.
Carroll, J. B., & Sapon, S. M. (1959). The modern language aptitude test. Psychological Corporation.
Chapple, J. (2015). Teaching in English is not necessarily the teaching of English. International Education Studies, 8, 1–13.
Coxhead, A., & Bourtowitch, 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.
Dörnyei, Z. (2005). The psychology of the language learner: Individual differences in second language acquisition. Lawrence Erlbaum.
El-Dakhs, D.A.S., Masrai, A. & Yahya, N. (2022). University student use and perception of listening comprehension strategies: The case of online EMI lectures. International Journal of Computer-Assisted Language Learning and Teaching, 12(1), 1–21. https://doi.org/10.4018/IJCALLT.291114
Fan, M. (2000). How big is the gap and how to narrow it? An investigation into the active and passive vocabulary knowledge of L2 learners. RELC Journal, 31, 105–119.
Harrington, M., & Roche, T. (2014). Word recognition skill and academic success across disciplines in an ELF university setting. Papers in Language Testing and Assessment, 3(2), 76–99.
Harsch, C., & Hartig, J. (2016). Comparing C-tests and yes/no vocabulary size tests as predictors of receptive language skills. Language Testing, 33(4), 555–575.
Lauffer, B. (1989). What percentage of text is essential for comprehension? In C. Lauren & M. Nordman (Eds.), Special

## Survey Results

| Question                                                                 | M (Mdn) | SD  | Range |
|-------------------------------------------------------------------------|---------|-----|-------|
| Q20. I can understand the author’s intention and attitude.              | 4.4 (4.0) | 1.2 | 1–6   |
| Q21. I can understand the syllabus and the assignment instructions written in English. | 5.7 (6.0) | 0.8 | 3–6   |
| Q22. I can understand the main points.                                  | 5.7 (6.0) | 0.6 | 4–6   |
| Q23. I can understand the content in detail.                            | 4.8 (5.0) | 1.0 | 2–6   |
| Q24. I can read critically on the basis of my prior knowledge and experience. | 4.6 (5.0) | 1.2 | 1–6   |
| **Writing (average)**                                                   | 4.8 (4.8) | 0.8 | 2.8–6.0 |
| Q25. I can write with accurate grammar.                                 | 5.0 (5.0) | 0.8 | 3–6   |
| Q26. I can connect sentences smoothly.                                  | 5.2 (5.0) | 0.8 | 3–6   |
| Q27. I can explain technical terms.                                     | 4.3 (4.0) | 1.1 | 2–6   |
| Q28. I can answer the quizzes and brief questions with confidence.      | 4.8 (5.0) | 1.0 | 3–6   |
| Q29. I can explain logically using long sentences and examples.         | 4.7 (5.0) | 1.2 | 2–6   |

Note. N = 35.
language: From humans thinking to thinking machines (pp. 316–323). Multilingual Matters.
Lauf er, B. (1998). The development of passive and active vocabulary in a second language: Same or different? Applied Linguistics, 19, 255–271.
Lauf er, B., & Paribakht, T. (1998). The relationship between passive and active vocabularies: Effect of language learning context. Language Learning, 48, 365–391.
Malmström, H., Pecorari, D., & Gustafsson, M. (2016). Coverage and development of academic vocabulary in assessment texts in English medium of instruction. In S. Göpferich & I. Neumann (Eds.), Developing and assessing academic and professional writing skills (pp. 46–71). Peter Lang.
McLean, S., Kramer, B., & Beglar, D. (2015). The creation and validation of a listening vocabulary levels test. Language Teaching Research, 19, 741–760.
Meara, P., Lightbown, P. M., & Halter, R. H. (1997). Classrooms as lexical environments. Language Teaching Research, 1(1), 28–47.
Milton, J. (2009) Measuring second language vocabulary acquisition. Multilingual Matters.
Nation, I. S. P. (1983). Testing and teaching vocabulary. Guidelines, 5(1), 12–25.
Nation, I. S. P. (1990). Teaching and learning vocabulary. Heinle & Heinle.
Nation, I. S. P. (2001). Learning vocabulary in another language. Cambridge University Press.
Nation, I. S. P., & Beglar, D. (2007). A vocabulary size test. The Language Teacher, 31(7), 9–13.
Rahman, A. (2020). Receptive vocabulary size as a predictor of undergraduates’ overall cumulative grade point average with English as medium of instruction in universities. Asia-Pacific Social Science Review, 20(3), 56–75.
Richards, J., & Rodgers, T. (2001). Approaches and methods in language teaching (2nd ed.). Cambridge University Press.
Roche, T., & Harrington, M. (2013). Recognition vocabulary knowledge as a predictor of academic performance in an English as a foreign language setting. Language Testing in Asia, 3(12), 1–13.
Schmitt, N. (2000). Vocabulary in language teaching. Cambridge University Press.
Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the vocabulary levels test. Language Testing, 18, 55–88.
Stehr, L. S. (2008). Vocabulary size and the skills of listening, reading and writing. Language Learning Journal, 36(2), 139–152.
Taguchi, N. (2014). Pragmatic socialization in an English-medium university in Japan. International Review of Applied Linguistics in Language Teaching, 52, 157–181.
Uchihara, T., & Harada, T. (2018). Roles of vocabulary knowledge for success in EMI: Self-perceptions and academic outcomes of Japanese undergraduates. TESOL Quarterly, 52(3), 564–587.
Webb, S. (2008). Receptive and productive vocabulary sizes of L2 learners. Studies in Second Language Acquisition, 30(1), 79–95.
Webb, S., Saso, Y., & Ballance, O. (2017). The updated vocabulary levels test: Developing and validating two new forms of the VLT. ITL - International Journal of Applied Linguistics, 168(1), 34–70.
Zhang, M., & Pladevall-Ballester, E. (2021). Discipline-specific language learning outcomes in EMI programs in the People’s Republic of China. Language and Education, 35(4), 357–374.