Recognition Vocabulary Knowledge and Intelligence as Predictors of Academic Achievement in EFL Context

Ahmed Masrai*
King Abdulaziz Military Academy, Saudi Arabia

James Milton
Swansea University, UK

Abstract

Research has shown that general vocabulary knowledge (e.g., Milton & Treffers-Daller, 2013), academic vocabulary knowledge (e.g., Townsend et al., 2012) and general intelligence (e.g., Laidra et al., 2007) are good predictors of academic achievement. While the effect of these factors has mostly been examined separately, Townsend et al. (2012) have tried to model the contribution of general and academic vocabulary to academic achievement and find academic vocabulary knowledge adds only marginally to the predictive ability of general vocabulary knowledge. This study, therefore, examines further factors as part of a more extensive predictive model of academic performance, including L1 vocabulary knowledge, L2 general and academic vocabulary knowledge, and intelligence (IQ) as predictors of overall academic achievement among learners of EFL. Performance on these measures was correlated with Grade Point Average (GPA) as a measure of academic achievement for undergraduate Arabic L1 users (N = 96). The results show positive significant correlations between all the measures and academic achievement. However, academic vocabulary knowledge shows the strongest correlation (r = .72) suggesting that the pedagogical use of this list remains important. To further explore the data, multiple regression and factor analyses were performed. The results show that academic and general vocabulary knowledge combined can explain about 56% of the variance in students' GPAs. The findings, thus, suggest that, in addition to L1 and L2 vocabulary size, and IQ, knowledge of academic vocabulary is an important factor that explains an additional variance in learners' academic achievement.

Keywords: academic achievement, academic vocabulary, general vocabulary, intelligence, L1 vocabulary

Introduction

Academic achievement is crucial in impacting students' future employability and the opportunity to obtain better jobs. It is also a major concern for higher education institutions. Thus, research which taps into modelling the potential factors that might influence student academic success is worthwhile. A number of studies have investigated factors which are thought to influence students' academic success in various contexts (e.g., Laidra, Pullmann, & Allik, 2007; Milton & Treffer-Daller, 2013; Roche & Harrington, 2013; Townsend, Filippini, Collins, & Biancarosa, 2012). Among the factors identified as being associated with learners' overall academic performance have been intelligence, general L2 vocabulary size, L2 academic vocabulary knowledge, and first language (L1) vocabulary size. Despite the influence of these factors on academic success, there is a scarcity of studies examining their effect on achievement with native Arabic learners in the Arab world, with the exception of two studies by Roche and Harrington (2013) and Harington and Roche (2014) who studied the effect of

**E-mail: ahmedalmasrai@hotmail.com; King Abdulaziz Military Academic, Riyadh, Saudi Arabia.*

vocabulary knowledge on students' academic success in Oman. Thus, this study is an attempt to explore the effect of vocabulary knowledge, in L1 and L2, and intelligence on academic performance with learners from L1 Arabic context. There are currently many schools and universities in the Middle East which deliver their programmes through the medium of English, and academic achievement is one of their main concerns. Thus, this study was motivated by both a desire to expand our understanding of the predictors of academic achievement in general and in the Arab world context in particular, and by the scarcity of research in L1 Arabic users studying at higher education institutions through the medium of English in an environment where English is not the primary language used outside the classroom.

Success when studying through a foreign language is likely to be influenced by a range of possible factors and while we have some understanding of the factors through studies which investigate these individually, examining multiple factors as part of an overall predictive model of academic performance is likely to be more useful. Few studies have attempted to place these various factors, including vocabulary knowledge, into an overall model for the prediction of academic success. This study, therefore, will consider incorporation of four independent variables into a model in order to predict academic achievement of native Arabic speakers studying through the medium of English in Saudi higher education institutions.

Vocabulary Knowledge and Academic Achievement

The concept of vocabulary knowledge is not unidimensional and this concept should be clearly defined, particularly when it is referred to in vocabulary testing. Various kinds of vocabulary knowledge can be distinguished (as in Milton & Fitzpatrick, 2014) and it is essential to be clear about what we mean by word knowledge in any study that involves an assessment of vocabulary knowledge. In this study, the vocabulary knowledge referred to is recognition of word form. The term vocabulary size is generally used for a measure of how many words are known this way and, as Milton and Fitzpatrick reflect, this measure of knowledge generally correlates well with measures of all other aspects of vocabulary knowledge.

Vocabulary knowledge is suggested by several studies to be closely linked with various measures of English language ability and academic achievement. Insufficient vocabulary knowledge can impair students' study success (e.g., Alderson, 2005; Daller & Phelan, 2013; Milton & Treffer-Daller, 2013; Roche & Harrington, 2013; Saville-Troike, 1984). Milton and Treffers-Daller (2013) examined the relationship between L1 vocabulary size and academic success with native English speakers (N = 178) at undergraduate level at British universities and found that vocabulary size positively correlated with students' academic attainment (r = .477). In an earlier study, Saville-Troike (1984) investigated the academic success of school children where English as a second language is used as the medium of instruction. Her study concluded that "vocabulary knowledge is the single most important area of L2 competence" (p. 199).

Research also points to the idea that moderate to high correlations exist between general vocabulary size measures and performance in the four skills as measured by tests of academic English such as IELTS. Thus, estimates of vocabulary knowledge correlate with scores in reading comprehension (e.g., Beglar & Hunt, 1999; Laufer, 1992; Qian, 1999; Stæhr, 2008), with writing ability (e.g., Astika, 1993; Laufer, 1998; Stæhr, 2008), with listening comprehension (e.g., Milton, Wade, & Hopkins, 2010; Stæhr, 2008; Zimmerman, 2004), and with oral fluency (Milton et al., 2010; Zimmerman, 2004). The correlations are typically between 0.6 and 0.8 so, not surprisingly, overall vocabulary size alone is often capable of explaining over 50% of variance in scores in foreign language performance (e.g., Stæhr, 2008). This clearly indicates that vocabulary size is a major factor, if not the major factor, in explaining differences in language performance (e.g., Milton, 2013; Stæhr, 2008).

More recently, Harington and Roche (2014) conducted a study to detect academically at-risk students in an undergraduate level studying through the medium of English in Oman. Their findings show that vocabulary size was the best predictor of students' performance, as measured by GPA. In the same vein, Daller and Yixin (2016) found that vocabulary knowledge, as measured by C-test, can explain about 21% of the variance in the international students' academic success.

Academic Vocabulary Knowledge and Academic Achievement

The Academic Word List (AWL) (Coxhead, 2000) is widely used in preparing non-native speakers for academic courses which are taught through the medium of English and it is thought that these words are essential for the understanding of English academic texts (Cobb & Horst, 2004). The rationale for the significance of the AWL comes mainly from the evidence of the contribution to coverage provided by the list. The AWL is generally thought to provide around 10% coverage of academic written texts (e.g., Chen & Ge, 2007; Cobb & Horst, 2004; Coxhead, 2000). Together with the knowledge of the words in West's (1953) General Service List (GSL), the AWL provides approximately 90% coverage of academic written text (e.g., Nation, 2004). A number of research studies have also emphasised the importance of the AWL in academic texts related to specific fields, such as medical research (Chen & Ge, 2007), engineering (Mudraya, 2006), and applied linguistics (Chung & Nation, 2003). Other arguments in support of the AWL's use include Nagy and Townsend (2012), who suggest that the AWL can be very useful in identifying the words and types of words that support learners to access academic texts, and Lesaux, Kieffer, Faller, and Kelley (2010), who propose that the AWL, or subsets of the list used as a goal of learning, promotes a significant improvement in learners' overall vocabulary knowledge. If this small number of words is really very important it should be no surprise if academic words included in the AWL are particularly identified in the setting of L2 programme objectives, the design of lexical syllabi, and in proposals for a learner lexical focus in various stages of L2 learning (e.g., Laufer, 1992; Nation, 2001).

Although words in the AWL are arguably so important in the handling of academic discourse, the number of words which the AWL and GSL combined contain (2,570 words and their derivations) is not adequate for learners to achieve good levels of comprehension in handling academic text. 90% coverage is insufficient for this. Laufer and Ravenhorst-Kalovski (2010) suggest two figures of coverage would be required for different levels of fluency. They suggest a minimal figure of 95% coverage is needed even for adequate comprehension, which they indicate would not satisfy most educators, and an optimal figure of 98% coverage is needed for significantly better comprehension associated with 'functional independence in reading' (p. 25). The minimal coverage figure they propose requires knowledge of the most frequent 4,000 to 5,000 words in English and the optimal coverage figure requires knowledge of the most frequent 8,000 words. Nation (2006) likewise reports a figure of 8,000 to 9,000 words is required for the ideal coverage of 98% for the comprehension of written text, but a slightly smaller figure, 6,000 to 7,000 words for 98% coverage and comprehension of spoken text. The conclusion of this is that learners are likely to need far greater volumes of vocabulary than that provided by the GSL and AWL to reach the levels of knowledge necessary for academic study. Research suggests that if these volumes of words are attained then the AWL will be most likely be known since the AWL falls predominantly within the most frequent 5,000 words in most well-constructed frequency lists.

A new study, Masrai and Milton (forthcoming), indicates that the impact on text coverage of the AWL per se may be overstated since its choice of the GSL as a basic vocabulary list, places underlying general vocabulary knowledge at around the 2,000 word mark. In their analysis of AWL against BNC/COCA, Masrai and Milton found that words from the AWL are heavily concentrated on the 3,000 most frequent words in BNC/COCA. They argued that if Schmitt and Schmitt's (2014) 3,000 word limit for basic vocabulary knowledge were applied in place of GSL, then only 86 AWL words would fall outside this range. The remaining 86 words are not disproportionately frequent in Coxhead's (2000) lists and fall into her least frequent groups, and the additional contribution these words will make to coverage, beyond BNC/COCA 3,000, is likely to be minimal. Masrai and Milton further report that factor analysis suggests that a test of the AWL is, in effect, a test of vocabulary size. Tests of AWL and of general vocabulary size are strongly correlated and appear to test the same factor. However, while the AWL is believed to be an important contributor to attainment in academic study there are few studies which can show that it has an impact additional to general vocabulary knowledge, or that can quantify how large this impact is.

A study that does attempt to quantify the importance of the AWL is by Townsend et al. (2012) who estimate both academic vocabulary knowledge and general vocabulary size in a group of school learners, and use

the scores on these measures to calculate the contribution to academic success that the two types of knowledge can make. The contribution of scores from the two measures to academic success was calculated both individually and combined. They conclude that academic vocabulary knowledge contributes unique variance to achievement across disciplines even when the overall breadth of vocabulary knowledge is controlled. The explanatory power of vocabulary size as a whole was larger than that of academic word knowledge, between 26% and 43% of variance according to discipline. However, academic word knowledge can still add an additional 2% to 7%, depending on discipline, to this explanatory power. These findings appear to suggest that developing a reasonably large vocabulary is more effective for success but that knowledge of the AWL has some additional and marginal influence on academic performance.

The findings from Townsend et al.'s (2012) study is supported by results from Roche and Harrington (2013), who attempt a variety of methodological changes to their test to understand better how vocabulary and academic performance are linked. Their results for the impact of vocabulary size are similar to those of Townsend et al. (2012) and in their study vocabulary size can explain about 25% of the variance in students' GPAs.

Intelligence and Academic Achievement

The question of the relationship between IQ and academic achievement has been addressed by researchers over many years. A number of studies have shown empirical evidence for the strong link between general cognitive ability and academic success. In a study by Jensen (1998), academic achievement of students in high school was found to strongly correlate with IQ, correlations ranged from .50 to .70. A similar correlation (r = .50) between IQ scores and students' grade was also found by Neisser et al. (1996). In another study, Laidra, Pullmann, and Allik (2007) investigate general intelligence with personality traits from the Five-Factor model as predictors of academic achievement in a large sample of Estonian schoolchildren from elementary to secondary school. The results from the study suggest that intelligence is the best single predictor of study success among the Estonian schoolchildren.

Despite the evidence from the literature on the relationship between general intelligence and academic achievement, over 50% of variance in students' academic performance remains unexplained by general intelligence alone. Rohde and Thompson (2007) point out that about 51% to 75% of the variance in academic achievement is not accounted for by the measures of general cognitive ability per se. From this perspective, the question is raised as to whether general intelligence, when examined along other factors, i.e., general vocabulary knowledge and academic vocabulary knowledge, remains the best predictor of students' academic achievement, over and above vocabulary knowledge.

The Study

The aim of this study is to model a number of factors as part of a predictive model of academic performance among native Arabic speakers from an undergraduate population. These factors are general vocabulary knowledge, academic vocabulary knowledge, L1 vocabulary knowledge, and general intelligence. To examine the effectiveness and the predictive power of these factors, individually and combined, on measure of academic achievement, four research questions were addressed:

1. What are the levels of correlation of general vocabulary, academic vocabulary, L1 vocabulary, and IQ with GPA?
2. What is the contribution of each of these variables to academic achievement?
3. Can general vocabulary knowledge and academic vocabulary knowledge explain a unique variance in academic achievement?
4. Can factor analysis allow us to identify whether the vocabulary based variables are identifying separate factors which contribute to GPA?

Method

Participants

Participants in this study were 96 undergraduate students (aged 20-22 years) from two universities in Saudi Arabia. The students were following degree courses in Languages and Translation. The two universities where the participants were drawn implement a very similar programme in English language and translation, so at least in part the input factor from language classroom is controlled. The participants in both institutions were attending levels two, three and four in a four-year degree programme when the data collection for this study took place. Informed consent was obtained from all participants. Also, as a monolingual Arabic vocabulary size test is administered to the participants in the study, only native Arabic speakers were included in the study. The participants' involvement was voluntary.

Instruments

Four measures were used to collect the required data for the current study.

1. The first was a general vocabulary size test (XK-Lex; Masrai & Milton, 2012), which was used to measure the receptive vocabulary knowledge of the participants in the most frequent 10,000 words in English. The XK-Lex is a *yes/no* test of decontextualised words sampled from the first ten 1000 frequency bands in English and includes non-words to control for guesswork.
2. The second was a written receptive vocabulary knowledge test (Arabic-Lex; Masrai & Milton, 2017), which was used to estimate the participants' L1 (Arabic) vocabulary size. This test is similar in its construct to the XK-Lex, but was designed to measure the knowledge of the most frequent 50,000 words in Arabic.
3. The third was a newly developed receptive academic vocabulary size test (AVST; Masrai & Milton, forthcoming) used to assess students' academic vocabulary knowledge of the 570 words from Coxhead's (2000) AWL. The test is similar in its design (frequency based test) to the English and Arabic vocabulary size tests.
4. The final tool was Raven's Standard Progressive Matrices (SPM), a non-verbal IQ test (Raven's matrices, 1998). We chose this version because it was developed to measure a wider range of mental ability and to be equally used with persons of all ages, regardless of their education, background and physical condition (Raven, Raven, & Court, 1998). The test consists of 60 problems divided into five sets (A, B, C, D, and E), each includes 12 problems. All the testing materials were delivered in pencil and paper format and were not timed. However, each of the three vocabulary measures should not take longer than 10-15 minutes to complete. The non-verbal IQ, on the other hand, should take about 45 minutes to finish.

Yes/No tests have been reported in the literature as being suitable, reliable and valid measure of breadth vocabulary knowledge (e.g., Harrington & Carey, 2009; Milton, 2009; Mochida & Harrington, 2006; Read, 2000). It allows for the sampling of a large number of items, and is easy and economical to administer and score. The scoring system of the three Yes/No tests used in the current study is far from being complicated. Yes responses to real words is calculated to represent a participant's raw score, and yes responses to non-words are false alarms. The false alarms result in a reduction in the participant's total score. The scoring matrix of Yes/No tests is presented in Table 1.

Table 1
Matrix of Possible Responses in XK-Lex, Arabic Lex, and AVST, Where UPPER CASE = Correct Responses

	Word	Non-word
Yes	HIT	False alarm
No	Miss	CORRECT REJECTION

Procedure

Participants were tested on two consecutive days at each institution to avoid testing fatigue. After instructions were delivered, participants were first presented with the two general vocabulary size measures (Arabic-Lex and XK-Lex) followed by the academic vocabulary size test (AVST), which was administered after a short break. On the second day, the non-verbal IQ test was delivered to the same participants. All the testing procedures were performed with help of volunteer lecturers at each institution.

Results

Correlation Analysis

In this study four predictor variables of academic success were used (XK-Lex as a measure of general English vocabulary size, Arabic-Lex as a measure of Arabic vocabulary size, AVST as a measure of English academic vocabulary knowledge, and SPM as a measure of non-verbal IQ). The descriptive statistics of the four predictor variables are shown in Table 2.

Table 2
Descriptive Statistics for Four Variables (IQ , Arabic-Lex, XK-Lex, and AVST)

	N	Minimum	Maximum	Mean	Std. Deviation
IQ	96	12	39	28.25	5.55
Arabic_Lex	96	8500	43000	30843.75	7784.03
XK_Lex	96	1100	6400	3125.00	1310.40
AVST	96	20	470	171.48	86.62

Table 3
Correlation between Variables in The Study

	IQ	Arabic_Lex	XK_Lex	AVST	GPA
IQ	-	.501**	.340**	.411**	.469**
Arabic_Lex		-	.512**	.446**	.590**
XK_Lex			-	.782**	.683**
AVST				-	.728**
GPA					-

Note. ** = Correlation is significant at the 0.01 level.

In order to examine research question 1 (i.e., the relationship between the four measures in our study and academic achievement), correlational analysis was conducted between the observed scores from the four measures and students' GPA. The correlation of the four predictor variables with GPA is shown in Table 3. All predictor variables appear to correlate significantly with GPA. This indicates the validity of these measures since they had all been identified on theoretical grounds to be related to academic success, although most of the previous studies have examined their relationship with academic success individually. XK-Lex and AVST scores show the strongest correlation with GPA, followed by Arabic-Lex and IQ. Also, the strongest correlation between the predictor variables is reported between XK-Lex and AVST (r = .782), which may indicate that a test of academic vocabulary (AVST) resembles very strongly a test of overall vocabulary size (XK-Lex) as suggested in Masrai and Milton (forthcoming). The correlation matrix reported in Table 3 provides a preliminary indication of the effect size (ES) of each independent variable on the dependent variable (GPA). However, to examine the ES in more depth, partial Eta Squared was calculated for each predictor variable.

Analysis of variance showed large ES of L2 general and academic vocabulary knowledge on academic achievement ($F(31, 64) = 6.79$, $p = .001$, $\eta_p^2 = .77$; $F(35, 60) = 63.54$, $p = .001$, $\eta_p^2 = .97$, respectively). The other two variables (Arabic-Lex scores and IQ scores) were also found to explain some levels of ES on academic performance, but to a lesser extent ($F(27, 68) = 5.75$, $p = .001$, $\eta_p^2 = .69$; $F(18, 77) = 3.31$, $p = .001$, $\eta_p^2 = .44$, respectively). Although correlation analysis and ES measures provide insight into how well independent variables relate to the dependent variable, a more detailed analysis is needed to gain further understanding of the predictive power of the different predictor variables. Thus, regression analysis was performed to calculate the explanatory power of the variables individually and combined.

Regression Analysis

Since some levels of high inter-correlations were observed between the predictor variables, multicollinearity diagnostics were preformed prior to reporting regression analysis. Result shows no indication for multicollinearity (all values for tolerance were > .02 and all values for VIF were < 5).

To examine research question 2 (i.e., the predictive power of scores on XK-Lex, Arabic-Lex, AVST, and IQ measures for students' academic achievement measured by GPA), regression analysis was performed.

Table 4

Explained Variance in The Regression Model Predicting GPA with The Four Measures Combined

Model	R	R^2	SE
1	.80[a]	.64	.49

Note. a. Predictors: (Constant), IQ, XK_Lex, Arabic_Lex, AVST.

First, a multiple regression was carried out with GPA as the dependent variable and XK-Lex, Arabic-Lex, AVST and IQ as independent variables, using Enter method. This led to a significant model ($F(4, 91) = 39.675$, $p < .001$) which explains about 64% of the variance in students' GPA.

However, we are also interested in the individual contribution of each predictor variable towards the predictive power of the regression model. To examine this, multiple regressions were carried out to compute the effect of each variable individually. The models summary is reported in Table 5.

Table 5

Explained Variance in The Regression Model Predicting GPA with Each of The Four Measures

Predictor	Model	R	R^2	SE
XK_Lex	1	.68	.47	.58
Arabic_Lex	2	.59	.35	.64
AVST	3	.73	.53	.55
IQ	4	.47	.22	.71

As shown in Table 5, each variable can explain variance in students' success. XK-lex and AVST scores explain the greatest variance in students' GPA, ($R^2 = .47$ and .53, respectively). The other variables, Arabic-Lex and IQ, also explain substantial amounts of variance in the students' achievement ($R^2 = .35$ and .22, respectively).

To further examine the explanatory power, we carried out hierarchical regression with L2 vocabulary size measures (XK-lex and AVST) in block 1, and L1 vocabulary size measure (Arabic-Lex) and IQ in block 2. Since XK-lex and AVST are the best predictors there is a danger the contribution of the less well correlated predictors, Arabic-Lex and IQ, will be lost in a combined model. Dividing the factors this way allows the contribution of these other, less well correlated variables, factors to the model, to be estimated. The result is shown in Table 6.

Table 6
Hierarchical Regression Models

				Change Statistics				
Model	R	R^2	SE	R^2 Change	F Change	df1	df2	Sig. F Change
1	.75[a]	.56	.53	.56	59.92	2	93	.000
2	.80[b]	.64	.49	.07	9.05	2	91	.000

Note. a. Predictors: (Constant), AVST, XK_Lex; b. Predictors: (Constant), AVST, XK_Lex, IQ, Arabic_Lex.

The variables in block 1 produce a significant model ($F(2, 93) = 59.92$, $p < 001$) which predicts about 56% of the variance in GPA and this is substantial. The other variables in block 2, however, can still be shown to contribute marginally to the predictive power of the regression model. The addition to R^2 is still significant ($F(2, 91) = 9.05$, $p < 001$) and these two factors appear to explain an additional 7% of the variance in GPA. These results indicate that when general L2 vocabulary knowledge (measured with XK-Lex) and L2 academic vocabulary knowledge (measured with AVST) are combined they can have a very strong positive effect on learners' performance when studying through the medium of English but that the predictive power of other factors can still improve on this result.

To provide an answer to research question 3 (i.e., whether L2 general vocabulary knowledge can explain a unique variance in academic success) we had to control first for the scores from AVST, the academic vocabulary knowledge test, as these two variables were highly correlated. Interpretation of this strong correlation will be provided in the discussion section of the study. However, to measure if a unique predictive power can be explained by L2 general vocabulary per se, a stepwise regression model was generated including the R^2 change, but with AVST scores removed from the model. The result is summarised in Table 7.

Table 7
Predictive Power of General L2 Vocabulary When Academic Vocabulary Is Controlled for

				Change Statistics				
Model	R	R^2	SE	R^2 Change	F Change	df1	df2	Sig. F Change
1	.68[a]	.47	.58	.47	82.35	1	94	.000
2	.74[b]	.55	.54	.08	16.07	1	93	.000
3	.75[c]	.57	.53	.02	4.89	1	92	.030

Note. a. Predictors: (Constant), XK_Lex; b. XK_Lex, Arabic_Lex; c. XK_Lex, Arabic_Lex, IQ.

The result in Table 7 shows a significant unique contribution of L2 general vocabulary knowledge in explaining academic success. The R^2 of .47, explaining about 47% of variance, therefore, has already been shown in Table 5. But the two other factors are able to enhance this and, combined, add a further 10% to the explanation of variance in GPA scores.

Factor Analysis

Factor analysis was run in an attempt to provide an answer to research question 4 (i.e., examining whether different factors can be discerned in the four sets of results). The factor analysis results are summarised in the Scree plot in Figure 1 and the component matrix in Table 8.

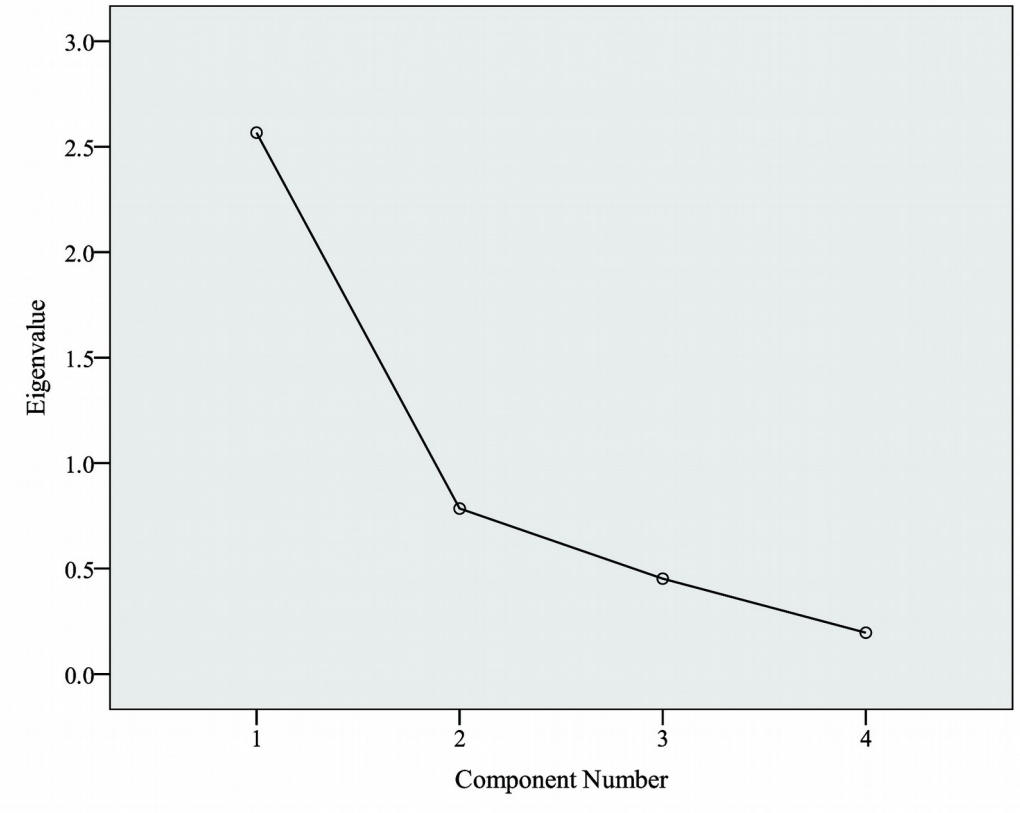

Figure 1. Scree Plot from The Four Sets of Results

Table 8
Component Matrix from The Four Sets of Data

	Component
	1
XK_Lex	.855
AVST	.854
Arabic_Lex	.767
IQ	.679

Note. Extraction Method: Principal Component Analysis; a = 1 Components Extracted.

There appears to be only one component extracted with an Eigen value above 1 and it is concluded that the four variables examined in this study are measuring the same construct.

Discussion

In this study, the contribution of four variables were investigated to assess their impact on students' academic performance measured in GPA. These variables were L2 general vocabulary knowledge, L2 academic vocabulary knowledge, L1 vocabulary knowledge, and non-verbal IQ. While the predictive power of these variables on academic achievement is widely reported in the literature, we investigated their power to predict academic performance among Arabic university students incorporating the four variables in one experimental

setting. The study also aimed at finding out whether including academic vocabulary knowledge among other factors will explain a unique variance and remain the greatest contributing factor towards students' academic success.

Research Question 1: The Relationship between the Four Measures in This Study

In answer to research question 1, all the measures show statistically significant correlations with students' academic performance, as measured by GPA (see Table 2). This finding is broadly in line with what is reported in the literature (e.g., Alderson 2005; Milton & Treffers-Daller, 2013; Laidra et al., 2007; Townsend et al., 2012). The strongest correlation (r = .728) is between L2 academic vocabulary knowledge and GPA. The correlation between L2 general vocabulary knowledge and GPA is moderate to strong (r = .683). The other two factors, L1 vocabulary knowledge, and non-verbal IQ, also display moderate correlations with GPA, which are less strong than the two L2 vocabulary knowledge factors. Since all four test variables correlate moderately to strongly with GPA, it should not be a surprise that they also correlate moderately with each other as is shown in Table 3. There is a particularly strong correlation between L2 general vocabulary knowledge and L2 academic vocabulary knowledge (r = .782). The way that the L2 academic vocabulary knowledge test is likely also to test general vocabulary knowledge has already been suggested in the second sub-section of the literature review. The L2 academic vocabulary knowledge test is based on the AWL and these words occur in general frequency lists spread across the most frequent bands. Good correlations should therefore be expected between any test based on the AWL and any well-formed general vocabulary size test, as is noted in Masrai and Milton (forthcoming). Although the four measures show significant correlations with GPA, multiple regression analyses were required to quantify the effect of each measure on academic performance.

Research Question 2: The Predictive Power of The Variables

The strength of the correlations between the four test variables and GPA means each of the variables can explain some levels of variance in students' GPAs, and this is shown in Table 5. In line with the strength of the correlations, L2 academic vocabulary knowledge and L2 general vocabulary knowledge explain the largest variance in GPA scores (about 53% and 47%, respectively). The least well correlated variable, IQ, still explains a substantial amount of variance, about 22%. These findings fit with the results of other studies reported in the literature. The predictive power of L2 vocabulary knowledge in explaining variance in GPA scores is noted in Townsend et al.'s (2012), Roche and Harrington (2013), and Daller and Yixin (2016). The results from this study suggest that vocabulary has a particularly high predictive power. In Daller and Yixin's study, for example, vocabulary knowledge explains about 21% of variance in GPA, and in Townsend et al.'s study the general vocabulary factor explained between 26% and 43% of variance. In this study the two L2 factors explain about 56%. The findings from this study also agree with other studies with regard to the impact of IQ on academic achievement. Rohde and Thompson (2007), for example, suggest that although IQ can influence academic achievement, about 51% to 75% of the variance in academic achievement is not accounted for by the measures of general cognitive ability *per se*. Thus, the current study suggests that possessing larger L2 vocabularies, both general and academic, can have a major impact on students' academic achievement, possibly greater in its effect than the academic ability of the learner.

The high predictiveness of L2 language factors beyond that reported by others, for example Townsend et al. (2012), is capable of various interpretations. It is possible that the language problems for the Arabic speaking learners in this study are sufficiently great in some cases to affect strongly the knowledge and understanding of the subject matter they have, greater than in the Townsend et al.' study. We have no baseline language knowledge scores to allow us to compare the language levels of the learners in the two studies but if the learners in this study were to include a tail of students with particularly weak English then this might explain the high correlations seen here. But these differences may equally mean that the examiners in this study are placing a high emphasis on language accuracy in the scores they give for academic achievement. GPA is not a well-defined construct and variation in what marks are awarded for is bound to occur.

A multiple regression, Table 4, indicates that the four variables combined can explain 64% of the variance in GPA scores. This combined result is greater than the two factor model investigated in Townsend et al. (2012), which examined only L2 language knowledge factors. This suggests that the variables examined in this study would be particularly useful in a practical setting where, for example, university and school teachers need to anticipate which of their students are at risk of low academic performance and are in need to support in their academic studies.

The further regression analyses carried out in this study are designed to examine the way the variables interact with each other, to better understand in what proportions these variables combine in their interactions with GPA. The strongest predictors among the four variables are the L2 vocabulary factors, academic vocabulary knowledge and general vocabulary size. It has been indicated above that these two tests may be testing a single factor and the hierarchical regression reported in Table 6 has therefore been carried out to separate out the L2 language factor from the potential contribution of the other variables in gaining good GPA scores. The results suggest that L1 Arabic vocabulary size and IQ combined can add slightly more than 7% to the predictiveness of the L2 language factor. The 56% of variance explained by general and academic vocabulary knowledge rises to 64% once IQ and L1 vocabulary size are added in (note there is some rounding of numbers in Table 6). The regression analysis summarised in Table 7 separates out the contribution of IQ and L1 vocabulary size and for this analysis scores from the L2 academic vocabulary size variable have been omitted because of their co-linearity with general vocabulary size, and the better to examine the effect of the other factors. The results in Table 7 suggest that both IQ and L1 vocabulary can make separate and unique contributions to the predictive ability of the model and with this combination of variables, L1 vocabulary size appears to add some 8% to the, 47% of variance explained by L2 general vocabulary size. IQ appears to add a further 2%. This last figure need not contradict the suggestion of Rohde & Thompson (2007) where between about 25% and 50% of variance in academic achievement can be explained by IQ alone since studies in the effect of IQ rarely include in their models the powerful effects of L2 vocabulary as measured with the sophistication of the most recent L2 tests.

Research Question 3: The Relationship between General and Academic Vocabulary Knowledge

The co-linearity of L2 vocabulary factors, L2 general vocabulary size and L2 academic vocabulary knowledge, has been noted above and has raised the question whether academic vocabulary knowledge is capable of making a unique contribution to the variance in GPA scores, over and above the impact of general vocabulary size. In this study, as distinct from the Townsend et al.'s (2012), it is the L2 academic vocabulary knowledge test which is the best individual predictor among the four variables, slightly better than L2 general vocabulary size. The difference between the results might largely be attributed to the academic word measures used in both studies. Townsend et al. (2012) used the academic part of the revised version of Vocabulary Levels Test (VLT) (Schmitt, Schmitt, & Clapham, 2001) which includes only 30 words sample of the 570 AWL (Coxhead, 2000). The low sampling rate and also the problematic sampling technique (see Schmitt et al., 2001) of this part of the VLT might explain, in part, why the predictive power of the test scores is lower than for their general vocabulary measure. On the other hand, the test used in the current study (AVST) is thought to produce more credible scores, as it features a high sampling rate (1:5) and is based on frequency selection of its items (Masrai & Milton, forthcoming).

Nonetheless, this result mirrors the findings reported in Masrai and Milton (forthcoming). This suggests that while the results produced by the L2 academic vocabulary knowledge test must include L2 general vocabulary size (its construction using words drawn from across the most frequent general vocabulary bands means it cannot avoid this), the two types of knowledge can nonetheless still be differentiated. Our best interpretation of the data is that L2 general vocabulary size is crucial to academic achievement and that academic vocabulary knowledge will add marginally to this. It appears that knowledge of the AWL specifically can add an additional 7% to L2 general vocabulary knowledge in explaining variance in GPA. This conclusion is strikingly similar to the results obtained in the studies by Townsend et al. (2012) and Masrai and Milton

(forthcoming) and similar too to other studies (e.g., Harington & Roche, 2014; Milton & Treffers-Daller, 2013; Saville-Troike, 1984).

Research Question 4. How Many Separate Factors Can Be Identified in These Variables?

One argument used in Masrai and Milton (forthcoming) to suggest that a test based on the AWL is likely to function also as a general vocabulary size test, is that when scores for the two different tests were subjected to factor analysis, only one component could be identified leading to the conclusion that they were testing the same construct. Figure 1 and Table 8 report the results of factor analysis with the four sets of data obtained in this study. In line with the earlier study, the results here also suggest that the two L2 factors, L2 general vocabulary size and L2 academic knowledge, are part of the same component. But the results in Figure 1 and Table 8 also suggest that the other two variables investigated in the study, IQ and L1 vocabulary size, are included in the same component and are also, in some way, measuring the same construct.

Perhaps it should not be a surprise if all three of the language related variables form part of the same component. L1 and L2 vocabulary size have been demonstrated to correlate closely among native Arabic speakers who use English as a foreign language (e.g., Masrai, 2015). But there are suggestions at the level of theory too, for example Cummins' Common Underlying Proficiency ideas (Cummins, 2000), that L1 and L2 vocabulary size should be related. There may be a general language ability factor at play here. However, it is not so clear why IQ scores should form part of the same factor. The tests used in this study have been deliberately chosen to be non-verbal assessments with the intention that this would avoid potential interference from language knowledge and ability. The tests used are abstract reasoning tasks which involve completing a pattern or figure with a part missing, by choosing the correct missing piece from among six alternatives. However, there is some evidence that these types of reasoning task can function well in predicting language learning aptitude in young children (Milton & Alexiou, 2006), and may pick up on the ability to infer rules and structures in language. It must be noted that all four variables correlate quite strongly with each other and there is a long-standing tradition that a wide variety of variables can all fall under a single general intelligence factor as in Spearman's G factor (Spearman, 1927). Nonetheless, this idea that the four variables may all be part of a single factor need not detract from the evidence of the regression analyses which suggests the four variables investigated here interact with academic performance as measured by GPA in slightly different ways and that a unique contribution to GPA for each of them can be found.

Conclusions

The attempt to use several factors to predict and explain academic performance has produced results which are very encouraging. The combined model of four variables in this study can predict nearly two-thirds of variance in academic performance as measured by GPA, stronger than any individual factor. This suggests greater predictiveness than most other studies even where several factors are combined in a predictive model (e.g. Townsend et al, 2012; Daller & Yixin, 2016; and Roche & Harrington, 2013). This may be the result of the particular circumstances of the learners and staff who provided the marks for GPA, involved in this study. The bulk of the explanatory power is provided by L2 knowledge factors but the regression analyses suggest it is possible to identify a unique, if sometimes marginal, contribution to variance in GPA scores for all the factors investigated here.

The L2 general vocabulary and L2 academic vocabulary scores are strongly correlated and it is difficult to decide how independently these two variables function. Our best interpretation of the results is to confirm Townsend et al.'s (2012) conclusion that knowledge of academic words provides some unique, albeit marginal, variance to general academic success as measured by GPA, in addition to general vocabulary size. A focus on the AWL in teaching, within this interpretation, appears a useful element of any English for academic purposes course provided it is implemented within the context of an overall programme of vocabulary development for learners to reach the size of lexicon necessary for fluent language use.

The factor analysis suggests all the variables here are closely related, and here our best interpretation is that there may be a general language ability factor at play which is linked to other factors, identified in other studies, like IQ. Even though these factors appear closely related, the use of multiple tests in combination appears to have potential for identifying learners at risk of academic failure. It may be possible to provide language support for students at risk. The prominence of L2 vocabulary knowledge in predicting academic success suggests that a wider use of vocabulary size tests specifically, in the acceptance process for learners at school or university, could help improve the selection process and ensure those entering education and studying through the medium of English as a foreign language have the skills to succeed academically.

While these results are encouraging, it must also be noted that this is a single study, drawing learners from a homogenous L1 Arabic speaking background, with results drawn from two institutions in Saudi Arabia. Further research is needed with larger samples, learners from different L1s, and including groups from different disciplines, to confirm the idea that combinations of factors can usefully predict students' academic attainment.

References

Alderson, J.C. (2005). *Diagnosing foreign language proficiency: The interface between learning and assessment*. London: Bloomsbury.

Astika, G. (1993). Analytical assessments of foreign students' writing. *RELC Journal*, *24*, 61–70.

Beglar, D., & Hunt, A. (1999). Revising and validating the 2000 word level and university word level vocabulary tests. *Language Testing*, *16*(2), 131–162.

Chen, Q., & Ge, C. (2007). A corpus-based lexical study on frequency and distribution of Coxhead's AWL word families in medical research articles. *English for Specific Purposes*, *26*, 502–514.

Chung, M., & Nation, P. (2003). Technical vocabulary in specialized texts. *Reading in a Foreign Language*, *15*, 103–116.

Cobb, T., & Horst, M. (2004). Is there room for an AWL in French? In B. Laufer & P. Bogaards (Eds.), *Vocabulary in a second language: Selection, acquisition, and testing* (pp. 15-38). Amsterdam: John Benjamins.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34*, 213-238.

Cummins, J. (2000). *Language, power and pedagogy: bilingual children in the crossfire.* Clevedon: Multilingual Matters.

Daller, M. H. & Phelan, D. (2013). Predicting international student study success. *Applied Linguistics Review*, *4*(1). 173–193.

Daller, M., & Yixin, W. (2016). Predicting study success of international students. *Applied Linguistics Review,* (ahead of print).

Harrington, M., & Carey, M. (2009). The online yes/no test as a placement tool. *System, 37*(4), 614–626.

Harrington, M. & Roche, T. (2014). Identifying academically at-risk students in an English-as-a-Lingua-Franca university setting. *Journal of English for Academic Purposes, 15*, 37–47.

Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.

Laidra, K., Pullmann, H., & Allik, J. (2007). Personality and intelligence as predictors of academic achievement: A cross-sectional study from elementary to secondary school. *Personality and Individual Differences, 42*(3), 441-451.

Laufer, B. (1992). How much lexis is necessary for reading comprehension? In H. Bejoint & P. Arnaud (Eds.), *Vocabulary and applied linguistics* (pp. 126-132). London: Macmillan.

Laufer, B. (1998). The development of passive and active vocabulary in a second language: Same or different? *Applied Linguistics, 19*(2), 255-271.

Laufer, B. & Ravenhorst-Kalovski, G. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language, 22*(1), 15-30.

Lesaux, N., Kieffer, M., Faller, S. E., & Kelley, J. G. (2010). The effectiveness and ease of implementation of an academic vocabulary intervention for linguistically diverse students in urban middle schools. *Reading Research Quarterly*, *45*, 196–228.

Masrai, A. (2015). *Investigating and explaining the relationship between L1 mental lexicon size and organisation and L2 vocabulary development*. Unpublished PhD thesis. Swansea University, UK.

Masrai, A., & Milton, J. (2012). The vocabulary knowledge of university students in Saudi Arabia. *TESOL Arabia Perspectives, 19*(3), 13-20.

Masrai, A., & Milton, J. (2017). How many words do you need to speak Arabic? An Arabic vocabulary size test. *Language learning Journal*, (ahead of print).

Masrai, A., & Milton, J. (forthcoming). Measuring the contribution of academic and general vocabulary knowledge to learners' academic achievement. *Journal of English for Academic Purposes*.

Masrai, A., & Milton, J. (forthcoming). Frequency distribution of the words in AWL in BNC and BNC/COCA.

Milton, J. (2009). *Measuring second language vocabulary acquisition*. Bristol, UK: Multilingual Matters.

Milton, J. (2013). Measuring the contribution of vocabulary knowledge to proficiency in the four skills. In C. Bardel, C. Lindqvist & B. Laufer (Eds.), *L2 vocabulary acquisition, knowledge and use: New perspectives on assessment and corpus analysis* (pp. 57-78): EUROSLA monograph 2.

Milton, J., & Alexiou, T. (2006). What makes a good young language learner? In A. Kavadia, M. Joanopoulou & A. Tsaggalidis (Eds.), *New directions in applied linguistics* (636-646). Thessaloniki, Greece: Aristotle University of Thessaloniki.

Milton, J. and Fitzpatrick, T. (eds.). (2014). *Dimensions of vocabulary knowledge*. Basingstoke, UK: Palgrave Macmillan.

Milton, J., & Treffers-Daller, J. (2013). Vocabulary size revisited: The link between vocabulary size and academic achievement. *Applied Linguistics Review, 4*(1), 151–172.

Milton, J., Wade, J., & Hopkins, N. (2010). Aural word recognition and oral competence in a foreign language. In R. Chacón-Beltrán, C. Abello-Contesse & M. Torreblanca-López (Eds.), *Further insights into non-native vocabulary teaching and learning* (pp. 83-98). Bristol: Multilingual Matters.

Mochida, A., & Harrington, M. (2006). The yes/no test as a measure of receptive vocabulary knowledge. *Language Testing*, *23*(1), 73–98.

Mudraya, O. (2006). Engineering English: A lexical frequency instructional model. *English for Specific Purposes*, *25*, 235–256.

Nagy, W., & Townsend, D. (2012). Words as tools: Learning academic vocabulary as language acquisition. *Reading Research Quarterly, 47*(1), 91-108.

Nation, I.S.P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Nation, I.S.P. (2004). A study of the most frequent word families in the British National Corpus. In Bogaards P. and Laufer B. (Eds), *Vocabulary in a second language: Selection, acquisition and testing* (pp. 3-13). Amsterdam: John Benjamins.

Nation, I.S.P. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review*, *63*(1), 59–81.

Neisser, U., Boodoo, G., Bouchard, T. J., Boykin, A. W., Brody, N., Cesi, S. J., et al. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, *51*, 77–101.

Qian, D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *Canadian Modern Language Review, 56*, 282–308.

Raven, J., Raven, J. C., & Court, J. H. (1998). Manual for Raven's Progressive Matrices and vocabulary scales: Section 4 Advanced Progressive Matrices sets I and II, 1998 ed. Oxford, UK: Oxford Psychologists Press Ltd.

Read, J. (2000). *Assessing vocabulary*. Cambridge: 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*(1), 1-13.

Rohde, T. E., & Thompson, L. A. (2007). Predicting academic achievement with cognitive ability. *Intelligence, 35*(1), 83-92.

Saville-Troike, M. (1984). What really matters in second language learning for academic achievement? *TESOL Quarterly*, *18*(2), 199–219.
Schmitt, N., & Schmitt, D. (2014). A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching, 47*(4), 484-503.
Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. *Language Testing*, *18*(1), 55–88.
Spearman, C. (1927). *The abilities of man.* Oxford: Macmillan
Stæhr, L. S. (2008). Vocabulary size and the skills of listening, reading and writing. *Language Learning Journal*, *36*(2), 139–152.
Townsend, D., Filippini, A., Collins, P., & Biancarosa, G. (2012). Evidence for the importance of academic word knowledge for the academic achievement of diverse middle school students. *The Elementary School Journal, 112*(3), 497-518.
West, M. (1953). *A general service list of English words*. London: Longman.
Zimmerman, K. (2004). *The role of vocabulary size in assessing second language proficiency.* Unpublished MA thesis. Brigham Young University.

About the Authors

Ahmed Masrai is Assistant Professor of Applied Linguistics at King Abdulaziz Academy, Saudi Arabia. He holds a PhD degree in Applied Linguistics from Swansea University, UK. His research interest includes second language acquisition, lexical processing, and vocabulary assessment. This interest has led to a number of publications in these areas.

James Milton is Professor of Applied Linguistics at Swansea University, UK. He worked in Nigeria and in Libya before coming to Swansea in 1985. A long-term interest in measuring lexical breadth and establishing normative data for learning has produced extensive publications including *Measuring Second Language Vocabulary Acquisition* (Multilingual Matters, 2009).