

Revisiting Language Threshold Hypothesis in Iranian EAP Context in Light of Content Specificity

¹Nazila Naghipoor

²Farhad Mazlum*

³Mostafa Janebi Enayat

Research Paper

IJEAP- 2402-2026 DOR: [20.1001.1.24763187.2024.13.1.5.0](https://doi.org/20.1001.1.24763187.2024.13.1.5.0)

Received: 2024-01-14

Accepted: 2024-03-10

Published: 2024-03-30

Abstract: EAP students' performance on reading comprehension tests is affected by English proficiency and discipline-specific background knowledge. Proponents of the Language Threshold Hypothesis (LTH) propose that the contributory effects of background knowledge are bound to two language thresholds: high and low proficiency levels; for high and low proficiency students, it does not make difference whether they take reading tests from inside or outside of their fields. On the other hand, it is suggested that the predictive power of the Hypothesis might fluctuate due to content specificity of reading texts. The study aims to, firstly, investigate the predictive potential of the LTH in the Iranian EAP context, and, secondly, to examine if the effects of language proficiency and background knowledge fluctuates with content specificity. Following convenience sampling, one-hundred sixty graduate and undergraduate students of biology and psychology took five tests, including an English proficiency test, two knowledge tests, and two cloze tests for psychology and biology, with each containing three paragraphs of varying levels of content specificity. A three-way factorial ANOVA indicated that the LTH was partially supported; no evidence was found for high proficiency threshold and partial evidence was observed for the lower threshold. It was also found that the contributory effect of English proficiency on Iranian students' reading comprehension significantly fluctuates with changes in content specificity of reading tests; as content specificity increased, high and low proficiency students' performances on field-related and field-unrelated tests became poorer and with a decrease in content specificity their performances improved. The implications of the study are discussed.

Keywords: Background Knowledge, Content Specificity, English Proficiency, Reading Comprehension

Introduction

Once criticized for lacking a solid and substantiated theoretical underpinning, assessing English for Specific Purposes (ESP) and, by extension, English for Academic Purposes (EAP) nowadays is an established territory in language assessment in general and ESP/EAP in particular. Historically speaking, theoretical sophistication of ESP/EAP assessment was partially triggered by research attempts concerning high-stakes, large scale EAP tests (e.g. TOEFL and IELTS) administered to non-English students studying in English-medium universities. Such attempts were associated with controversies regarding the extent different factors (i.e., English proficiency vs. discipline-related topical knowledge) contribute to test-takers' performances on different modules of the tests (Clapham, 2000). This was a theoretical challenge with practical consequences since the choice of reading texts in international EAP tests, for instance, could favor disciplinarily close test-takers and disfavor distant ones (Darabi Bazvand,

¹ MA graduate of TEFL, nazilanaghipoor@gmail.com; English Department, Faculty of Humanities, University of Maragheh, Maragheh, Iran.

² Assistant Professor of Applied Linguistics (Corresponding Author), fmazlum@maragheh.ac.ir; English Department, Faculty of Humanities, University of Maragheh, Maragheh, Iran.

³ Assistant Professor of Applied Linguistics, m.enayat@maragheh.ac.ir; English Department, Faculty of Humanities, University of Maragheh, Maragheh, Iran.

2019). Some decades ago, this was a contentious issue as some scholars argued for more field-wise modularization of such EAP tests whilst others maintained that since English proficiency alone explains a large share of the variance in tests purportedly targeting both language and content knowledge, IELTS and TEEP officials do not need to design separate and discipline specific tests for different EAP test takers with different educational backgrounds (Smith et al., 2021).

Since then, the issue, i.e., contributory effects of English proficiency and background content knowledge on EAP test-takers' performances on different reading comprehension (RC) tests, has been studied extensively and there are currently several key issues clustering around it. For example, some research findings suggest that the influence of background knowledge on EAP students' scores on different RC tests depends on students' English proficiency level; for low and advanced level students, it does not make difference whether they take RC tests from inside or outside their academic fields while it is the opposite for intermediate level test-takers. The argument, rooted in LTH, has engendered a continuing line of inquiry the concomitant of which has been our more in-depth understanding of the complexities involved, on a positive note, and, less promisingly, an inconclusive picture of how English proficiency and disciplinary content knowledge influence EAP test-takers' RC scores (Ridgway, 1997; Krekeler, 2006). The potential causes of such inconclusiveness are attributed to factors such as RC test types following EAP reading tests (Taghizadeh Vahed & Alavi, 2019), different measures of topic familiarity (Krekeler, 2006), and textual features such as readability and genre of reading texts (Tarhani-Aliabad et al., 2022), to name a few.

Investigating whether the effects of content knowledge differ based on the level of English proficiency remains an ongoing line of research in EAP (Cai & Kunnan, 2019). One potential explanation of such a continuing interest lies with how EAP construct is defined. Douglas (2013), for instance, proposes that, unlike other non-specific tests in which topical knowledge is regarded as an extraneous variable the effect of which should be decreased, in ESP/EAP, background knowledge and English proficiency are inextricably intertwined and make up a composite construct. His argument, known as Inseparability Hypothesis in ESP/EAP assessment literature, implies that two knowledge sources, i.e., discipline-based background knowledge and English proficiency, will be competing in exerting effects on EAP test-takers' performances. Therefore, researchers are interested to understand how such factors jointly influence students' scores on not only reading comprehension tests but other tests of different language skills. The LTH, albeit contentious and still open to investigation, has been advanced to explain how the facilitative role of background knowledge can be explained vis-à-vis EAP test-takers' English proficiency level.

This study attempts to, first, examine the predictive potential of the LTH with Iranian EAP students. More specifically, it is intended to see if the facilitative role of Iranian EAP students' background knowledge is bound to linguistic thresholds suggested by the LTH when they take different RC tests. This is of theoretical and pedagogical relevance. Theoretically, findings can contribute to the current debates pertaining to the LTH and its explanatory power in EAP assessment in general. Pedagogically, and in line with Krekeler's argument (2006), the interpretation of EAP students' test performance depends on whether such language thresholds exist or not. If research evidence endorses the explanatory potential of the LTH, test users will interpret low and high proficiency students' performances differently, as predicted by the LTH, and this, in turn, might affect EAP test construction. Secondly, and probably more importantly, the study attempts to investigate if with changes in content specificity of texts the effect of English proficiency on RC EAP tests fluctuates. As predicted by the LTH, low-proficient test-takers, for example, are likely to perform equally poorly on disciplinarily neighboring and non-neighboring RC tests since they are too concentrated on low level, bottom-up text decoding processes; however, such LTH-informed predictions might be prone to fluctuations due to several features one of which is content specificity. The rationale to focus on content specificity as a variable in this study is rooted in prior research studies ending with the question 'Can the predictive power of the LTH fluctuate with changes in content specificity of texts in RC EAP tests?' The review below paves the ground further for the current work.

Literature Review

Theoretical Framework

As noted above, the interest to see if the impact of background knowledge on EAP tests of RC can be predicted by English proficiency level, has been an ongoing line of research in EAP. Earlier scholars such as Clapham (2000) proposed that the LTH can be used to make the prediction. The LTH is rooted in *Short Circuit Hypothesis* suggested by Clarke (1980); an attempt to explain if and how first language reading strategies transfer to reading in a second language. Simply put, Clarke argued that such an advantageous transfer is feasible if students' proficiency in L2 reaches a threshold level, otherwise their available L1 knowledge is obstructed or short circuited. Earlier scholars in EAP extended Clarke's argument to EAP assessment and held that background knowledge is facilitative when test-takers reach a certain L2 proficiency threshold. Later findings suggested there exist two such thresholds: the effect of background knowledge in second language reading is blocked for students with *lower* L2 proficiency levels and unnecessary for *advanced* level students. This is so because the first group is "too concerned with bottom-up skills such as decoding the text" (Clapham, 2000, p. 515) to benefit from their disciplinary knowledge and the second group takes maximum advantage of their L2 proficiency so that the need to background knowledge is circumvented. For medium level readers, however, background knowledge is a facilitative factor (Tarlani-Aliabad et al., 2022).

Since then, several researchers have examined the predictive potential of the LTH in EAP reading tests. Ridgway (1997), for example, gave advanced and low proficiency students texts from outside and inside their own academic discipline to see if LTH predictions were tenable. He found that students with low L2 proficiency levels did equally poorly on field-related and field-unrelated RC texts—evidence for endorsing the LTH—while advanced level students did significantly better on texts from their own subject area. He, therefore, concluded that the lower threshold of the LTH may exist but the upper level may not. Clapham (1996) did a similar study and, using students' grammar scores as an indicator of L2 proficiency, divided students to three groups (low, intermediate, and advanced). Similar to Ridgway's study, Clapham's initial analysis indicated that for low proficiency students the means on different RC tests (Business Studies & Social Sciences vs. Medical & Life Sciences) were not statistically significant while they were different for medium and advanced level students. While this initial analysis seemed to be partially in line with the LTH, further detailed analyses on texts and students from a third subject area (Physics) showed that the effect of background knowledge on academic reading comprehension was inconsistent. Content specificity of texts was speculated to be a potential cause for such inconsistencies.

Krekeler (2006) employed C-test to test language proficiency and administered field-related and field-unrelated RC tests to his participants. Similar to prior studies, he found that low proficiency EAP test takers did similarly on disciplinarily remote and close reading tests. He further noticed that students with high proficiency levels were not heavily dependent on their background knowledge. Although such findings might be regarded as empirical evidence advocating the LTH, Krekeler did not interpret his findings as such since his further analyses suggested that most EAP students, irrespective of their English proficiency level, could take advantage of their background knowledge especially when content specificity of reading texts was allegedly low.

Unlike prior studies, and motivated by inconsistencies in research results, Usó-Juan (2006) argued that instead of using categorical approaches to English proficiency and dividing EAP test-takers into three (i.e. low, intermediate, & advanced) or two (high & low) proficiency levels based on already defined scales, it makes more sense to view it as a continuous variable, use regression analysis, and give more accurate information regarding how much of EAP RC test performance is accounted for by English proficiency and background knowledge. Her results suggested that 58% to 68% of students' EAP reading comprehension could be accounted for by English proficiency whereas 21% to 31% could be explained by their subject-related knowledge. Based on her further analyses, Usó-Juan (2006) maintained that her advanced and intermediate level participants could successfully read EAP texts without discipline-related knowledge, but low proficiency test-takers could do so if they made it to a linguistic threshold and had field-related content knowledge.

Local studies have also been conducted to address the issue. Salmani-Nodoushan (2003), for example, examined the effects language proficiency, task type, and text familiarity on Iranian EAP students' performances. Using regression analysis, he concluded that language proficiency, compared to background knowledge, accounts for more of their performances (i.e., not less than 50% of their performance). He also found that with both disciplinarily familiar and unfamiliar texts, low and high proficiency level students performed differently (Salmani-Nodoushan, 2003). In another study, Taghizadeh Vahed and Alavi (2019) gave EAP RC tests to 281 Iranian civil engineering students; two texts (cementitious materials & strength of materials) were field adjacent and the other two (Chomsky's theory of language & the audio-lingual method) were field remote. Their results showed that topical knowledge was not facilitative of low proficiency students' performance whereas the interplay between disciplinary knowledge and language proficiency was significant for intermediate and advanced students. Interestingly, they observed that the influence of English proficiency on EAP test performance varies with task types. Mazlum, Shamameh, and Salimi (2020) used TOEFL to divide 110 English majors into low, intermediate, and advanced students and administered two neighboring (psychology & law) and two non-neighboring (chemistry & geology) EAP tests to see if taking advantage of academic adjacency is bound to any linguistic thresholds. Unlike some previous studies (e.g., Usó-Juan, 2006), for all three proficiency levels, it made a statistical difference whether EAP tests came from close versus distant academic neighborhoods. Finally, Tarlani-Aliabadi et al. (2022) studied the role of language knowledge and background knowledge in ESP reading comprehension tests with varying degrees of text difficulty and concluded that "background knowledge had marked effects on the performances of the participants across the three reading texts" adding that in the most difficult text, "neither general language knowledge nor general background knowledge were strong indicators" (p.1). The three reading texts varied in text difficulty determined on the basis of Flesch Readability Index, number of passive verbs and participles as attributes.

The Current Study

In some of the reviewed studies, content specificity has been speculated to explain the inconsistencies of the findings. Although reference is made to it as a factor that might exert influence on the interaction between subject knowledge and language proficiency when discussing EAP test-takers' performances on RC tests, prior studies tend to control its potential effect by choosing texts that are "thought to be equally accessible for all test-takers" (Krekeler, 2006, p. 122) or "texts that could be understood by any educated layman" (Taghizadeh Vahed & Alavi, 2019, p. 5). Such text selection rationales are often accompanied by reports of readability index or word length to address and control the likely effect of content specificity. However, readability indexes are measures of text difficulty and might not give us a valid basis for content specificity. Additionally, the reliability of using readability for text difficulty has been questioned in many studies (e.g., Benjamin, 2012; Janebi Enayat & Babaii, 2018). Setting *any educated layman's understanding or equal accessibility of texts for all test takers* as criteria to address content specificity is equally contentious as such criteria are prone to open and subjective interpretations. Motivated by Usó-Juan (2006)'s proposal, we argue that content specificity of EAP texts should be negotiated with content specialists. In this study, the choice of such texts is made in close collaboration with subject specialists in content departments. As a result, unlike prior studies in which textual features and quantitative measures (i.e., readability indexes) are used to choose EAP texts, we employ more in-depth qualitative measures to address content specificity to see if the above speculation in LTH studies should be dismissed or retained.

The current study, then, addressed the predictive potential of the LTH with Iranian EAP students with a specific focus on content specificity. More specifically, it is, firstly, intended to see if there are linguistic thresholds to which the impacts of background knowledge on EAP RC test performance are bound. Unlike some prior studies (e.g., Darabi Bazvand, 2019; Tarlani-Aliabadi, et al., 2022) in which the subjects' field-related knowledge was not tested but presumed on the ground of students' field of study, we followed Usó-Juan (2006)'s proposal and, collaborating with subject specialists in content departments, designed knowledge tests to measure it. Secondly, the study

examined whether the contributory effects of subject-related background knowledge and English proficiency would fluctuate with content specificity. The research questions were:

Research Question One: Do English language proficiency and field-related knowledge affect Iranian EAP reading test performance?

Research Question Two: Do the contributory effects of field-related background knowledge and English proficiency on Iranian EAP reading test performance fluctuate with changes in content specificity?

Methodology

Design of the Study

This is a descriptive study in which attempts are made to shed light on the relationships between Iranian EAP students' L2 proficiency and their performances on RC tests coming from inside versus outside their area of study.

Participants

One hundred eighty-one male and female graduate and undergraduate students, with BA, MA, and PhD degrees at different universities of Iran, including Tehran, Tabriz, Semnan, Zanzan, Urmia, Maragheh, Babolshahr, and Saveh took part in the study. They were aged from 22 to 50. Some of the respondents were monolingual with Persian as their native language. Others were bilingual with Azerbaijani Turkish or Kurdish as their first language and Persian as their second language. Twenty-one out of the 181 were kept out of the final data due to reluctance to attend in the next steps of the research. There were 160 students in our final sample. All the participants were given five tests, including one English proficiency test, two Farsi knowledge tests on biology and psychology, and two English content tests in cloze format on biology and psychology. The following table summarizes information pertaining to our participants.

Table 1

Demographic Information of Research Participants

University	Academic degrees			Total
	PhD	MA	BA	
Tehran	2	4	13	19
Tabriz	3	6	20	29
Semnan	1	6	15	22
Zanzan	--	7	12	19
Urmia	3	6	14	23
Maragheh	1	7	17	25
Babolshahr	--	2	10	12
Saveh	--	3	8	11

Instruments

To collect data, the researchers used two content cloze tests, two Farsi knowledge tests, and Vocabulary Level Test (VLT). They also interviewed subject specialist informants (SSIs) to get information for developing appropriate knowledge and cloze tests. An account of each is provided below.

Vocabulary Level Test (VLT)

VLT (2nd version), designed and adapted by Schmitt et al. (2001), was used to measure the respondent's general English ability. Previous studies (e.g., Read, 2009; Akbarian, 2010; Derakhshan & Janebi Enayat, 2020; Janebi Enayat & Derakhshan, 2021; Janebi Enayat et al., 2018) suggest that the VLT can be legitimately used for such purposes. The reliability index was 0.86 using Cronbach's Alpha.

Cloze Tests

Two English cloze tests, one biology (cell division) and the other psychology (autism), were administered as EAP RC tests. There were three paragraphs in each, ordered from less content specific (paragraph 1) to most content specific (paragraph 3). Ordering was based on our interviews and negotiations with content specialists.

Knowledge Tests

Two Farsi background knowledge tests of biology and psychology developed in collaboration with subject specialists were used to assess the participants' content knowledge of cell division and autism. The questions were based on the content of psychology and biology texts used for designing the cloze tests.

Interview

Semi-structured interviews were conducted with nine field specialists (five professors of Biology and four professors of Psychology). The interviewees were given the corresponding texts before the interview to study. During interviews, they were required to order paragraphs in terms of content specificity (from 1, less specific, to 3, more specific) and explain their ordering. High inter-coder agreement (0.9) was observed among field specialists ordering the texts. Extended involvement with the interviewees was employed to cater for credibility and trustworthiness. In other words, the researchers were involved in the process of text ordering asking field specialists to extensively explain the rationales behind their views. Finally, as regards researcher positionality, the researchers believe that content specificity cannot be determined by textual features (e.g. readability) alone and field specialists' views can provide an alternative explanation of the variable.

Procedure

At first, the interviews with field specialists in biology and psychology departments were conducted to select appropriate texts and order them in terms of content specificity. Then cloze and content tests were collaboratively designed. As for test administration, Farsi knowledge tests were given first. To avoid the likely effects of students' remembering the content of the first tests on the next ones, EAP RC cloze tests and the proficiency test (VLT) were administered after a two-week time interval. It should be noted that the participants did not know there would be subsequent EAP RC tests which could further minimize memory effect.

Results

To investigate if content knowledge (discipline-related and discipline un-related knowledge) and English proficiency (low and high) influenced our participants reading comprehension and whether the effects of discipline-related background knowledge and English proficiency interacted with content specificity (three levels), a mixed factorial ANOVA was run. In the analysis of the study, there were 2 within-subjects variables (discipline-relatedness with 2 levels and content specificity with 3 levels) and 2 between-subjects variables (level of language proficiency and major). First, we took the measurements of the difference between the performances of our participants. Then, we subjected them to repeated measures respecting content specificity and discipline-relatedness of the texts. Table 2

provides the list of all the between- and within-subjects variables in our analysis. The descriptive statistics, including mean and standard deviation, for these variables are provided in Table 2.

Table 2*Within- and Between-subjects Variables in the ANOVA Design*

Type	Variable	Levels
Within-subjects	discipline-relatedness	1. discipline-related 2. discipline-unrelated
Within-subjects	content specificity	1. Low content specificity 2. Medium content specificity 3. High content specificity
Between-subjects	English proficiency level	1. Low 2. High
Between-subjects	Major	1. Psychology 2. Biology

Table 3*Descriptive Statistics for the Combinations of Within-subjects and Between-subjects Variables*

	Proficiency	Major	Mean	SD	N
PS1	Low	Psychology	5.500	1.831	32
		Biology	4.200	1.181	40
	High	Psychology	7.833	.694	48
		Biology	6.400	2.085	40
BI1	Low	Psychology	4.500	2.094	32
		Biology	4.800	1.343	40
	High	Psychology	5.666	1.905	48
		Biology	7.000	2.717	40
PS2	Low	Psychology	5.000	3.048	32
		Biology	4.200	1.620	40
	High	Psychology	8.666	2.309	48
		Biology	7.000	2.935	40
BI2	Low	Psychology	3.750	3.079	32
		Biology	3.600	1.373	40
	High	Psychology	6.000	1.166	48
		Biology	5.600	2.529	40
PS3	Low	Psychology	3.750	1.665	32
		Biology	4.200	1.963	40
	High	Psychology	5.000	1.845	48
		Biology	6.600	2.609	40
BI3	Low	Psychology	2.750	1.502	32
		Biology	3.400	.496	40
	High	Psychology	3.166	1.883	48
		Biology	5.200	1.181	40

Note: PS = Psychology; BI = Biology; SD = Standard Deviation

Before conducting the mixed factorial ANOVA, the researchers checked each required ANOVA assumption for each combination of variables. In so doing, we tested plots of normal probability for repeated measures and some considerable deviations were observed. The result of the Levene's test was statistically significant, showing that statistically meaningful differences existed between the variances of the groups. It means that the assumption of homogeneity of variances was violated. To address it, we carried out Mauchly's test. The results were significant ($p < .05$), suggesting the violation of assumption of sphericity. Therefore, we employed the Huynh-Feldt correction to confront any Type I error. The results of Factorial ANOVA are provided in Table three.

The Effects of Discipline-related Knowledge and English Proficiency on EAP Reading

Comprehension

The first research question addressed the role of English proficiency and discipline-related knowledge in Iranian EAP reading test. The simple effect results (Table 4), used to break down the interaction to understand precisely what happens, manifested that subject-related knowledge significantly helped the students of psychology in the first paragraph with lowest content specificity ($p < .001$, partial $\eta^2 = .17$) and the second paragraph with medium content specificity ($p < .01$, partial $\eta^2 = .05$). Discipline-related knowledge also had a significant influence on the reading performance of the students of biology in the first paragraph with lowest content specificity ($p < .05$, partial $\eta^2 = .03$) and the third paragraph with highest content specificity ($p < .001$, partial $\eta^2 = .19$).

Table 4

Simple Effect Results for Discipline-related Knowledge and EAP Reading Tests

Dependent variable	(I) Major	(J) Major	Mean Difference (I-J)	Std. Error	Sig.	F	Partial η^2
PS1	Psychology	Biology	1.36	.24	.000	32.55	.173
PS2	Psychology	Biology	1.23	.40	.002	9.51	.057
PS3	Psychology	Biology	-1.02	.32	.002	9.70	.059
BI1	Psychology	Biology	-.81	.33	.014	6.13	.038
BI2	Psychology	Biology	.27	.33	.411	.67	.004
BI3	Psychology	Biology	-1.34	.22	.000	36.59	.190

Note: PS = Psychology; BI = Biology

As for the role of English proficiency in Iranian EAP RC performance, the results (Table 5) indicated that it significantly contributed to our participants' performance on EAP RC in all the subtests. In the first paragraph with lowest content specificity ($p < .001$, partial $\eta^2 = .36$), the second paragraph with medium content specificity ($p < .001$, partial $\eta^2 = .29$), and the third paragraph with highest content specificity ($p < .001$, partial $\eta^2 = .16$), the contribution of language proficiency to reading comprehension (psychology) was statistically significant with large effect sizes. As for the subtests of biology, similar results were obtained for the first ($p < .001$, partial $\eta^2 = .14$), second ($p < .001$, partial $\eta^2 = .20$), and third ($p < .001$, partial $\eta^2 = .13$) paragraphs.

Table 5

Simple Effect Results for English Language Proficiency and EAP Reading Tests

Dependent variable	(I) Proficiency	(J) Proficiency	Mean Difference (I-J)	Std. Error	Sig.	F	Partial η^2
PS1	Low	High	-2.26	.24	.000	89.55	.365
PS2	Low	High	-3.23	.40	.000	65.40	.295
PS3	Low	High	-1.82	.32	.000	30.75	.165
BI1	Low	High	-1.68	.33	.000	26.04	.143
BI2	Low	High	-2.12	.33	.000	40.52	.206
BI3	Low	High	-1.10	.22	.000	24.97	.138

Note: PS = Psychology; BI = Biology

The Interaction between English Proficiency, Discipline-related Background Knowledge, and Content Specificity in EAP Reading Comprehension

The second research question addressed the interaction between English proficiency, field-related background knowledge, and content specificity in EAP reading comprehension to find if the contribution of the first two variables fluctuates with changes in content specificity of the passages. The results (Table 6) showed that the interaction between the variables was significant ($F(1.49) = 8.29, p < .01$, partial $\eta^2 = .051$), which shows that the influences of the content specificity on the participants' EAP test performance fluctuated with changes in proficiency level and discipline-related knowledge. For Cohen et al. (2007), the size of the effect is small for this interaction. The results further corroborate the findings for the effects of language proficiency and field-related knowledge. The two-way interaction between the students' language proficiency and the EAP reading texts with three levels of content-specificity was significant ($F(1.49) = 14.72, p < .001$, partial $\eta^2 = .86$), indicating that the effect of proficiency level on their performance varied for the 3 texts. Results in Table 4 indicated that subject-related knowledge did not affect the performance of the low proficiency students while it significantly affected the performance of the high proficiency students. Disciplinary knowledge also had a significant interaction with the comprehension of passages with three levels of content specificity ($F(1.72) = 8.64, p < .001$, partial $\eta^2 = .05$), showing that knowledge of content contributed to the EAP reading tests.

Table 6

Comparison of Discipline-related Knowledge and Language Proficiency across Three Levels of Content Specificity

Source	Sum of Squares	df	Mean Square	F	Sig.	Partial η^2
Content * Proficiency	58.10	1.49	38.85	14.72	.000	0.86
Content*Discipline	42.10	1.72	24.41	8.64	.000	.052
Content * Proficiency * Discipline	32.74	1.49	21.89	8.29	.001	.051
Error (Content)	615.68	233.32	2.63			
Discipline	272.39	1.00	272.32	164.54	.000	.513
Discipline * Proficiency	37.87	1.00	37.78	22.88	.000	.128
Error (Discipline)	258.24	156.00	1.65			
Error (Content * Discipline)	759.95	759.95	2.82			

Discussion

The main aim of the investigation was to study the contributory effect of field-related knowledge and language knowledge on Iranian EAP students' performances with regard to changes in the degree of content specificity. Further, considering the LTH, the researchers intended to find out how Iranian EAP students with different proficiency levels perform on disciplinarily close and remote EAP RC tests. To this end, an English language proficiency test, two Farsi knowledge tests, and two English content tests relevant to biology and psychology were used.

Firstly, findings suggest that the LTH is not completely confirmed regarding Iranian EAP students (at least in the case of our participants). In fact, the researchers failed to find enough evidence for *all proficiency levels* in order to support the LTH. More specifically, with regards to high proficiency biology and psychology test takers, both groups performed statistically differently on EAP tests coming from inside vs. outside their academic neighborhoods. This runs counter to one of the predictions of the LTH: for high proficiency EAP students it makes no difference where RC texts come from since they take maximum advantage of their L2 proficiency so that the need to background knowledge is bypassed. As for low proficiency biology and psychology test takers, it was for the latter group only that we found support for the LTH. Otherwise stated, for low proficiency psychology test takers, it made no difference whether they took psychology or biology EAP RC tests—an endorsement of the LTH—while low proficiency biology participants did significantly better on EAP tests corresponding to their own academic discipline. In a nutshell, our findings disagree with the prediction of the LTH for high proficiency EAP test takers, and partially agree with it for low proficiency students.

Consistent with Krekeler (2006) who failed to find a clear-cut picture of the LTH as a powerful predictor of the participants' performance on EAP RC tests in Germany, we use our findings to maintain that the explanatory power of the LTH is prone to challenges as the contributory effects of students' background knowledge on EAP RC test performance were found not to be *strictly* bound to two language thresholds (i.e., high and low). Of course, similar to Krekeler (2006) who concludes that his findings do not “disprove the two thresholds hypothesis” but indicate that the application and relevance of the LTH to LSP (Language for Specific Purposes) test construction is uncertain and open to question, we suggest that our findings neither endorses nor dismisses the LTH completely. As such, the LTH seems not to be rigorous enough to *explain* EAP students' performance on RC tests and it is inadvisable to interpret EAP test results with reference to the LTH. Despite this, since our findings *partially* lend support to the LTH with *low* proficiency EAP test takers and given that studies (Carrell, 1991; Ridgway, 1997; Taghizadeh Vahed & Alavi; 2019) reporting similar results with different low proficiency students outnumber studies suggesting otherwise, one might cautiously suggest that a *weak version* of the LTH might be tenable. Of course, further studies are needed to examine if the weaker version proves to be more explanatory of EAP test performance.

Secondly, and as far as the role of content specificity in biology and psychology students' test performances on RC EAP tests are concerned, our general findings indicated that with fluctuations in content specificity, high and low proficiency students' performances on field-related and field-unrelated were affected. Overall, as content specificity increased, both groups' performances on field-related and field-unrelated tests became poorer and vice versa. This general pattern was observed at inter- and intra-discipline levels; with gradual increase of content specificity, low and high proficiency biology students' performance was negatively affected in EAP tests of both biology and psychology. The same pattern was evident for psychology students as well. Our findings might help explain one of the common speculations in similar studies. Earlier researchers, when explaining why the LTH failed to accurately predict low and high proficiency students' EAP test performance, speculated that the predictive force of the LTH might be seen more evident should content/subject specificity of selected texts were high (Krekeler, 2006). In this study, texts were ordered in terms of content specificity—as determined by content specialists—to see if the predictive potential of the LTH improves as EAP texts become more and more specific. The results in Table 5 suggest that as content specificity of EAP reading texts increases, the predictive force of the LTH improves.

Thirdly, we use our qualitative data (i.e., interviews with content specialists) to suggest that content specificity of EAP reading texts seems to be a complicated notion that cannot be accounted for by quantitative measures such as readability indexes. More specifically, content specificity and text difficulty are not necessarily the same. Despite the fact that measures of text complexity/difficulty such as readability indexes, number of passive sentences, sentences with participles, etc. have been used (Tarlani-Aliabadi, et al., 2022; Taghizadeh Vahed & Alavi, 2019) to discuss content specificity, we believe there might or might not be a correspondence between measures of text difficulty/complexity and content specificity. What lends support to this argument is the lack of complete correspondence between quantitative measures (i.e., Flesch Readability Index) of our texts and qualitative measures we used for determining content specificity. Of course, as explained before, we followed the latter in this study due to controversies associated with readability indexes in general and Usó-Juan's (2006) advice pertaining to how to determine which EAP text is more content-specific. Admittedly, more studies are needed to gain an in-depth understanding of what content specificity is and how it relates to EAP reading comprehension assessment.

Conclusion and Implications

Interest in predicting EAP students' performances on reading comprehension tests is almost four decades old. The interest is justified on theoretical and pedagogical grounds as shedding light on factors contributing to students' performance explains the theoretical construct underpinning EAP reading comprehension, on the one hand, and helps with better EAP test construction, use, and interpretation, on the other (Cai & Kunnan, 2019).

The LTH has been proposed to explain and predict EAP students' performances on RC tests. According to the Hypothesis, the performance is bound to two language thresholds: high and low, although in some studies (Smith et al, 2021; Kendeou & Van Den Broek, 2007) intermediate proficiency level is also addressed. Our findings indicate that the LTH fails to accurately predict EAP students' performances on RC tests in general; it could do so with one group of low-proficiency test takers. We interpreted such an overall inconclusive picture neither as an endorsement of the LTH nor its rejection since its predictive force changed not only with language proficiency level (failure to predict high proficiency students' performance but explaining intermediate level students' scores) but also with EAP students' discipline background; the LTH could predict low proficiency *psychology* students' performances but not that of *biology* students. Moreover, findings revealed that the contributory effects of English knowledge and background knowledge fluctuate with changes in content specificity of EAP texts; the predictive value of the LTH becomes more evident with more content-specific EAP texts.

The implications of the study are both theoretical and pedagogical. Theoretically, the LTH seems to fall short of fully explaining and predicting Iranian EAP students' performance on RC tests. Given the fact that there is a generally similar inconclusive picture of the LTH in other studies, and regarding how content specificity results in fluctuations in high and low proficiency EAP students' performances on RC tests, it seems safe to propose that what underpins EAP RC construct is too multifaceted to be accurately accounted for by the LTH and alternative theoretical postulations might be proposed in future. We use our results to argue that the contributory effects of the two constructs—discipline specific knowledge and English proficiency—are likely to vary depending on a host of factors including not only the focus of this study, i.e. content specificity, but also variables already documented to be significant mediators between the two. The implication is that attempts to make predictions of EAP students' performances on RC tests are likely to face challenges since not only such predictions on the basis of L2 proficiency are inaccurate but also the predictions on the basis of L2 proficiency are prone to further challenges when other pertinent variables, e.g. content specificity, are brought to surface. Therefore, EAP test users are suggested not to interpret EAP RC test scores based on the LTH according to which the effects of test takers' background knowledge on EAP RC are bound to proficiency thresholds. Pedagogically, and as far as EAP RC test construction is concerned, we call on EAP teacher-testers to critically address the extent background knowledge affects students' performances in EAP RC and if such effects are predictable by students' English proficiency level. Our

findings suggest that the LTH fails to accurately account for their performances and, as a result, it is advisable to dispense with the claims the LTH makes. More specifically, the claim that for low and high proficiency students it does not make a difference if RC texts come from neighboring or non-neighboring academic disciplines seems not to be tenable. Therefore, EAP test designers need to consider test takers' academic disciplines while bearing in mind that EAP assessment in general is fundamentally a language rather than a content-focused enterprise (Airey, 2016; Karimpour & Mazlum, 2024; Poorebrahim & Mazlum, 2013).

Findings of the study are not claimed to be generalizable. We followed Usó-Juan's (2006) proposal to order the texts in terms of content specificity. More studies are needed to provide us with more established criteria in this regard. We also used cloze tests as measures of EAP reading comprehension ability; other test formats need to be designed and used in similar studies as test method facets are important in assessing language skills in general. We selected EAP reading topics (cell division and autism) in collaboration with subject specialists in biology and psychology departments. Corpus-based criteria (e.g. students' EAP reading materials) can be used for such purposes as topic selection in such studies is a source of debate. Interviews with EAP test takers can help us understand their lived experiences with EAP RC tests coming from close vs. distant academic territories, test methods employed, and topics selected. Finally, and on a broader note, whether it is necessary or valuable for different discourse communities (Swales, 2016) in EAP to predict would-be members' performances on different EAP tests needs more in-depth qualitative inquiries.

Acknowledgement

We would like to express our gratitude to two reviewers of the IJEAP who meticulously examined the paper and provided insightful feedback. We are also thankful to students who accepted to participate in the study. Finally, we are grateful to psychology and biology professors at the University of Maragheh who helped us with text choice and qualitative data collection.

Declaration of Conflicting Interests

We have no conflicts of interest to disclose.

Funding Details

We received no funding for the study.

References

- Akbarian, I. (2010). The relationship between vocabulary size and depth for ESP/EAP learners. *System*, 38, 391-401.
- Airey, J. (2016). EAP, EMI or CLIL. In P. Shaw & K. Hyland (Eds.), *The Routledge handbook of English for academic purposes*. Routledge.
- Alderson, J.C. (1993). The relationship between grammar and reading in an English for Academic Purposes test battery. In D. Douglas & C. Chapelle (Eds.) *A new decade of language testing research* (pp. 203-219). TESOL, Alexandria, VA.
- Benjamin, R.G. (2012). Reconstructing readability: Recent developments and recommendations in the analysis of text difficulty. *Educational Psychology Review*, 24, 63-88.
- Cai, Y., & Kunnan, A. J. (2019). Detecting the language thresholds of the effect of background knowledge on an LSP reading performance: A case of the island ridge curve. *Journal of English for Academic Purposes*, 42, 1-32.

- Carrell, P. (1991). Second language reading: Reading ability or language proficiency? *Applied Linguistics*, 12(2), 159-179.
- Clapham, C. (1996). *The development of IELTS: A study of the effect of background knowledge on reading comprehension*. Cambridge: Cambridge University Press.
- Clapham, C. (2000). Assessment for academic purposes: Where next? *System*, 28(4), 511-521.
- Clarke, M. A. (1980). The short circuit hypothesis of ESL reading-or when language competence interferes with reading performance. *The Modern Language Journal*, 64(2), 203-209. <https://doi.org/10.2307/325304>.
- Cohen, L., Manion, L., & Morrison, K. (2007). *Research Methods in Education* (6th ed.). London & New York, NY: Routledge.
- Darabi Bazvand, A. (2019). L1 domain-specific knowledge as predictor of reading comprehension in L2 domain-specific texts: the case of ELT student teachers. *Cogent Education*, 6, 1-11.
- Derakhshan, A., & Janebi Enayat, M. (2020). High- and Mid-Frequency Vocabulary Size as Predictors of Iranian University EFL Students' Speaking Performance (Research Paper). *Iranian Journal of English for Academic Purposes*, 9(3), 1-13.
- Douglas, D. (2013). ESP and assessment. In B. Paltridge & S. Starfield (Eds.), *The handbook of English for specific purposes* (pp. 367-384). Chichester, UK: Wiley-Blackwell.
- Fulcher, G. (1999). Assessment in English for academic purposes: Putting content validity in its place, *Applied Linguistics*, 20(2), 221-236.
- Janebi Enayat, M., & Babaii, E. (2018). Reliable predictors of reduced redundancy test performance: The interaction between lexical bonds and test takers' depth and breadth of vocabulary knowledge. *Language Testing*, 35(1), 121-144. <https://doi.org/10.1177/0265532216683223>
- Janebi Enayat, M., & Derakhshan, A. (2021). Vocabulary size and depth as predictors of second language speaking ability. *System*, 99, 102521. <https://doi.org/10.1016/j.system.2021.102521>
- Janebi Enayat, M., Amirian, S. M. R., Zareian, G., & Ghaniabadi, S. (2018). Reliable measure of written receptive vocabulary size: using the L2 depth of vocabulary knowledge as a yardstick. *SAGE Open*, 8(1), 215824401775222. <https://doi.org/10.1177/2158244017752221>
- Karimpour, P., & Mazlum, F. (2024). EAP practitioners' assessment behavior: Bringing the hidden-away to light. *Journal of English for Academic Purposes*, 67, 101321. <https://doi.org/10.1016/j.jeap.2023.101321>
- Kendeou, P., & Van Den Broek, P. (2007). The effects of prior knowledge and text structure on comprehension processes during reading of scientific texts. *Memory & Cognition*, 35(7), 1567-1577. <https://doi.org/10.3758/bf03193491>.
- Krekeler, C. (2006). Language for special academic purposes (LSAP) testing: the effect of background knowledge revisited. *Language Testing*, 23 (9), 99-130.
- Mazlum, F., Shamameh, S., & Salimi, A. (2020). The Relationship between Disciplinary Distance, Language Proficiency and EAP Test Performance. *Iranian Journal of English for Academic Purposes*, 9(4), 31-44.
- Peretz, A. S. & Shoham, M. (1990). Testing reading comprehension in LSP: Does topic familiarity affect assessed difficulty and actual performance? *Reading in a Foreign Language* 7, 447-455.
- Poorebrahim, F., & Mazlum, F. (2013). An ESP approach to course and material design for the students of agriculture. *Taiwan International ESP Journal*, 5(2), 21-46.
- Read, J. (2009). *Assessing vocabulary*. Cambridge: Cambridge University Press.
- Ridgway, T. (1997). Thresholds of the background knowledge effect in foreign language reading. *Reading in a Foreign Language*, 11, 151-168.

- Salmani-Nodoushan, M. A. (2003). Text familiarity, reading tasks, and ESP test performance: A study on Iranian LEP and non-LEP university students, *The Reading Matrix*, 3(1), 1-14.
- Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring behaviour of two new versions of the vocabulary levels test. *Language Testing*, 18(1), 55-88.
- Smith, R., Snow, P., Serry, T., & Hammond, L. (2021). The role of background knowledge in reading comprehension: a critical review. *Reading Psychology*, 42(3), 214-240. <https://doi.org/10.1080/02702711.2021.1888348>
- Swales, J. (2016). Reflections on the concept of discourse community. *La Revue Du Geras ASP* 69, 7-19.
- Taghizadeh Vahed, Sh. & Alavi, S. M. (2019). The role of discipline-related knowledge and test task objectivity in assessing reading for academic purposes. *Language Assessment Quarterly*, 17(1), 1-17.
- Tarlani-Aliabadi, H., Tazik, Kh., & Azizi, Z. (2022). Exploring the role of language knowledge and background knowledge in reading comprehension of specific-purpose tests in higher education. *Language Testing in Asia*, 12, 48. <https://doi.org/10.1186/s40468-022-00198-x>
- Usó-Juan, E. (2006). The compensatory nature of discipline-related knowledge and English- language proficiency in reading English for academic purposes. *The Modern Language Journal*, 90(2), 210-227.