Effect of cultural familiarity on reading comprehension performance:

A case-study of Vietnamese and Chilean EFL learners

Nguyen Ngoc Bao Chau*, Sean Liu, Katia Lopez

Abstract

This paper investigates the influence of cultural background knowledge in a reading comprehension test of English as a foreign language. In order to explore the general assumption that cultural background knowledge would help students to better comprehend a text, a reading comprehension test including two different passages describing a Vietnamese and a Chilean national holiday was developed. This test was undertaken by a group of 41 Vietnamese and 35 Chilean students in order to observe whether there was an effect of background knowledge in their scores for each passage. In addition, a brief questionnaire was used to examine the students' perceptions of their familiarity with the texts based on their background knowledge in order to relate it with their scores. Results from the t-tests revealed statistically significant differences between the two groups. Vietnamese students performed better in the section about the Vietnamese holiday and Chilean students outperformed Vietnamese students in test section about the Chilean holiday, which suggested a positive influence of cultural background knowledge. Weak but significant correlations were found between the students' reported familiarity with the texts and their scores in each section. The findings of this study provide some relevant implications for language testing and also second language instruction.

Introduction

The ability to understand, extract and utilise information through reading a text is included in many test constructs to ascertain the linguistic proficiency of people who use English as a second language. Reading components are included both for English tests for academic purposes, such as the PTE and the TOEFL, as well as tests for more general purposes, such as the IELTS (General Training) and the Cambridge PET. A test taker has to utilise many different skills in order to read a text, depending on what he or she is asked to do with the information

*Corresponding author: Nguyen Ngoc Bao Chau – Hue University of Foreign Languages, Hue University
chaunguyen@hueuni.edu.vn

presented within the text (Hughes, 2002). Different skills and cognitive processes are activated when different types of texts are read.

There are many factors that contribute to an individual's performance on a reading test such as language proficiency, reading strategies and background knowledge (Alderson, 2000). It is generally understood that prior or background knowledge has some degree of influence on an individual's ability to comprehend and extrapolate from a text; more knowledge of the topic of the text will lead to better comprehension and learning (Braasch & Goldman, 2010). Background knowledge can be reflected through topic familiarity (i.e. a scientist reading a text about biology) or cultural familiarity (i.e. an Australian reading about cricket). A good lexical resource, which helps in reading comprehension, is also in some sense an indication of an individual's topical background knowledge, especially when considering domain-specific vocabulary.

This paper sets forth the methodology undertaken to investigate whether cultural background knowledge has any substantial influence on the scores of second language learners on reading comprehension tests. We begin by reviewing the literature on reading and comprehension, and other studies done pertaining to the role of background knowledge on test scores. Next, we set out our methodology in developing the instruments, selecting participants and collation of data. And finally we end by presenting and discussing our findings before discussing the implications of our results on test fairness and the teaching of English as a second language.

Literature Review

Cognitive processes in reading

Before critically reviewing literature pertaining to the role of cultural background knowledge on the comprehension of texts and resultant scores on reading tests, it is fruitful to briefly examine literature on the cognitive processes involved in reading. Van Den Broek, Young, Tzeng and Linderholm (1999) more than a decade ago have conceptualised the Landscape Model of Reading where they have identified four sources of activation that may happen when a person reads a particular text. Later, research done by Van Den Broek, Virtue,

Everson, Tzeng and Sung (2002) on comprehension of science texts note that a "coherent mental representation of a text consists of a network of semantic relations between text elements and the reader's background knowledge" (p. 133). Around the same time, Nassaji (2002), concluded that models of L2 reading comprehension are interactive because they involve processes that are both data-driven which is analysis of the text per se, and reader-driven that can be understood as an individual using background knowledge and linguistic proficiency to help comprehend the text.

In terms of reader-driven processes, schema theory in reading divides schemata into two categories: formal schema, related to knowledge of language, and content schema, related to domain-specific knowledge (Bernhardt, 1991). Ketchum (2006) later defined cultural schema as a culture specific branch of content schema and articulated the fact that it is an instrumental process that a reader needs to use to fully understand a text. Cultural background knowledge plays a role in reading comprehension, but it is unclear whether or not it influences comprehension to an extent where it would actually affect test scores.

Effect of cultural background knowledge

Cultural background knowledge does not only become manifest in the form of tangible content knowledge in an individual's endeavour to comprehend a text in his or her second language; sometimes culturally salient reading strategies, for example a preference for top down or bottom up reader-driven processes, also influence text comprehension and scores. Certain items types that are included in reading tests, for example those that require bottom-up processing, would favour certain cultural groups because they had been taught to read in this way (Nassaji, 2002). Abbott (2006) found that Chinese teachers tend to use approaches that result in Chinese students mostly paying attention first to form and later to meaning while the opposite is true for students in the Arab nations. Differential item functioning was used to determine the probability of scoring the correct answer for the same item for different groups of people and it was found that even though there were minimal score differences between the two groups of examinees on bottom-up (form first) and top-down (meaning first) bundles of items, different strategies were in fact used by the different groups (Abbott, 2006). This shows that cultural background knowledge about reading strategies is a significant consideration in assessing

reading because certain cultural groups may approach test items differently; a test with more bottom-up items, for example, may in fact favour some cultural groups.

Research has also been done on nativising texts so that even though the content remains the same, the actual elements become more familiar to test takers. Alptekin (2006) and later Erten and Razi (2009) both designed an experiment where they nativised an original text by changing many elements to culturally familiar ones for Turkish test takers. Both studies found that the groups that received the nativised version scored better than those who did not in terms of comprehension of texts. While Alptekin's study made use of multiple choice questions to assess comprehension of the text, Erten and Razi (2009) used free recall tasks to ascertain if students remembered and comprehended content. Here it becomes evident that there is a significant difference in comprehension and performance on comprehension-related reading tests even if only basic elements are adapted to make the texts more culturally familiar.

Studies have also been done where completely different texts that dealt with culturally different topics were given to different groups of test takers. Lee (2007) conducted research on topic familiarity on Korean EFL students' reading comprehension and learning of passive form and found that groups that were given texts on more culturally familiar topics scored better when asked to recall content, but did not score better when asked to complete a form correction task on the passive voice. Dehghan and Sadighi (2011) designed a study where Iranian students were asked to read two texts that contained culturally familiar themes and two texts that did not. Assessment consisted of a mixture of questions that required local (bottom-up) and global (top-down) processing. The results showed that readers outperformed themselves on familiar texts as compared to unfamiliar texts and it was concluded that cultural familiarity (that is not always related to linguistic competence) does have a significant influence on reading comprehension. Li and Lai (2012) wanted to test if culturally more familiar texts resulted in a shorter reading time and a better score on a cloze text and designed an experiment where the independent variable was two texts, one that was more culturally familiar and another that was more culturally unfamiliar, and where the dependent variables were time spent in reading the text and performance on the cloze test. They found that culturally more familiar texts does in fact result in a higher degree of comprehension of the texts and also takes a shorter time to read. These

three studies all indicate that culturally more familiar texts facilitate comprehension and result in better scores on reading comprehension tests.

However, a study done by Stott (2004) did not support this hypothesis. In this study, two groups of Japanese students were given the same slightly simplified English translation of a Japanese novel, but one group was told the origin and source of the text while the other group was left in the dark. Results showed that the group who knew the text origin did not outperform the other group. The design of this experiment differs from those previously mentioned in that the two groups were given the same text and that there was no obvious difference between the two groups in terms of cultural knowledge. However, the results here do not support the argument that cultural knowledge is activated in a reading of a text in a second language to a large degree.

In fact, some studies have found that language proficiency is a stronger predictor of academic reading performance than background knowledge (Hill & Liu, 2012). Chan (2003) had Hong Kong (HK) and mainland Chinese students of differing proficiency levels complete a cloze test on both a culturally neutral text and a text that would be more culturally familiar to the HK students. Results showed that for the lower proficiency groups, HK students scored significantly higher than mainland Chinese students on the more culturally familiar text, but there was no significant difference between the higher proficiency students for performance on either text (Chan, 2003). Hill and Liu (2012), selected three passages on cultural topics and two on physical science for a study that utilised differential item functioning to ascertain the relationship between background knowledge and language proficiency in TOEFL iBT tests. Results showed some level of interaction between background knowledge and proficiency level at the item level, but there was no passage-level differential functioning. Furthermore, the study found evidence that background knowledge does not always work to an individual's advantage, especially when individuals rely too heavily on their background knowledge in completing the test when answering items that require local (bottom-up) processing. This study showed that participants used background knowledge to answer the questions, but did not indicate that it has in fact helped them comprehend the intended meaning of the writing.

Given the mixed findings in the literature pertaining to the influence of cultural familiarity on reading comprehension test scores, the current study is motivated to further investigate this interaction. Specifically, the study examines the effect of cultural familiarity on reading performance of Vietnamese and Chilean students at A2 to B1 level (Common European Framework of Reference).

Research questions

The following research questions inform this study:

1. *Is there a significant difference between students from different backgrounds on their test scores on respectively more culturally familiar texts?*
2. *What is the relationship between students' perception of their cultural familiarity and their test scores?*

Methodology

Participants

Two groups of learners of English as an additional language from Vietnamese and Chilean backgrounds participated in this study. The Chilean participants were originally 50 native speakers of Spanish with ages ranging between 17 and 18 years old. The teacher reported the students to be at an English proficiency level of A2 to B1, according to the CEFR. They belonged to two different year 12 classes in the same public school located in Viña del Mar, one of the main cities in Chile.

The Vietnamese students participating in the study were 119 first year university students majoring in English at a Vietnamese public university. Their ages range from 18 to 19. Their proficiency levels were also reported by their teacher to be at an A2 to B1 level. The students belonged to 4 different classes of the same subject Reading 1. 65 students from two classes participated in the pilot of the test instrument, and 54 students from the other two classes participated in our final test version.

Due to the limited time and resources of this study, we were not able to test students' proficiency prior to data collection and had to rely on teachers' reports. We realised the use of

reported proficiency levels from teachers can be problematic as it was subjective and there could be inconsistency between the Chilean and Vietnamese teachers' interpretation of the CEFR levels. Therefore, we added an extra complementary reading section to our test (See *Instruments* section) and used the total test scores to omit outliers (i.e. those whose test scores are too high or too low). In our final selection of participants for analysis, we chose 41 Vietnamese (M=3.68, SD=1.19) and 35 Chilean students (M=3.71, SD=1.22). We believe the selection has resulted in a more homogenous group of participants in regards to proficiency level.

Table 1. Summary of participants' background for each stage

	N	Gender	Nationality	Proficiency	Age
Piloting stage	65	6M, 59F	Vietnamese	A2 - B1	18 - 19
Data collection	54	8M, 46F	Vietnamese	A2 - B1	18 - 19
	50	23M, 27F	Chilean	A2 - B1	17 - 18
Data analysis	41	5M, 36F	Vietnamese	upper-A2	18 - 19
	35	15M, 20F	Chilean	upper-A2	17 - 18

Instruments

Designed Test

The target domain for our designed test was the reading that needs to be done by an individual in everyday situations when living in a country where English is the dominant, if not only, language. For this, the universe of generalisation was the ability to comprehend information presented within an informational text about cultural events.

Texts

In order to answer our first research question we designed a test that deliberately included two texts that were about culturally familiar topics to Vietnamese and Chilean participants. Text A was about *Tet Nguyen Dan* (Vietnamese New Year) and Text B was about *Fiestas Patrias*

(Chilean Independence Day); these events are considered to be significant in the respective countries and so we were able to assume that the participants in both Vietnam and Chile would have some cultural familiarity with one of the texts. The texts were drawn from the same source (Henderson, 2005) and edited to retain proper nouns in the original language, such as *Tet Nguyen Dan* and *banh chung* for text A and *Fiesta Patrias* and *cueca* for text B, as well as to facilitate the creation of multiple choice items. Care was also taken to ensure that both texts were of similar length, approximately 270 words each. The texts had to be long enough for 5 multiple-choice items each, but short enough so that participants could complete the test and survey in the allotted class time.

Item Type

We chose multiple choice questions due to the fact that they would provide us with more discrete data from which to work. This item type is also prevalent in reading tests taken by students at the same proficiency level as our participants, such as the Cambridge PET.

Item Descriptors

Care was taken to ensure that each text included both items that required global and local reading processes. Each text has five of the same item types to ensure the similarity between the two sets of questions. Table 2 shows a summary of the item descriptors.

Table 2. Item descriptors

Item type	Specific task description	Example
Local reading processes:		
Vocabulary item	Infer meaning of an unknown word from context	*Which word is closest in meaning to "evil" in the following sentence in paragraph 2?*
Pronominal reference	Assign previously mentioned subjects or objects to pronouns.	*What does the word 'it' in the following sentence in paragraph 3 refer to?*
Organisation of information	Insert a sentence into the correct place within a paragraph, maintaining the coherence.	*Where does this sentence fit best?*
Global reading processes:		

| Identifying implicitly stated main ideas | Infer implicitly stated main ideas in a paragraph. | *What can be inferred from the first paragraph?* |
| Identifying explicitly stated main ideas | Identify explicitly stated supporting ideas in a paragraph. | *Why do the Vietnamese people release a special type of fish into the river?* |

Complementary Reading Section

Because our test was relatively short with only ten items, we included an additional section with five questions to obtain a more trustworthy measure of students' proficiency. For this section, we used a sample PET reading section with a culturally neutral text and similar test item format. The scores from this section aid us in limiting our participants to a more homogenous range to minimise the proficiency variable.

Piloting

We piloted our test version A with a panel of three Chilean and four Vietnamese language professionals, and adjusted items for difficulty to develop version B. Version B was piloted with a sample from the participant population (65 students in Vietnam). From this piloting we carried out item analysis and distractor analysis. The analyses revealed several problematic test items. The level of difficulty was too low for items 1, 4 and 8, and too high for items 6 and 7. In addition, items 3, 4, 6 and 7 did not provide much discrimination between students. Distractor analysis helped us identified problematic distractors to modify. For example, items 1, 4 and 8 had a disproportionate distribution of distractor selection. We edited these items by modifying difficulty level of questions and changed problematic distractors. In developing Version C, we were aware that our pilot test was only completed by Vietnamese participants and therefore tried to allow for any inconsistencies that might have arisen as a result of cultural familiarity with Text A or the lack thereof for Text B. The edited test was our test instrument that was sent out to participants.

Validation of version C:

After sending out version C to all our participants, we analysed the data received for reliability and found that Cronbach's alpha was low (α= 0.4). We also conducted test item analysis; see Table 3 for the IF and R values for each of the items.

Table 3. Item analysis for Test version C

Item No.	IF	r pbi
1	0.83	0.43
2	0.31	0.4
3	0.34	0.42
4	0.24	0.17
5	0.33	0.43
6	0.1	0.22
7	0.53	0.38
8	0.61	0.4
9	0.22	0.35
10	0.57	0.55

From Table 3 we can see that item 4 was too difficult (IF = 0.24) and also does not discriminate well (r=0.17) between different levels of students. Item 6 was similar in terms of difficulty (IF = 0.1) and discrimination (r=0.22). A decision was made to disclude scores from items 4 and 6 in final analysis. Both items involved global reading processes where test takers had to identify implicitly stated ideas, and this particular reading process was thought to be too difficult for our participants.

Item 1 was relatively easier (IF = 0.83) than the resulting 7 items, but we retained it because it was able to discriminate well (r = 0.43) between different levels of students. Also, based on the figures, the average difficulty for both the Vietnamese text and the Chilean text was

10

similar. From this we were able to assume that overall difficulty for the Vietnamese text was the same as overall difficulty for the Chilean text.

The final version for analysis had 8 questions (Version D) with a new reliability coefficient of $\alpha = 0.5$. This was deemed to be acceptable due to the low number of test items and homogeneity of participants.

Questionnaire

In order to answer our second research question which addresses the relationship between students' perceptions of their cultural familiarity and their scores, we designed a questionnaire consisting of eight statements. This aimed to ascertain the extent to which participants felt they activated their cultural knowledge while answering the questions. Following Dörnyei and Taguchi (2009), the questionnaire was translated into Vietnamese and Spanish in order to assure participants' appropriate understanding of its content. Participants selected their answers using a Likert scale from 1 (completely disagree) to 5 (completely agree). Statements 1, 2, 3, and 4 formed the variable "Familiarity with Text A", and statements 5, 6, 7, and 8 formed the variable "Familiarity with Text B".

In the process of piloting our test instrument, we also piloted our questionnaire. Reliability analyses showed that the Cronbach's alpha for both variables were $\alpha = 0.5$, which is moderately low. However, we kept in mind that only Vietnamese participants completed the pilot questionnaire and would engender similar responses. Therefore, we assumed the reliability to be higher with a more heterogeneous pool of participants that would include Chilean learners. Table 4 shows examples of the statements for each variable.

Table 4. Questionnaire examples

Familiarity Variables	
Familiarity with text A (Vietnam)	Example (statement 1): *In general, I think the text about Vietnam is easier to read than the text about Chile because of my familiarity with the Vietnamese text.*
Familiarity with text B (Chile)	Example (statement 6): *I did not have to read the text about Chile very carefully to answer the questions, due to my*

Data collection procedure

Once the preliminary versions of the test and questionnaire were piloted and revised, the final version was generated. We used *Google Forms* due to the geographical distance between our research team and the two sites of the study. This free online tool allowed us to develop a computer-based test and questionnaire and facilitated the administration, management and scoring of participants' responses. The link to the online test was provided to two teachers, one in Vietnam and one in Chile, who helped with its administration. Students at each site of the study completed the test electronically during school hours. They were given 45 minutes to finish the test and answer the questionnaire. The timing was informed by the initial piloting session.

Data analysis

Scores from the online test and questionnaire results were extracted into spreadsheets. The data was then transferred to the Statistical Package for the Social Sciences (SPSS) for analysis. We conducted a participant selection process whereby we chose 41 Vietnamese and 35 Chilean participants, all achieving overall scores from 4 to 9 out of a total of 13 (Text A, text B and the Complementary Text). The mean scores and standard deviations for selected Vietnamese and Chilean participants were at similar values (see *Results* section).

Through our participants' saturation process as well as test instrument validation, we were able to claim the students were at similar proficiency level and that the two texts and questions were of similar difficulty. Therefore, we conducted a series of dependent and independent t-tests to observe any significant differences within-group and across-group performances. For the questionnaire, Spearman rho correlation was used to determine the relationship between students' reported cultural familiarity and their scores.

12

Results

Analysis of test data

Descriptive statistics
Table 4 illustrates Vietnamese and Chilean students' performance on the whole test and on each text A and B, respectively about Vietnamese and Chilean traditional holidays. Overall, the test is difficult as the mean scores for the whole test is less than 50% for both Vietnamese (M=3.68) and Chilean students (M=3.71). Each group respectively scored higher on the text about their national holiday.

Table 5. Descriptive statistics of participants' performance

Group	N	Whole test		Text A (Tet Nguyen Dan)		Text B (Fiestas Patrias)	
		M	*SD*	*M*	*SD*	*M*	*SD*
Vietnamese students	41	3.68	1.19	2.07	0.82	1.61	0.92
Chilean students	35	3.71	1.22	1.51	0.78	2.20	0.90

In order to answer the first research question and determine whether cultural familiarity affected students' performance, a series of t-test analyses were conducted.

Within-group comparisons
Two dependent t-tests were conducted to determine whether there are any significant differences between performances on questions on Text A and Text B within each group of participants. Table 6 shows results from these analyses.

Table 6. Dependent t-tests between Texts A and B

	N	t	df	p
Vietnamese students	41	2.34	40	p <.05
Chilean students	35	3.51	34	p <.001

Results show significant differences in performances on Text A and Text B for both Vietnamese group ($p < .05$) and Chilean group ($p < .001$). Vietnamese students performed significantly higher on questions about the Vietnamese holiday, Tet Nguyen Dan, $t(40) = 2.34$, $p < .05$. Similarly, Chilean students performed better on the text about Fiestas Patrias, $t(34) = 3.51$, $p < .001$. Effect sizes were calculated for these differences and corrected with correlations between the groups' mean scores for text A and B. The calculations were conducted with Lenhard and Lendhard's (2016) online calculator, based on the formulas reported by Borenstein (2009). The effect size for the Vietnamese group is $d = .501$ and for the Chilean group is $d = .814$. These effect sizes indicate that the difference in scores between the two texts is of medium size for the Vietnamese group and is of large size for the Chilean group.

Across-group comparisons

In order to investigate whether Vietnamese or Chileans students performed significantly better on each text, two independent t-tests were conducted. Table 7 shows results of these analyses.

Table 7. Independent t-tests between Texts A and B

	t	df	p
Text A	2.82	73.0	p<.01
Text B	3.04	72.6	p<.01

Results from the independent t-tests show that there are significant differences between Vietnamese and Chilean students in their mean scores for text A ($p < .01$) and text B ($p < .01$). For text A about the Vietnamese New Year, Vietnamese students performed significantly higher than Chilean students, $t(73) = 2.82$, $p < .01$. For text B about the Chilean Independence Day, Chilean students significantly outperformed Vietnamese students, $t(72.6) = 3.04$, $p < .01$. Effect sizes were calculated and results show that the differences in performance were at medium-large values for both text A ($d = 0.71$) and text B ($d = .66$).

Analysis of questionnaire data

Internal consistency reliability of the two variables for familiarity was calculated using Cronbach's alpha (α) resulting in a coefficient of .80 for the variable "Familiarity with text A"

and .81 for the variable "Familiarity with text B" which, according to Dörnyei and Taguchi (2009), is a sufficient reliability level for scales of 3-4 items.

Descriptive statistics

Overall, as seen in Table 8, the mean for the variable "Familiarity" is higher in each group for the texts describing their own national holidays than the mean for familiarity with the text of the other culture. The Chilean group mean for the variable "Familiarity with text A" was $M=2.21$ whereas for "Familiarity with text B" was considerably higher $M= 3.77$. On the other hand, the mean for the Vietnamese students in the variable "Familiarity with text B" was $M=2.07$ whilst for "Familiarity with text A" was much higher with $M=3.87$. From questionnaire results, we can see that students reported higher familiarity with the text about their culture, and perceived this familiarity to be advantageous in comprehending the text and answering the questions.

Table 8. Descriptive statistics for the questionnaire results

Group		Familiarity with text A		Familiarity with text B	
	N	M	SD	M	SD
Vietnamese	41	3.87	.515	2.07	.384
Chilean	35	2.21	.466	3.77	.512

Correlations between students' perceptions of their familiarity with the texts and their scores. Results showed significant positive correlations between participants' familiarity with the text and their performance. Despite the fact that both groups showed high levels of familiarity with the text about the respective national holidays, the relationships to scores were weak. The correlation for the reported familiarity with text A and the scores the participants obtained in the corresponding section A was $r_s=.319$, $p<.01$. For students' perceived familiarity with text B and their scores in section B, the correlation was $r_s=.234$, $p = p<.05$. These results suggest that, to at least some extent, the more familiar a student perceived the text, the higher the score obtained in that test section, or from another perspective, the less familiar the text is perceived, the lower the score in that section.

Discussion

Results from the t-test analyses show that there are significant differences in how Vietnamese and Chilean students performed in the test. The overall conclusion is that the Vietnamese students did better on the text about the *Tet Nguyen Dan* and the Chileans achieved higher performance on the questions about *Fiestas Patrias*. The high effect sizes of these differences mean that the difference between the two groups is large. Since our test instrument was designed to achieve a similar level of difficulty for the two texts and questions, and our participants were selected to be at a similar level of proficiency, we can assume that the influencing variable is students' cultural background. The findings of the current study is congruent with previous research that has claimed that cultural background knowledge gives an advantage for reading comprehension (e.g. Alptekin, 2006; Dehghan & Sadighi, 2011; Erten & Razi, 2009; Lee, 2007; Li & Lai, 2012). It is noteworthy that for high proficiency learners, cultural background knowledge does not give as much of an advantage (e.g. Chan, 2003). However, the participants in the current study were at a low proficiency level (A2 - B1). We believe that due to their limited language resources, our participants had to rely more heavily on their cultural background knowledge.

There were significant but weak relationships between perceived cultural familiarity of participants and their final scores. A possible explanation of these results is that questionnaire answers were self-reported and may not be an accurate representation of their actual cultural knowledge. For example, a student reporting a high level of familiarity may have informed this because it is socially valued in their culture to have a good knowledge about their national traditions. In addition, the students may have had the impression that the text about their culture would be easier to answer. This may in part be due to enhanced confidence when presented with a foreign language text with more familiar cultural elements (e.g. Erten & Razi, 2009). Furthermore, the activation of background cultural knowledge only provides a certain amount of support, especially with lower proficiency learners (e.g. Chan, 2003; Hill & Liu, 2012). These correlation results, albeit not strong, support the results from the t-tests in that there can be a positive influence from background knowledge on students' scores.

16

A significant limitation of this research would be the length of the test. Due to limited time and resources, our final instrument, Version D, had only four multiple choice questions per text. We were not able to include items that tested for the full range of global and local processes involved in reading. We were only able to get a gauge of student proficiency levels through their performance on 13 items. Further research can be done on a wider scale with a larger population of participants of different backgrounds and proficiency levels. It is suggested that future research include other test instruments to ascertain actual, as opposed to perceived, cultural background knowledge.

Conclusion

The study has implications on research in validating existing language tests, especially those that are not L1 independent. Fairness review guidelines (for example by ETS) have stressed the representation of diverse ethnic groups in tests. However, in the process of ensuring diversity, test developers have to take care not to inadvertently advantage certain groups of test takers.

Apart from the implications in the field of second language testing, the results in this study can also have some implications in second language learning and teaching. In line with Alderson (2000), the results in our test and questionnaire suggested that more a familiar text would facilitate reading, comprehension and processing of the content in a text. Exposure to more familiar topics would also help students to recall information more easily (Erten and Razi, 2009; Lee, 2007). Familiarity can also benefit vocabulary learning since students can remember new words better when they associate them to content that is more significant (Lightbown & Spada, 2013), as in the case of the national holidays in our test. These considerations can be taken into account in terms the selection of teaching material for language classes. This does not mean that unfamiliar topics should be avoided; on the contrary, students need to be provided with a range of different topics that reflect the scenarios that they will face in real life when exposed to the second language. However, in the first stages of learning, when learners' proficiency is low, we propose that using culturally familiar topics can be a good strategy to reinforce learning process and the development of their language skills in L2, as well as to help promote learners' confidence and motivation.

In summary, the study investigated whether or not cultural background knowledge affects performance on reading comprehension tests. A text on a Vietnamese cultural event and a text on a Chilean cultural event were chosen for Vietnamese and Chilean participants to complete. Results showed that the participants, who were of the same proficiency levels, performed better on the respective texts. This conclusion supports previous literature that cultural background knowledge, especially for lower-proficiency students, has a significant impact on reading comprehension tests. The findings of this research also provide important implications for fairness in language testing and material selection in second language instruction.

References

Abbott, M. L. (2006). ESL reading strategies: differences in Arabic and Mandarin speaker test performance. *Language Learning*, *56*(4), 633 - 670.

Alderson, J. C. (2000). *Assessing reading*. Cambridge: Cambridge University Press.

Alptekin, C. (2006). Cultural familiarity in inferential and literal comprehension in l2 reading. *System: An International Journal of Educational Technology and Applied Linguistics*, *34*(4), 494 to 508.

Bernhardt, E. (1991). *Reading development in a second language*. New York: Ablex.

Borenstein, M. (2009). Effect sizes for continuous data. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.). *The Handbook of Research Synthesis and Meta Analysis (pp. 221-237)*. New York: Russell Sage Foundation.

Braasch, J.L., Goldman, S.R. (2010). The role of prior knowledge in learning from analogies in science texts. *Discourse Processes*, *47*(6), 447 to 479.

Chan, C. H. (2003). Cultural content and reading proficiency: A comparison of mainland Chinese and Hong Kong learners of English. *Language, Culture And Curriculum*, *16*(1), 60-69.

Dehghan, F., & Sadighi, F. (2011). On the cultural schema and Iranian EFL learners' reading performance: A case of local and global items. *Journal Of Pan-Pacific Association Of Applied Linguistics*, *15*(2), 97-108.

Dörnyei, Z., & Taguchi, T. (2009). *Questionnaires in second language research: Construction, administration, and processing*. Abington on Thames: Routledge.

Erten, I. H., & Razi, S. (2009). The effects of cultural familiarity on reading comprehension. *Reading in a Foreign Language*, *21*(1), 60-77.

Exam English Ltd. (2014). Cambridge English Preliminary (PET) Reading Part 4 Comprehension Questions. Retrieved from www.examenglish.com/PET/pet_reading4_test2.html

Hill, Y. Z. & Liu, O. L. (2012). Is there any interaction between background knowledge and language proficiency that affects TOEFL iBT reading performance?. *ETS Research Report Series, 2012*(2), i - 34.

Hughes, A. (2002) *Testing for language teachers.* Cambridge: Cambridge University Press.

Ketchum E. M. (2006). The cultural baggage of second language reading: An approach to

understanding the practices and perspectives of a nonnative products. *Foreign Language Annals, 39*(1), 22 - 42.

Lee, S. (2007). Effects of textual enhancement and topic familiarity on Korean EFL students' reading comprehension and learning of passive form. *Language Learning*, *57*(1), 87-118. DOI:10.1111/j.1467-9922.2007.00400.x

Lenhard, W., & Lenhard, A. (2016). Calculation of effect sizes. Retrieved from http://www.psychometrica.de/effect_size.html. DOI: 10.13140/RG.2.1.3478.4245

Li, C., & Lai, S. (2012). The functions of cultural schemata in the Chinese reading comprehension and reading time of college students in Taiwan. *Journal Of International Education Research*, *8*(2), 105-112.

Lightbown, P., & Spada, N. (2013). *How languages are learned*. Oxford: Oxford University Press.

Nassaji, H. (2002). Schema theory and knowledge-based processes in second language reading comprehension: A need for alternative perspectives. *Language Learning, 52* (2), 439-481.

Stott, N. (2004). Familiarity breeds contempt: reading texts from learners' own cultures does not guarantee recall. *TESOL Quarterly*, *38*(2), 345-352.

Henderson, H. (Ed). (2005). *Holidays, festivals, and celebration of the world dictionary*. Detroit, MI: Omnigraphics Inc.

Van Den Broek, P., Virtue, S., Everson M.G., Tzeng Y., & Sung Y.C. (2002). Comprehension and memory of science texts: Inferential processes and the construction of a mental representation. In A. Graesser, J. Otero, J. Leon (Eds.). *The Psychology of Science Text Comprehension (pp. 131 - 154).* Mahwah: L. Erlbaum Associates.

Van Den Broek, P., Young, M., Tzeng, Y., & Linderholm, T. (1999). The landscape model of reading: Inferences and the online construction of a memory representation. In H. Van Oostendorp & S.R. Goldman (Eds.). *The Construction of Mental Representations During Reading (pp. 71 - 98).* Abingdon on Thames: Psychology Press.