

Prejudicial Assessment? The Inherent Flaws of Reading Comprehension Test for EFL Students

Tikamporn Wuttiornpong¹

*Liberal Arts Faculty, Ubon Ratchathani University,
Ubon Ratchathani, Thailand*

Abstract

This paper presents and discusses an analysis of the validity of using L2 to test learners' L2 reading comprehension, a case study of a multiple choice test type. Participants were 242 Thai university students who took a foundation English course at a university in Thailand. They were asked to complete two sets of English reading comprehension tests in which both shared an identical reading passage. However, the two sets differed in that the questions and their coinciding choices came in an English version and a Thai version. Results showed that participants received significantly higher scores in the Thai version than in the English version. More specifically, students with low English proficiency benefitted more by taking the test in Thai (L1) while students with higher English proficiency did not receive any benefit. Furthermore, using L1 helped students when the question was detailed with multiple choice options containing long phrases, clauses or sentences. However, this was not the case in an inference question, or a detailed question with only one-word multiple choice options. Thus, this paper argues in favor of using L1 in multiple choice questions and their multiple choice options to assess learners' L2 reading comprehension as opposed to the presently existent practice of using the L2, which by itself frequently poses understanding difficulties for low proficiency students.

Key Words: L2 reading comprehension test; test validity; test bias; multiple choice test

¹ Email: tikamporn@gmail.com

บทคัดย่อ

บทความนี้อภิปรายการวิเคราะห์ความเที่ยงตรงของแบบทดสอบเมื่อใช้ภาษาที่สองในการทดสอบการอ่านภาษาที่สองเพื่อความเข้าใจของผู้เรียนกรณีศึกษาแบบทดสอบแบบปรนัยที่มีตัวเลือก ผู้เข้าร่วมในการวิจัยเป็นนักศึกษามหาวิทยาลัยจำนวน 242 คนที่เรียนรายวิชาภาษาอังกฤษพื้นฐานในมหาวิทยาลัยแห่งหนึ่งในประเทศไทย นักศึกษากลุ่มนี้ได้ทำแบบทดสอบการอ่านภาษาอังกฤษเพื่อความเข้าใจสองชุด เนื้อเรื่องที่อ่านในแบบทดสอบทั้งสองชุดเหมือนกันทุกประการ เว้นแต่ส่วนของคำถามและตัวเลือกเป็นภาษาอังกฤษในแบบทดสอบชุดหนึ่ง (ชุดภาษาอังกฤษ) และเป็นภาษาไทยในอีกชุดหนึ่ง (ชุดภาษาไทย) ผลปรากฏว่าผู้เข้าร่วมในการวิจัยสอบได้คะแนนสูงขึ้นอย่างมีนัยสำคัญในชุดภาษาไทย โดยเฉพาะอย่างยิ่งผู้เรียนที่มีความรู้ภาษาอังกฤษในระดับต่ำได้รับประโยชน์มากกว่าผู้เรียนที่มีความรู้ภาษาอังกฤษในระดับที่สูงกว่าเมื่อทำแบบทดสอบชุดภาษาไทย แต่อย่างไรก็ตามการใช้ภาษาที่หนึ่งช่วยผู้เรียนได้เมื่อคำถามเป็นคำถามเกี่ยวกับรายละเอียดเฉพาะที่มีตัวเลือกเป็นกลุ่มคำ อนุประโยค หรือประโยคที่ยาว แต่ไม่ช่วยเมื่อเป็นคำถามประเภทที่ต้องอนุมานหรือคำถามเกี่ยวกับรายละเอียดที่มีตัวเลือกเป็นคำๆ เดียว บทความนี้แสดงให้เห็นว่าควรใช้ภาษาที่หนึ่งในส่วนของคำถามและตัวเลือกในการทดสอบการอ่านภาษาที่สองเพื่อความเข้าใจแทนการใช้ภาษาที่สองซึ่งมักจะเป็นอุปสรรคต่อความเข้าใจสำหรับผู้เรียนที่มีความสามารถทางภาษาในระดับต่ำ

คำสำคัญ: แบบทดสอบการอ่านภาษาที่สองเพื่อความเข้าใจ ความเที่ยงตรงของแบบทดสอบ อดคิดของแบบทดสอบ แบบทดสอบแบบปรนัยที่มีตัวเลือก

Prejudicial Assessment? The Inherent Flaws of Reading Comprehension Test for EFL Students

Tikamporn Wuttiornpong

*Liberal Arts Faculty, Ubon Ratchathani University,
Ubon Ratchathani, Thailand*

1. Introduction

In terms of test validity, a test is supposed to test only those skills for which it is designed. But in a number of EFL contexts, where learners are not proficient in the target foreign language, this may not be the case. For example, in Thailand, multiple choice reading comprehension tests are common in universities and high schools, and students are expected to take English reading comprehension tests where the question and its coinciding choices are in English. Having to answer such imposes an additional level of complexity on students that is not directly related to the purpose of the test, which is to determine the students' comprehension of the reading passage rather than their comprehension of the questions and their coinciding choices. This *test within a test* is hypothesized to pose a greater burden on students with lower English proficiency than on their proficient classmates. Both the failure to accurately test students' comprehension of the passage and the resulting disparity in scores has bothered students as well as educators in Thailand.

The present study aims to investigate whether or not and to what extent using L2 (L2 here refers to second language and foreign language) to formulate the test questions may inaccurately measure students' comprehension of a reading text in the target language. The results of this study contribute some empirical evidence in an attempt to raise awareness of a test bias which has been overlooked in many ESL/EFL contexts.

2. Literature review

2.1 Reading comprehension and its processes

There are several concepts for reading comprehension. In the view of the bottom-up approach, reading comprehension is generally referred to as a linguistic decoding process which starts from analyzing linguistic units as small as individual letters to larger units, i.e. words, phrases, clauses and sentences to achieve meaning (Nunan, 1995, p. 33). The top-down approach, on the other hand, involves the construction of the textual meaning is a result of the interaction between the reader and the text, based on experience and world knowledge, language knowledge and expectations about how language works, interest, motivation, and attitudes toward the text, rather than on the decoding of the linguistic forms (Nunan, 1985). Finally, there is the interactive approach where reading comprehension is seen as an interactive process between accurate linguistic processing and previous experience which facilitates the further processing of added information to understand the text (Celce-Murcia, 1991).

Although meaning or comprehension is the fundamental goal of the reading process, reading comprehension is much more cognitively complex than one would assume, especially for less skilled L2 readers. Since we read for different purposes, we have different ways and strategies to read a text. In this light, Grabe and Stoller (2002, p. 4) propose that “reading comprehension is remarkably complex, involving many processing skills that are coordinated in very efficient combinations.” Thus, it should be rather seen as a type of expertise that develops over time.

Adopting similar view, this paper largely uses Grabe and Stoller’s (2002) framework for reading comprehension processes for L1 reading for discussion, as these authors explain that “at very advanced levels, L1 and L2 reading abilities tend to merge and appear to be quite similar’ (p. 11).

Thus, to have a more complete understanding about reading comprehension abilities, i.e. to the extent achieved by native-speaking readers or fluent and critical readers, it is essential to consult L1 reading processing and development. However, since there are differences between L1 and L2 reading, the differences are also covered and discussed.

In explaining reading comprehension components and processes for L1 reading, Grabe and Stoller (2002) divide the processes into two phases: lower-level processes and higher-level processes. The lower-level processes largely involve linguistic processes, which require linguistic knowledge, gradually developing over time. The higher-level processes involve the reader's use of background knowledge, making predictions, understanding inferences, etc. The labels, lower-level and higher-level, however, do not mean that the former can be mastered more easily than the latter. Instead, the former serves as the basis for the latter.

2.1.1 The lower-level processes

According to Grabe and Stoller (2002), the lower-level processes include (1) lexical access, (2) syntactic parsing, (3) semantic proposition formation, and (4) working memory activation. First, lexical access, or word recognition (Grabe, 2009), refers to the ability to recognize individual words. For fluent readers, this process is carried out subconsciously and virtually without effort as a result of frequent exposure and practice over time whereas L2 readers with limited vocabulary are likely to have difficulty in understanding L2 text. Lexical access or word recognition is so crucial that differing abilities in recognizing individual words can result in varied levels of comprehension (Perfetti, 2007, in Grabe, 2009). A number of studies over the past two decades point to word recognition as a major predictive factor for reading abilities (Adams, 1990, 1999; Perfetti, 1999, 2007; Perfetti, Landi, & Oakhill, 2005, in Grabe, 2009). Most researchers agree that fluent readers

perform “rapid and automatic word recognition of a large vocabulary” (Grabe, 2009, p. 23). In general, fluent readers recognize almost all words, approximately 80% of content words and 50% of function words (Adams, 1990; Perfetti, 1999; Pressley, 2006; Stanovich, 2000, in Grabe, 2009).

The word recognition processing takes place when readers perceive the written input and successfully activate lexical items in their mental lexicon, which contains well-established orthographic, phonological, semantic, and syntactic information for the stored lexical items or words. When facing word recognition difficulty, context may be of great help in identifying words (Perfetti, 1994, 1999; Perfetti & Hart, 2001; Stanovich, 2000, in Grabe, 2009). It should be understood that the reader’s well-established words in the mental lexicon may not always match the correct words in the target language system. Thus, word recognition difficulty may occur due to mismatches between the reader’s words and the correct words. Mismatches may even be caused by trivial differences at any of the four types of information mentioned.

Another process is syntactic parsing. A competent reader uses grammatical information to understand word groups at the phrase, clause and sentence levels as combinations of smaller units. In understanding a long text, the reader needs to be able to recognize word ordering information and a number of phrases and clauses. Syntactic parsing allows the reader to discern the correct meaning of words in different contexts. For example, ‘*well*’ in the sentence “The father ordered the son to get some water from the *well*” is a noun meaning ‘a deep hole to obtain water’, not an adverb, meaning in a satisfactory manner. Another example, the compound verb ‘have damaged’ is a present perfect form of ‘to damage’ and cannot be translated separately as ‘have’ meaning ‘to possess’ and ‘damaged’ meaning ‘to harm or break something physically’. An L2 reader who is not proficient enough tends to have

tremendous difficulty in processing syntactic parsing, and as a consequence, cannot decode the meaning of a long string of words or a long text. Findings from several studies indicate that sentence structures and grammatical features can cause readers to spend a noticeable amount of time on reading processing (Carpenter, Miyake & Just, 1994; Fender, 2001, in Grabe, 2009).

The third lower-level process is semantic proposition formation. In this process, the parsed words and their syntactic information are recombined to create semantic meaning units or ideas, generally realized at the phrase, clause and sentence levels. These semantic meaning units or ideas are referred to as semantic propositions. Word recognition and syntactic parsing do not operate separately, but at the same time to create semantic propositions, which are “the building blocks of text comprehension” (Grabe, 2009, p. 31).

The fourth and last process is working memory activation which can be explained as the time when the above three processes are active for a very short time. Grabe and Stoller (2002, pp. 24-25) explain the role of working memory in the reading process:

Working memory keeps information active for one to two seconds while it carries out the appropriate processes. [...] If processing of active information is not done quickly enough, the information fades from memory and must be reactivated, taking more resources and making the reading process inefficient.

This is why reading comprehension is a skill that requires practice, and why frequent exposure is necessary to reach automaticity of the lower-level processes for one to become a fluent reader.

In addition to the lower-level processes, which involve rather primarily linguistic mechanics, the reader also needs to see beyond the letters and words and process at a higher level, for example, possessing a purpose for reading, using reading strategies, understanding the main points (i.e. the model of the text), interpreting information in the text correctly, using background knowledge to make appropriate inferences, and evaluating success in comprehending texts. All of these are referred to as the higher-level processes.

2.1.2 The higher-level processes

The higher-level processes are what people generally think of when talking about reading comprehension. These processes are beyond the relationship of language mechanics, or the relationship between elements of letters and words. They include (1) text model of reading comprehension, (2) situation model of reader interpretation, (3) background knowledge use and inferencing, and (4) executive control processes.

The first and most fundamental higher-level process for reading comprehension is the text model of reading comprehension. It is basically an organization of ideas from the text which characterizes the main and supporting ideas. Once the reader finishes working out the lower-level processes and the meaning units at the clause-level are formed, more information is added, and a network of connected ideas develops. Repetitive ideas which form usable linkages to other information are noticed and discerned as the main ideas of the text. These repeated ideas become active and remain in the reader's network of ideas, whereas ideas which are not important to connect new information or to facilitate inferencing become less important and fade from the network. In this way, the reader starts to draw main ideas out of the added and active information in the process of reading. This set of the main ideas from the text is called the text model of

comprehension. Related to this process is the use of background knowledge, also referred to as world knowledge, schema theory, mental model, etc. Background knowledge helps the readers define the context, predict the text's discourse organization, and minimize disambiguation of the meanings of the words from the word-level up to the clause-level, when new information is inserted. In short, the text model is what the author wishes to convey to the reader, which when achieved, enables the reader to make a summary of the text.

While reading and creating the text model of comprehension, the reader starts predicting where the text is going, based on personal background knowledge, attitudes, inferences, goals, task and text difficulty, etc. That is, as soon as the building of text model begins, the reader also starts interpreting what is being read and creating a situation model of reader interpretation. In effect, the situation model justifies "how a reader can understand both what the author is trying to say [...] and how the reader interprets that information for his or her own purposes" (Grabe & Stoller, 2002, p. 28).

Elaborating further, to comprehend the text as a complete process, the fluent reader has to set the situation model of reader interpretation as the initial goal. To achieve this goal, the reader needs to integrate the information in the text with a well-developed network of ideas, supported by personal background knowledge. Background knowledge is important for the situational model because it helps draw correct inferences, generally essential for correct comprehension of the text. Without adequate background knowledge and correct inferences, misinterpretation may occur, resulting in comprehension failure. If there is a success in performing the situation model of reader interpretation, the reader is not only able to make a summary of the text but is also able to provide an appropriate critique on the text.

The fourth and final process is the executive control processing which helps assess how well the reader is comprehending the text. This process operates as a feedback system, monitoring comprehension accuracy, assessing and reassessing reading purposes, repairing comprehension problems, and evaluating the degree of the success in understanding the text being read to optimize comprehension.

All of these four higher-level processes work together, and the interplay occurs fast in working memory. However, these processes will not function efficiently and effortlessly if the texts are beyond the reader's comprehension ability. Sources of reading difficulties are likely to be inadequate linguistic knowledge, lack of background knowledge, and poor reading skills due to not reading enough (Grabe & Stoller, 2002).

At times during the reading, readers fail to correctly interpret certain parts of the text, resulting in comprehension problems. This is likely because they do not recognize the unstated ideas, i.e. the ideas they need to infer. To successfully make inferences is to, first, correctly connect elements in the text or maintain the text coherence, which is mandatory for comprehension. Causes of failure to make inferences are difficult to specify due to several factors, e.g. the different levels of the readers' lower-level processes and working memory capacity to make inferences. A study by Cain and Oakhill (1999, cited in Perfetti, Landi, & Oakhill, 2005) found an association between inference making ability and the readers' differences in comprehension skill, which is based on the readers' assessed level of comprehension, not their age. Yuill and Oakhill (1991, cited in Perfetti, Landi, & Oakhill 2005) offer three possibilities to explain differences in making inferences between skilled and less-skilled readers:

(1) Less-skilled comprehension readers cannot make correct inferences due to their insufficiency of background or world knowledge.

(2) Less-skilled comprehension readers cannot discern when they need to draw meaning through inferences.

(3) Limited processing capacity obstructs less-skilled comprehension readers from inferring the hidden message and integrating ideas in the text with prior knowledge.

In other words, L2 readers with insufficient background knowledge and limited processing capacity, regardless of their age, are not likely to know when it is appropriate to make inferences.

As stated earlier, Grabe and Stoller's framework is outlined for L1 and fluent reading development. It is a broad sketch designed to explicitly portray a complete process of the reading processes. Yet for L2 reading, there are unaccommodating conditions and limitations for reading development due to the numerous differences between L1 and L2 reading contexts, which are covered in the next part.

2.2 Differences between L1 and L2 reading

In addition to the framework for L1 reading comprehension processes, Grabe and Stoller (2002) propose that L2 reading differs from L1 reading in three main aspects: (1) linguistic and processing differences, (2) individual and experiential differences, and (3) socio-cultural institutional differences. These are discussed below.

2.2.1 Linguistic and processing differences between L1 and L2 readers

Linguistic and processing differences to be discussed include vocabulary, grammar, discourse, orthography, metalinguistic and metacognitive issues, and amount of exposure to L2.

2.2.1.1 Different amounts of lexical, grammatical and discourse knowledge

As Grabe and Stoller (2002) point out, when L1 readers learn to read, they already have tacit knowledge acquired from having learned the language orally for 4-5 years. Thus, prior to the time they learn to read, they have already acquired most of the basic grammar related to their L1 and a personal vocabulary ranging from 5,000 – 7,000 words. All of these linguistic resources provide the adequate foundation necessary for successful reading. On the contrary, many L2 learners do not have the same tacit knowledge or oral abilities as L1 readers do when they start to read in L2. Actually, many even begin to learn L2 orally at precisely the same time they start to read in L2, beginning with simple sentences and simple texts simultaneously. In addition, in many L2 settings, these learners even start to learn L2 grammatical structures in their reading texts. Thus, unlike L1 readers, L2 readers do not have adequate language knowledge of L2 prior to their learning to read, which is a major disadvantage because “one benefit of developing accurate letter-sound correspondences as a support for reading is lost” (p. 43). For L2 learners to succeed in reading comprehension, they need a good foundation of L2 language knowledge for a good start.

2.2.1.2 Metalinguistic and metacognitive awareness in L2 settings

It is widely observable that L2 learners do not have the same type of tacit knowledge L1 learners do. In L2 settings, learners largely gain knowledge about L2 through formal instruction. They also practice reading through learning tasks, projects, and external reading. In the classroom, learners discuss and reflect on L2 linguistic resources, including vocabulary, grammar, and discourse knowledge. Grabe and Stoller (2002) explain that through this method of learning, these learners have metalinguistic awareness, and often make

use of their metalinguistic knowledge to make sense of a reading text. Moreover, since L2 learners have already developed literacy skills and content knowledge from learning their L1, they have obtained metacognitive awareness in L1. Both metalinguistic and metacognitive awareness work together to help them find strategic support to understand the information in the text, and assess and identify comprehension failure (Grabe & Stoller, 2002; Paris & Winograd, 1990; Raphael & Pearson, 1985).

2.2.1.3 Difference in amounts of exposure to L2 reading in comparison to L1 reading

According to Grabe and Stoller (2002), a vital foundation for reading comprehension is a large amount of vocabulary and fluency in syntactic processing. However, most L2 learners simply do not read enough to get sufficient exposures to L2 written input and develop automaticity in word recognition and fluent syntactic processing (Koda, 1996; Lundberg, 1999, cited in Grabe & Stoller, 2002). As a result, they often have difficulties in comprehending L2 texts.

2.2.1.4 Linguistic differences

A major factor that may cause difficulty to L2 readers is linguistic differences between L1 and L2, which mainly include grammatical and orthography differences, and the role of cognates. Simply speaking, any L1 and L2 pair of languages which has substantial differences in the above mentioned aspects has a tendency to pose reading difficulty to L2 readers. Apart from that, L2 proficiency is an essential factor for being successful in L2 reading. According to the Language Threshold Hypothesis, which holds that L1 reading strategies and skills may be used to facilitate L2 reading only when L2 readers possess adequate L2 knowledge (i.e. grammar and vocabulary), therefore making L2 knowledge more crucial to reading fluency than L1 reading skills at an early stage. This is because beginning L2 readers have to initially use most of

their cognitive resources to understand vocabulary and language structures (the lower-level processes), which, in turn, leaves insufficient cognitive resources for the higher processes which help them read more strategically in order to decipher what is beyond words and grammar. To cross this threshold, L2 readers have to understand nearly all the words in order to process the text effortlessly.

Another issue related to the Language Threshold Hypothesis is that of transfer. L1 transfer in L2 reading can either be constructive, in cases where it supports L2 reading, or destructive, in cases where it interferes with L2 reading. The latter obviously posing the challenger to the reader. Beginning L2 readers, and sometimes even intermediate-level readers, equipped with inadequate L2 knowledge, are prone to rely heavily on their own world knowledge--L1 knowledge and L1 reading abilities when faced with difficult part in the reading text. Sometimes, L1 resources help them correctly understand the text, but at other times mislead them or impede the L2 reading processing (occurring during the situation model of reader interpretation). Some researchers argue that incompetent L2 readers may not benefit from their world knowledge and L1 reading abilities if they do not have adequate L2 knowledge to trigger L1 resources (e.g. Bossers, 1992; Carrell, 1991; Clarke, 1980; Taillefer, 1996).

Suggestion to circumvent this unwanted situation and to encourage reading enthusiasm is selecting reading texts which are not too difficult and are instead pleasant to read to encourage L2 readers to do more reading practice, enlarge vocabulary repertoire, and build L2 reading processing fluency. Consequently, L2 readers will rely less and less on L1 resources.

2.2.2 Individual and experiential differences in L2 reading

Grabe and Stoller (2002) contend that L2 readers with weak L1 literacy abilities are unlikely to transfer L1 resources to L2 reading contexts. They explain that these readers also have different motivational levels for reading L2 texts from depending on their self-esteem, emotional attitudes and responses to reading, interests in specific topics, to willingness to keep on reading and learn from the texts. In addition, text types are important in developing L2 reading ability in that simple texts do not allow L2 readers to experience the comparable cognitive-ability levels that L1 readers do. A rather interesting point here is that very difficult texts which may seem to be similar to authentic texts, are often found to be much shorter. Thus, L2 readers appear to encounter a narrower range of text genres and text lengths, when compared to those experienced by L1 readers. As a consequence, they have restricted exposure to a wider range of text genres that would otherwise expand their lexicon, and exercise their reading ability.

2.2.3 Socio-cultural institutional differences

Regarding socio-cultural differences in reading development, Grabe and Stoller (2002) point out some discrepancies between L1 and L2 reading comprehension by primarily asserting that, in most cases, reading L2 texts will be difficult in cases where framing assumptions in L2 texts rely on cultural assumptions which L2 readers have no clue of. Another discrepancy mentioned is the different conventions by which speakers of a particular language organize text discourse. This means that when the discourse organization in L2 texts does not match that of their L1, they will encounter a problem. Furthermore, in relation to institutional differences, Grabe & Stoller (2002, p.61) clearly state that, “distinction between L1 and L2 reading is shaped by different attitudes,

resources, and expectations of L1 and L2 educational structures” meaning that L2 reading is basically performed for different purposes from that of L1 reading. Students engaged in L2 reading engage in formal classroom settings where they read, for example, in response to a teacher’s instructions, and they do so for the purpose of taking exams rather than for mere acquisition of information and reception of communication as would be the case for L1 readers. In L2 settings, many L2 readers may have positive attitudes toward their L2 reading material in much the same way as they have for their L1 reading material. However, many others may have a completely negative attitude toward the same L2 reading material, viewing it simply as a utilitarian tool for learning L2. This may hinder them from engaging willingly to reading and therefore becoming fluent L2 readers.

In sum, although L1 and L2 readers undergo all lower-level and higher-level reading processes to reach the goal of reading increased comprehension ability, they do not experience the same conditions, given that L2 readers have a more significant number of limitations to deal with based on several factors in the L2 context.

The above discussion of the lower-level and the higher-level processes for reading comprehension has shown how reading comprehension is so complex that it requires multiple cognitive skills operating simultaneously. On top of that, differences between L1 and L2 reading pose additional obstacles for L2 readers. For fluent readers, these processes are rapid and almost effortless. However, for less competent L2 readers, reading comprehension can be very problematic.

When encountering comprehension problems, it is a common practice for many L2 readers to slow the spontaneity of the reading process by trying to translate the text into L1. They also tend to try forming a situation model of reader interpretation from past experiences and force the text to fit

their own preconceived notions. Unfortunately, both strategies result in meager comprehension, since working memory efficiencies will not function well in the case of translation process, and in the situation model the reader may access background information that does not correctly match the context (Grabe & Stoller, 2002). If these discouraging experiences continue, L2 readers are likely to lose the motivation necessary to become fluent. To prevent this from happening, the reading texts should be at a level appropriate to the reader's ability, and L2 readers have to read regularly for hours (Grabe & Stoller, 2002). Much evidence from studies on the effects of extensive reading on reading comprehension shows that a number of exposures to written input result in improved reading comprehension as well as reading speed (Bell, 2001), vocabulary enlargement (Grabe & Stoller, 1997; Horst, 2005; Pigada & Schmitt, 2006), and both expanded vocabulary repertoire and comprehension improvement (Stanovich, 2000).

Having discussed reading comprehension processes, factors leading to test bias will now be reviewed. Since the present study focuses on the test itself, test reliability is excluded, and only factors violating test validity are discussed.

2.3 Descriptions of test validity

Bachman (1991, p. 20) defines a test as “a measurement instrument designed to elicit a specific sample of an individual's behavior...[and]... necessarily quantifies characteristics of individuals according to explicit procedures”. As for language tests, they are expected to measure specific language abilities.

Test validity is directly related to the test itself and refers to “the appropriateness of a given test or any of its component parts as a measure of what it is purported to measure” (Henning, 1987, p. 89). It is generally considered the most

important quality of a test for interpretation of results and use (Bachman, 1991).

Bachman states further that a test will be valid if the inferences or decisions we make based on obtained test scores are applicable, meaningful, and useful. In principle, to ensure a test is applicable or valid and that the test scores are meaningful and useful, the test must not be chiefly affected by other factors than the ability being measured. Otherwise, that particular test will be considered invalid.

Traditionally, validity can be classified into several types, for instance, content, criterion, face, construct, and concurrent validity. For the purpose of the present study, which is investigating the validity of the test itself and not any external factors, only content validity is concerned and described below.

Hughes (1998, p. 22) proposes that a test is content valid if “its content constitutes a representative sample of the language skills, structures, etc. with which it is meant to be concerned.” In other words, a valid test is able to measure the particular knowledge, skills or abilities it is designed or intended for. For example, a grammar test must be able to elicit the test taker’s knowledge about specific grammar structures and not vocabulary knowledge.

As discussed above, reading comprehension requires several complex processes, which cause difficulties particularly for L2 readers. As for a reading comprehension test, which is supposed to measure the reader’s comprehension of the given text and nothing else, there are factors which can interfere with test validity. One possible influential factor is language difficulty in a multiple choice section of the test regarding both the question and its choice options, given in the target language of a reading comprehension test. Therefore, the present study has been designed to investigate whether and to what extent a multiple choice L2 reading comprehension test

using L2 in the comprehension question section it may affect the validity of the test.

3. Methods

3.1 Participants

Participants were 242 non-English major students from four sections of Foundation English II in the first semester of the academic year 2009 at a Thai University. They were of diverse English proficiency levels. For the purpose of analysis, they were divided into three groups according to their scores (out of 9 points from 9 test items) in the English version of the reading comprehension test (discussed in section 3.2): the low-score (scores from 1% up to 33%), mid-score (scores from 44% up to 66%), and high-score (scores above 66%).

3.2 Instruments

The instruments used in the present study were two sets of multiple-choice reading comprehension tests. In both sets, the reading text of 364 words was identical, and was about women and their improved political rights in terms of the increasing number of successful women in politics in many Arab countries (See Appendix). The passage was followed by ten comprehension questions (However, item 8 had a flaw in its multiple choice options; therefore, it has been discarded from the analysis), asking about the topic/main idea, details, references and interpretations/inferences. The passage and some test items were in fact taken from an actual final examination paper. However, questions asking for the meaning of a word or a phrase were taken out, as the students were allowed to ask for the meanings of any words in the passage, thus rendering these types of questions inappropriate. To compensate for questions that were taken out, some questions were designed by the researcher in the same manner as tests given for grades at the university.

The two sets of tests differed in that one set had comprehension questions and their multiple choice options in English (hereafter called the “English version”), whereas in the other they were in Thai (hereafter called the “Thai version”). The questions and the multiple choice options in the Thai version were translations of those in the English version. However, to avoid and/or minimize the possibility of students’ memorizing the letter of the answers for the multiple choice options in the English version, which was taken first, the order of the multiple choice options in the Thai version was shuffled. The questions and their type are summarized in Table 1 below.

Table 1. Summary of the questions and their type.

Question	Type of Question	Description of M/C
1. What is the topic of the passage?	Topic/main idea	Sentence
2. What problem did Mrs. Al-Gerifi have when she ran for an election in Qatar?	Detail	Sentence
3. What is NOT TRUE about Qatari women?	Detail	Sentence
4. What does “ <u>this</u> ” in P2 refer to?	Reference	Clause
5. What is TRUE?	Detail	Sentence
6. Which is TRUE about the election in Bahrain in 2002 in P4?	Inference/interpretation	Sentence
7. According to the passage, in what country are women the least successful in politics?	Detail	One word
9. What is the main idea of P4?	Topic/main idea	Sentence
10. In P5, which country has the smallest number of women in parliament?	Detail	One word

Note: M/C refers to ‘multiple choice’.

The two versions of the complete test are provided in the appendix.

3.3 Procedure and data analysis

The research procedure for the present study was carried out in two phases. First, students were asked to do the English version of the reading comprehension test in the final week of the semester in reading comprehension by means of grammar-translation instruction. Then in the consecutive session in the same week, the participants were required to take the Thai version. Each test session was 45 minutes long.

Before students started the test, the teacher of each class, who had received training prior to administering the test, gave them the meaning of difficult words in the text line by line. During the test, the students were not allowed to use a dictionary, but they were permitted to ask for the meaning of any word in the text. This was to minimize the possibility of students scores varying due to the difficulty of the vocabulary in the text and to further limit the variable to only the difference of the languages used in the questions and the multiple choices. However, students were not allowed to consult each other and no help was given on word meaning from the questions and the multiple choice options. Finally, none of the students needed extra time to complete the test.

After the data was collected, each group's mean scores were calculated. Then, paired t-tests were performed to determine whether there was any significant difference in scores between the English and the Thai versions for each group. Analysis also explored whether or not the Thai version could significantly increase the number of students who chose the correct answers for each item. Results are described and discussed in section 4 below.

4. Results and discussion

4.1 Overall results

The overall statistical results, including mean scores, standard deviations, percentage of gained scores in the Thai version compared to the English version, t-value, and p-values for all groups are summarized in the Table 2 (Table 2) below.

Table 2. Comparison of mean scores between the English and the Thai versions for all three groups.

Group	Number	English version		Thai version		Score gained %	T-value	P-value
		Mean	SD	Mean	SD			
High-score	(n = 7)	7.00	0.00	7.43	0.79	0.06	1.44	0.200
Mid-score	(n = 138)	4.75	0.73	5.22	1.61	9.89	3.39	0.001*
Low-score	(n = 97)	2.46	0.65	4.44	12.23	80.49	12.23	0.000*

* $p < 0.01$

** $p < 0.001$

As summarized in Table 2, the paired t-tests showed an insignificant difference between the English version and the Thai version for the high-score group at $p > 0.05$ ($t = 1.44$, $p = 0.200$). Their mean scores for the English and the Thai versions are almost the same, with the gained score of 0.43 points or only 0.06%. These results indicate that it made no difference whether the questions and the multiple choice options were in English or Thai for the students in this group. On the other hand, there was a significant variation effect in scores for the mid-score group between the English version and the Thai version at $p < 0.01$ ($t = 3.39$, $p = 0.001$), and an even greater effect for the low-score group at $p < 0.001$ ($t = 12.23$, $p = 0.000$). Based on the paired t-test results, both the mid and the low-score groups performed significantly better in the Thai version.

The scores gained for each group showed that the Thai version favored the low-score group the most, followed by the mid-score group, with the high-score group not benefitting at all. Thus, it can be concluded that the less proficient students tested scored significantly lower when faced with questions and multiple choice options in L2.

4.2 Significant effects of using L1 in place of L2 on students' performance

Section 4.2 explores whether there is any significant difference in the percentage of students who chose the correct answers between the English and the Thai versions for each item. This is to determine significant effects of using L1 in the questions and the multiple choice options that may significantly helped these Thai students improve their scores. Table 3 presents the paired t-test results of comparison of percentages of students in each group who selected the correct answer between the Thai and the English versions for each item.

From Table 3, the question items which show significantly higher percentages of students who selected the correct answer in the Thai version than in the English version are Q1², Q2, Q3, Q4, Q5, and Q9 for the low-score group, and Q2 and Q5 for the mid-score group. The high-score group did not make any significant difference at all. Among these items, Q2, Q3, Q4, Q5 are detail questions with lengthy multiple choice options while Q1 and Q9 are topic/main idea questions with long complex noun phrases and long complex clauses, respectively. The statistical results of comparison of percentages of students in each group who chose the correct answers between the Thai and the English versions confirm the overall results presented in section 4.1 that the low-score group

² Q refers to 'question', thus Q1 refers to 'question 1'.

benefited the most from the Thai version , the mid-score group benefited less, while the high score group did not benefit at all in any of the items (Table 3).

Table 3. Comparison of percentages of students in each group who selected the correct answer between the Thai and the English versions for each item.

Q item	Group	English		Thai		T-value	P-value (2-tailed)
		M	SD	M	SD		
Q1	High (n = 7)	0.86	0.38	0.43	0.54	2.121	0.078
	Mid (n = 138)	0.20	0.40	0.29	0.46	-1.745	0.083
	Low (n = 97)	0.12	0.33	0.29	0.46	-3.161	0.002**
Q2	High (n = 7)	0.86	0.38	0.86	0.38	0	1.000
	Mid (n = 138)	0.86	0.35	0.93	0.25	-2.269	0.025***
	Low (n = 97)	0.59	0.50	0.99	0.10	-8.034	0.000*
Q3	High (n = 7)	0.71	0.49	0.86	0.38	-1.000	0.356
	Mid (n = 138)	0.54	0.50	0.58	0.50	0.699	0.486
	Low (n = 97)	0.20	0.40	0.42	0.50	-3.598	0.001**
Q4	High (n = 7)	0.57	0.54	0.86	0.38	-1.549	0.172
	Mid (n = 138)	0.49	0.50	0.57	0.50	-1.328	0.186
	Low (n = 97)	0.15	0.36	0.52	0.50	-6.120	0.000*
Q5	High (n = 7)	1.00	0.00	1.00	0.00	0	1.000
	Mid (n = 138)	0.49	0.50	0.76	0.43	-5.175	0.000*
	Low (n = 97)	0.16	0.37	0.60	0.50	-7.185	0.000*
Q6	High (n = 7)	0.71	0.49	0.86	0.38	-1.000	0.356
	Mid (n = 138)	0.26	0.44	0.28	0.49	-0.315	0.753
	Low (n = 97)	0.09	0.29	0.13	0.34	-0.942	0.348
Q7	High (n = 7)	0.86	0.38	0.86	0.38	0	1.000
	Mid (n = 138)	0.78	0.41	0.70	0.46	1.643	0.103
	Low (n = 97)	0.62	0.49	0.61	0.49	0.148	0.882
Q9	High (n = 7)	0.86	0.38	0.86	0.38	0	1.000
	Mid (n = 138)	0.44	0.50	0.47	0.50	-0.491	0.624
	Low (n = 97)	0.15	0.36	0.40	0.49	-4.36	0.000*
Q10	High (n = 7)	0.57	0.54	0.86	0.38	-1.549	0.172
	Mid (n = 138)	0.68	0.47	0.64	0.48	0.787	0.433
	Low (n = 97)	0.37	0.49	0.48	0.50	-1.654	0.101

* $p < 0.001$, ** $p < 0.01$, *** $p < 0.05$

In reference to Q1, Q2, Q3, Q4, Q5, and Q9, the most likely reason why a greater and more statistically significant number of students in the low-score and mid-score groups chose the correct answer in the Thai version may be due to one quality these items have in common, that is, the multiple choice options are lengthy and, thus, complicated for many of these L2 students. Therefore, using L1 in the questions and multiple choice options increased the numbers of students selecting the correct answer in those questions. The Thai version seemed to have been helpful for these two lower score groups in the sense that it brought out the information these students had decoded from the L2 text. Consequently, because the mid-score students had a higher proficiency, they were able to decode more information in the L2 text than the low-score students did. That is, the multiple choice question section given in L1 serves as ‘true’ multiple choice question to check students’ text comprehension according to what they have decoded from the L2 text, unlike the question section given in L2 which behaves like a test within a test. However, L1 in the multiple choice question section is certainly not helpful since if students cannot decipher the information in the text written in L2. On the contrary, the high-score group did not receive any benefit from the Thai version in any multiple choice question item which likely indicates that they had sufficient knowledge of L2 to help them understand both the L2 text and the L2 questions, therefore, the presence of the L1 question section was of no use to them in providing better answers.

4.3 Insignificant effects of using L1 in place of L2 on students’ performance

Based on the statistical results in Table 3, the question items which do not show any significant difference in percentages of students choosing the correct answer between the Thai and the English versions are Q6, an inference

question, and Q7 and Q10, two detail questions. Obviously, the questions and the multiple choice options in Q7 and Q10 are at one-word level while those in Q6 are phrases of four to eight words. Hence, detail questions Q7 and Q10 will be explored first, followed by inference question Q6.

As described in section 3, the questions and the multiple choice options in the Thai version are simply the translations of those in the English version. Thus, only examples of the English version are given for discussion.

Question 7

Q7 is given below for discussion.

Q7. According to the passage, in what country are women the least successful in politics?

- a. Morocco b. Bahrain c. Algeria d. Lebanon

The key sentences where the answer lies in are “*In Bahrain, for example, 39 women ran for local and national office in 2002. Not a single woman won.*” According to the passage, elsewhere, women either won some seats or were appointed in political offices, but no women got elected in Bahrain. Therefore, the correct choice is (b). Q7 is not problematic for the students because it asks for a short and straightforward answer, the name of the country which has the smallest number of woman politicians. The sets of information for each country described in the text were not very complicated and basically consisted of names and numbers. The key statement “Not a single woman won” does not seem difficult for the majority of these students, as 62% of the low-score group, 78% of the mid-score group, and 86% of the high-score group selected the correct answer in the English version. Most importantly, the multiple choice options contained easy one-word vocabulary items, the names of the countries, which were easy to understand.

Question 10

Q10 is given below for discussion.

10. In P5, which country has the smallest number of women in parliament?

- a. Saudi Arabia b. Jordan c. Morocco d. Algeria

Like Q7, Q10 asks for a short and straightforward answer. However, the sentences in the text are more complex due to the word choices and sentence structures used. The key sentences are:

Jordan set a quota to ensure that at least six women were elected to parliament.

Morocco and Algeria have relatively high numbers of women in parliament compared to the rest of the region. And even Saudi Arabia, which gives women fewer political rights than any other countries on earth, recently had two women elected to the board of directors of the Jeddah chamber of commerce.

The sentences above provide information, in simpler words, that Jordan had at least six women, Morocco and Algeria had quite a few women compared to other Arabic countries, and Saudi Arabia had none in its parliament, but two in a chamber of commerce. As none or zero is a whole real number, thus, the answer is option (a). Apparently, the sentences in paragraph 5 are more difficult than those in paragraph 4, which may be why noticeably fewer students in all groups scored on this item, compared to Q7 in the English version, i.e. 37% of the low-score, 68% for the mid-score, and 57% for the high-score groups. It is seen that the mean score of the high-score group is lower than that of the mid-score group, which is not expected. This was most likely due to the very small number of students in this group ($n = 7$), thus when only

three students chose the wrong answer, the percentage dropped dramatically.

From Table 3, the statistical results show a insignificant difference in the number of students in all groups who selected the correct answers between the English and the Thai versions. As discussed in Q7, this was most likely due to the fact that all of the multiple choice options were simple one-word vocabulary items, so there was no room for L1 to better improve the students' understanding of the multiple choice options.

Based on the results and discussion for Q7 and Q10 above, it can be concluded that using L1 is not very helpful for detail questions where the information in the multiple choice options is restricted to simple one-word level vocabulary items, e.g. the names of countries. This is most likely because such type of multiple choice options does not require much cognitive work.

Question 6

Q6 is given below for discussion.

Q6. Which is **TRUE** about the election in Bahrain in 2002 in P4?

- a. More women ran for the election than men.
- b. Thirty-nine women won the election.
- c. Only male candidates were elected.
- d. One woman was elected.

Item Q6 was intended to check the students' understanding of details by asking them to identify the correct choice from multiple choice options giving wrong information. The key sentences in the passage are "*In Bahrain, for example, 39 women ran for local and national office in 2002. Not a single woman won*", the correct option being (c) '*Only male candidates were elected*'. Since the answer must be inferred

from the two key sentences, specifically the latter, however, this question does not only elicit the students' understanding about what the words explicitly say but also their ability to interpret or infer what is not explicitly stated. Mean score results revealed that while 71% of the high-score group chose the correct answer in the English version, the numbers of the students in the mid-score and the low-score groups who did the same were far fewer at 26% for the mid-score and 9% for the low-score groups. There was no significant increase in number of students who chose the correct answer in the Thai version for the mid-score and the low-score groups, suggesting that using L1 in the multiple choice question section of the test did not help.

The fact that most of the mid-score and nearly all of the low-score students chose the wrong option (b) is noteworthy. These students were most probably lured by the number 39 which is the exact same number of the women who ran for the election, not who won the election. This finding suggests that when the students could not rely on their understanding of the text, they tended to rely on and seek help from some "clue" words instead. It must be noted that the text does not use a word for the male gender, e.g. 'men' or 'male'. This might have caused students forget or overlook the fact that there were also male candidates. Or else, the lack of "clue" word explicitly written in the text might have made them feel unsure whether there were also male candidates "out there" and decided to exclude them. They did not seem to make use of their background or world knowledge by analytical thinking or inferring that there were also male candidates and only male candidates won the election. Therefore, being unable to use background knowledge, the students failed to make a correct interpretation or inference of the information in the text (Yuill & Oakhill, 1991).

The assumption that the students did not seem to use their background knowledge to infer the unstated information suggests that even a simple higher-level process like using one's background knowledge, which is in fact readily available from life's experience, may not function adequately if the lower-level processes are not successfully operational. Based on the Language Threshold Hypothesis, it is only when L2 readers have sufficient L2 knowledge that they can adopt L1 reading strategies and skills to help comprehend L2 texts, as they spend most of their cognitive resources to figure out vocabulary and grammatical structures, leaving insufficient cognitive resources to understand and interpret the information or ideas represented by words and sentences. To pass this threshold, L2 readers have to know almost all of the words encountered in order to process the text effortlessly. In addition, the above finding indicates that students who could not tap into their knowledge of grammar and vocabulary, i.e. the lower-level processes, to help them understand what they were reading tended to resort to other strategies, in particular lexical form matching.

Given that the students were provided with the contextualized meanings in Thai for any of the English words in the passage, this result from item 6 also stresses that lexical access or the word recognition process is not sufficient to succeed in reading comprehension, and raises the importance of the role of grammar knowledge and the parsing process. This is because readers need grammar knowledge of the L2 to syntactically parse the recognized words to understand them before they can recombine them to form meaningful mental propositions or ideas (Grabe & Stoller, 2002).

On the whole, what renders many students in both lower score groups incapable of understanding the text well enough to draw correct inferences in both English and Thai test versions is likely due to an interplay of several factors. These

mainly including their limitations in operating the lower-level processes involving linguistic resources, limited working memory capacity (Cain & Oakhill, 1999; Grabe & Stoller, 2002; Yuill & Oakhill, 1991), not knowing when it is appropriate to make an inference, and/or incapacity to make adequate use of background knowledge (Yuill & Oakhill, 1991). The causes of these limitations appear to be due to linguistic and processing differences, unsupportive L2 learning environments (covering individual and experiential differences, and socio-cultural institutional differences), and poor reading skills as a result of not reading enough (Grabe & Stoller, 2002).

In summary, the results from the present study show that using L1 in the questions and multiple choice options benefits low-score students the most, followed by mid-score students, with not benefit to those in the high-score group. The results also suggest that using L1 in place of L2 may be most helpful to elicit L2 readers' comprehension of L2 texts at times when the questions simply ask for details, and the multiple choice options contain long and complex phrases, clauses or sentences. However, this does not seem to help incompetent L2 readers when the questions prompt them to draw an inference, as doing so requires them to grasp almost all the words in the text of which they are already deficient in, and to have sufficient cognitive resources to trigger and confirm the background knowledge to correctly interpret what is not explicitly stated. Likewise, using L1 is not much beneficial for questions where the information in the multiple choice options is restricted to short simple words, that do not require much cognitive work.

5. Conclusion

The present study has demonstrated that using L2 in the questions and the multiple choice options in L2 reading

comprehension tests can lower L2 students' genuine performance on L2 reading comprehension ability. The study has pointed out while students with high proficiency in L2 are not much likely to be affected, those with low proficiency suffer greatly. In other words, the less proficient a student is, the more they suffered. It is then argued that when a reading test intends only to target reading comprehension in L2 students, everything else on the test should be in L1 except for the L2 reading text so that no additional degree of difficulty will be encountered. As such, a reading comprehension test will not violate content validity. Having presented these results, it is hoped that the findings may raise awareness of a test bias that has been overlooked in several L2 and EFL contexts, and that these test biases be properly attended to.

As the findings from the present study suggest, multiple choice questions and their coinciding choice options in L2 reading comprehension tests should be presented in L1. However, in a classroom context where this may be considered unacceptable, translations or glossaries of difficult or unfamiliar words in the questions and multiple choice options that do not appear in the passage should be provided. In addition, words and structures in the questions and multiple choice options should be limited to only those appearing in the text books and learning materials to ensure that at least they are not completely new to students. However, the problem may still persist as some learned words may not have been internalized and thus continue to remain difficult for students.

Biodata

Dr. Tikamporn Wuttiornpong holds her Ed.D. in Teaching English to Speakers of Other Languages (TESOL) from the University of Western Australia. She is a lecturer at

the Liberal Arts Faculty, Ubon Ratchathani University, Thailand. She teaches undergraduate courses, and supervises postgraduate students in the TEFL program. Her research interests include L2 reading comprehension and development, L2 writing development, L2 acquisition of tense and aspect, and L2 vocabulary acquisition.

References

- Adams, M. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.
- Adams, M. (1999). Afterword: The science and politics of beginning reading practices. In J. Oakhill & R. Beard (Eds.), *Reading development and the teaching of reading* (pp. 213–227). Oxford: Blackwell.
- Bachman, F. L. (1991). *Fundamental considerations in language testing*. Oxford University Press.
- Bell, T. (2001). Extensive reading: speed and comprehension. *The Reading Matrix 1*. Retrieved from <http://www.readingmatrix.com/articles/bell/index.html>
- Bossers, B. (1992). *Reading in two languages: A study of reading comprehension in Dutch as a second language and in Turkish as a first language*. Rotterdam: The Netherlands: Drukkerij Van Driel.
- Cain, K., & Oakhill, J. V. (1999). Inference ability and its relation to comprehension failure in young children. *Reading and Writing, 11*, 489–503.
- Carpenter, P., Miyake, A., & Just, M. (1994). Working memory constraints in comprehension: Evidence from individual differences, aphasia, and aging (pp. 1075–1122). In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 849–94). San Diego: Academic Press.

- Carrell, P. (1991). Second language reading: Reading ability or language proficiency? *Applied Linguistics*, 12, 159–179.
- Celce-Murcia, M. (1991). *Teaching English as a second or foreign language*. USA: Heinle & Heinle Publishers.
- Clarke, M. A. (1980). The short-circuit hypothesis of ESL reading—or when language competence interferes with reading performance. *Modern Language Journal*, 64, 203–209.
- Fender, M. (2001). A review of L1 and L2/ESL word integration skills and the nature of L2/ESL word integration development involved in lower-level text processing. *Language Learning*, 51, 319–396.
- Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. New York, NY: Cambridge University Press.
- Grabe, W., & Stoller, F. L. (1997). Reading and vocabulary development in a second language: A case study. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 98–122). Cambridge University Press.
- Grabe, W., & Stoller, F. L. (2002). *Teaching and researching reading*. Edinburgh: Pearson Education.
- Haris, M., & Hatano, G. (Eds.). (1999). *Learning to read and write: A cross-linguistic perspective*. New York, NY: Cambridge University Press.
- Henning, G. (1987). *A guide to language testing*. Cambridge, MA: Newbury House.
- Horst, M. (2005). Learning L2 vocabulary through extensive reading: a measurement study. *The Canadian Modern Language Review*, 61, 355–382.
- Hughes, A. (1998). *Testing for language teachers*. New York, NY: Cambridge University Press.
- Koda, K. (1996). L2 word recognition research: A critical review. *Modern Language Journal*, 80, 450–460.

- Lundberg, I. (1999). Learning to read in Scandinavia. In M. Harris & G. Hatano (Eds.), *Learning to read and write: A cross-linguistic perspective* (pp. 157–172). New York, NY: Cambridge University Press.
- Nunan, D. (1985). Content familiarity and the perception of textual relationships in second language reading. *RELC*, 16(1), 43-51.
- Nunan, D. (1995). *Designing tasks for the communicative classroom*. New York, NY: Cambridge University Press.
- Paris, S. G., & Winograd, P. (1990). How metacognition can promote academic learning and instruction. In B. F. Jones & L. Idol (Eds.), *Dimensions of thinking and cognitive instruction* (pp. 15–51). Hillsdale, NJ: Erlbaum.
- Perfetti, C. (1994). Psycholinguistic and reading ability. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics*. (pp. 849–894). San Diego, CA: Academic Press.
- Perfetti, C. (1999). Comprehending written language: A blueprint for the reader. In C. Brown & P. Hagoort (Eds.), *Neurocognition of language* (pp. 167–208). New York, NY: Oxford University Press.
- Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, 11, 357–383.
- Perfetti, C., & Hart, L. (2001). The lexical basis of comprehension skill. In D. Gorfien (Ed.), *On the consequences of meaning selection* (pp. 189–213). Amsterdam: John Benjamins.
- Perfetti, C., Landi, N., & Oakhill, J. (2005). The acquisition of reading comprehension skill. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook*. (pp. 227-247). Oxford: Blackwell.
- Pigada, M., & Schmitt, N. (2006). Vocabulary acquisition from extensive reading: A case study. *Reading in a ForeignLanguage*, 18(1). Retrieved from <http://www.nflrc.hawaii.edu/rfl>
- Pressley, M. (2006). *Reading instruction that works* (3rd ed.). New York: Guilford Press.

- Raphael, T. E., & Pearson, P. D.(1985). Increasing student's awareness of sources of information for answering questions. *American Educational Research Journal*, 22, 217–235.
- Snowling, M. J., & Hulme, C. (Eds.). (2005). *The science of reading: A handbook*. Oxford: Blackwell.
- Stanovich, K. (2000). *Progress in understanding reading: Scientific foundations and new frontiers*. New York: Guilford Press.
- Taillefer, G. F. (1996). L2 reading ability: Further insight into the short-circuit hypothesis. *The Modern Language Journal*, 80, 461–477.
- Yuill, N., & Oakhill, J. (1991). *Children's problems in text comprehension*. New York, NY: Cambridge University Press.

Appendix

Reading Comprehension Test (English version)

Name..... Number..... Section.....

Instructions: Read the text below and choose the best answer.

P1 In some parts of the world, women have almost no voice in politics and government. When Sheikha Yousef Hasan Al-Gerifi was running for city council in Qatar, her family did not allow her to put pictures in the campaign advertisements. Her family also refused to let her put pictures in newspapers and on posters.

P2 The tiny Persian Gulf nation of Qatar is a conservative place. As in neighboring Saudi Arabia, most Qatari women cover themselves completely in public, including their faces. So, because of this, it is really no surprise that Mrs. AL Gerifi's family would object to the idea of her photo being put on campaign ads. In the end, it did not matter. She won the election anyway.

P3 The fact that she was able to run for, and be elected to, public office is a sign of how things are changing in the most conservative corner of the globe. But that change is coming slowly. Political scientist Hala Mustafa of Egypt al-Ahram Foundation says women in the Arab world have little political power, and few Arab countries have any significant number of women in parliament.

P4 Although women may experience fewer difficulties in Lebanon, Morocco and Algeria, women in much of the Arab world have a very hard time getting elected. In Bahrain, for example, 39 women ran for local and national office in 2002. Not a single woman won. In Egypt's parliament election this year, only a very small

number of candidates were women, and four of them were elected. The president increased the overall number of women in parliament by giving them five of the 10 appointed seats after the election.

P5 However, some changes are happening. Two years ago, Jordan set a quota to ensure that at least six women were elected to parliament. Morocco and Algeria have relatively high numbers of women in parliament compared to the rest of the region. And even Saudi Arabia, which gives women fewer political rights than any other countries on earth, recently had two women elected to the board of directors of the Jeddah chamber of commerce. It is a small step, but in the eyes of many Arab women, a very important one.

1. What is the topic of the passage?
 - a. A better understanding of women in political positions
 - b. Qatari women and their political rights
 - c. Number of Arab women in political positions
 - d. Arab women and their better opportunities in politics

2. What problem did Mrs. Al-Gerifi have when she ran for an election in Qatar?
 - a. Her family did not allow her to place her photos on the campaign ads.
 - b. Her family did not allow her to run for the city council election.
 - c. She was not accepted by male candidates.
 - d. She was not allowed to appear in public places.

3. What is **NOT TRUE** about Qatari women?
 - a. They cannot show their faces in public.
 - b. They have equal political power to men.
 - c. They live in a conservative country.
 - d. They are allowed to run for an election.

4. What does “**this**” in P2 refer to?
- a. that most Qatari women cover up their bodies and faces outside homes
 - b. that Qatar is a conservative country where married women cover their bodies.
 - c. that most Qatari women are not surprised her photos are not shown on the ads
 - d. that Qatar is influenced by a more powerful country, Saudi Arabia
5. What is **TRUE**?
- a. Newspapers did not want Mrs. AL-Gerifi’s pictures.
 - b. Mrs. AL-Gerifi succeeded in the election.
 - c. The politics in Mrs. Al-Gerifi’s country rapidly changes.
 - d. Mrs. Al-Gerifi did not allow anyone to take her pictures.
6. Which is **TRUE** about the election in Bahrain in 2002 in P4?
- a. More women ran for the election than men.
 - b. Thirty-nine women won the election.
 - c. Only male candidates were elected.
 - d. One woman was elected.
7. According to the passage, in what country are women the least successful in politics?
- a. Morocco b. Bahrain c. Algeria d. Lebanon
9. What is the main idea of P4?
- a. Women in Lebanon, Morocco and Algeria are more successful than those in other Arab countries.
 - b. Women in Bahrain and Egypt have more difficulty in getting elected than those in Lebanon, Morocco and Algeria.
 - c. There should be a quota for women in the parliaments in Arab countries.

d. It is very difficult for women in most Arab countries to win an election.

10. In P5, which country has the smallest number of women in parliament?

- a. Saudi Arabia b. Jordan c. Morocco d. Algeria

Reading Comprehension Test (Thai version)

Name..... Number..... Section.....

คำชี้แจง: อ่านเนื้อเรื่องข้างล่างนี้ แล้วเลือกคำตอบที่ดีที่สุด

P1 In some parts of the world, women have almost no voice in politics and government. When Sheikha Yousef Hasan Al-Gerifi was running for city council in Qatar, her family did not allow her to put pictures in the campaign advertisements. Her family also refused to let her put pictures in newspapers and on posters.

P2 The tiny Persian Gulf nation of Qatar is a conservative place. As in heighboring Saudi Arabia, most Qatari women cover themselves completely in public, including their faces. So, because of this, it is really no surprise that Mrs. AL Gerifi's family would object to the idea of her photo being put on campaign ads. In the end, it did not matter. She won the election anyway.

P3 The fact that she was able to run for, and be elected to, public office is a sign of how things are changing in the most conservative corner of the globe. But that change is coming slowly. Political scientist Hala Mustafa of Egypt al-Ahram Foundation says women in the Arab world have little political power, and few Arab countries have any significant number of women in parliament.

P4 Although women may experience fewer difficulties in Lebanon, Morocco and Algeria, women in much of the Arab world have a very hard time getting elected. In Bahrain, for example, 39 women ran for local and national office in 2002. Not a single woman won. In Egypt's parliament election this year, only a very small number of candidates were women, and four of them were elected.

The president increased the overall number of women in parliament by giving them five of the 10 appointed seats after the election.

P5 However, some changes are happening. Two years ago, Jordan set a quota to ensure that at least six women were elected to parliament. Morocco and Algeria have relatively high numbers of women in parliament compared to the rest of the region. And even Saudi Arabia, which gives women fewer political rights than any other countries on earth, recently had two women elected to the board of directors of the Jeddah chamber of commerce. It is a small step, but in the eyes of many Arab women, a very important one.

1. หัวข้อของเรื่องนี้คืออะไร

- ก. ผู้หญิงกาตาร์กับสิทธิทางการเมือง
- ข. จำนวนผู้หญิงในตำแหน่งทางการเมือง
- ค. ความเข้าใจที่ดีขึ้นเกี่ยวกับผู้หญิงที่อยู่ในตำแหน่งทางการเมือง
- ง. ผู้หญิงอาหรับกับโอกาสทางการเมืองที่ดีขึ้น

2. นาง Al Gerifi มีปัญหาอะไรเมื่อเธอลงหาเสียงเลือกตั้งในกาตาร์

- ก. เธอไม่ได้รับการยอมรับจากผู้สมัครรับเลือกตั้งที่เป็นผู้ชาย
- ข. เธอไม่ได้รับอนุญาตให้ปรากฏตัวในที่สาธารณะ
- ค. ครอบครัวของเธอไม่อนุญาตให้เธอลงภาพของเธอบนสื่อโฆษณาหาเสียง
- ง. ครอบครัวของเธอไม่อนุญาตให้เธอลงหาเสียงเลือกตั้งสภาเมือง

3. ข้อใดไม่ถูกต้องเกี่ยวกับผู้หญิงกาตาร์

- ก. ผู้หญิงกาตาร์มีอำนาจทางการเมืองเท่าเทียมกับผู้ชาย
- ข. ผู้หญิงกาตาร์ได้รับอนุญาตให้หาเสียงเลือกตั้งได้
- ค. ผู้หญิงกาตาร์ไม่สามารถเปิดเผยใบหน้าในที่สาธารณะได้
- ง. ผู้หญิงกาตาร์ดำรงชีวิตอยู่ในประเทศอนุรักษ์นิยม

4. คำว่า “this” ในย่อหน้าที่ 2 หมายถึงข้อใด

- ก. การที่ผู้หญิงกาตาร์ส่วนใหญ่ไม่ได้รับรู้สิทธิประหลาดใจที่ภานาง Al Gerifi ไม่ปรากฏบนสื่อโฆษณาหาเสียง
- ข. การที่ผู้หญิงกาตาร์ส่วนใหญ่ปกปิดร่างกายและใบหน้าเมื่ออยู่นอกบ้าน

- ค. การที่กาดาร์ได้รับอิทธิพลจากประเทศซาอุดีอาระเบียที่มีอำนาจมากกว่า
- ง. การที่กาดาร์เป็นประเทศอนุรักษนิยมที่ซึ่งหญิงที่แต่งงานแล้วต้องปกปิดร่างกาย

5. ข้อใดถูกต้อง

- ก. นาง Al Gerifi ไม่อนุญาตให้ใครถ่ายภาพของเธอ
- ข. หนังสือพิมพ์หลายฉบับไม่ต้องการลงภาพของ นาง Al Gerifi
- ค. การเมืองในประเทศของ นาง Al Gerifi เปลี่ยนแปลงไปอย่างรวดเร็ว
- ง. นาง Al Gerifi ประสบความสำเร็จในการเลือกตั้ง

6. ข้อใดถูกต้องเกี่ยวกับการเลือกตั้งในบาห์เรนในปี 2002 ในย่อหน้าที่ 4

- ก. มีผู้สมัครรับเลือกตั้งที่เป็นผู้ชายเท่านั้นที่ได้รับการเลือกตั้ง
- ข. ผู้หญิง 39 คน ชนะการเลือกตั้ง
- ค. มีผู้หญิงเพียงคนเดียวที่ได้รับการเลือกตั้ง
- ง. มีจำนวนผู้หญิงลงหาเสียงเลือกตั้งมากกว่าผู้ชาย

7. จากเรื่องที่อ่าน ในประเทศใดที่ผู้หญิงประสบความสำเร็จทางการเมืองน้อยที่สุด

- ก. เลบานอน ข. โมร็อกโก ค. บาห์เรน ง. ออสเตรเลีย

8. ข้อใดเป็นใจความสำคัญของย่อหน้าที่ 4

- ก. น่าจะมีโควตาสำหรับผู้หญิงในรัฐสภาของประเทศอาหรับต่าง ๆ
- ข. ผู้หญิงในเลบานอน โมร็อกโก และออสเตรีย ประสบความสำเร็จมากกว่าผู้หญิงในประเทศอาหรับอื่น ๆ
- ค. มันเป็นเรื่องที่ยากลำบากยิ่งสำหรับผู้หญิงในประเทศอาหรับส่วนใหญ่ที่จะชนะการเลือกตั้ง
- ง. ผู้หญิงในบาห์เรนและอียิปต์มีความยากลำบากมากกว่าผู้หญิงในเลบานอน โมร็อกโก และออสเตรีย ในการได้รับการเลือกตั้ง

9. ในย่อหน้าที่ 5 ประเทศใดมีจำนวนผู้หญิงในรัฐสภาน้อยที่สุด

- ก. โมร็อกโก ข. จอร์แดน ค. ออสเตรเลีย ง. ซาอุดีอาระเบีย