Validity of Translation Tests as a Measure of Reading Comprehension

Hiroyuki OKAZAKI

E-mail: hokazaki@edu.u-toyama.ac.jp

Abstract

Identifying the main idea of an expository text is one of the most important skills in reading comprehension (Tomitch, 1996; 2000). In English entrance examinations of many Japanese national universities, expository texts are commonly used to measure reading comprehension. However, a translation-into-Japanese test method (TIJ-T) is still employed as a measure of reading comprehension despite being a target of criticism. This study attempts to examine to what extent this method can measure main idea reading comprehension. Two experiments have revealed that a TIJ-T is not valid as a means for measuring main idea comprehension of an expository text. Additionally, the findings have also indicated that some test takers can identify a main idea but cannot translate it into Japanese, while others who cannot identify a main idea (do not comprehend the text) are still able to provide a literal translation.

keywords：Reading Comprehension, Main Idea Reading Comprehension, Translation Ability, Test Validity

1. Introduction and Literature Review

A translation-into-Japanese test (TIJ-T) is a well-known method of measuring comprehension of a text. However, the reliability and validity of translation as a testing technique to measure reading comprehension have been questionable thus far. Shizuka (2006) claims that translation tests in university entrance exams are one of the main factors hindering the improvement of English education in Japan, particularly in high school. Nevertheless, many national universities adopt this testing technique for screening high school student applicants.

The relationship between translation and reading comprehension ability has been investigated and declared an invalid measure for reading comprehension (Aoki, 2000; Minowa, 2006; Ushiro et al, 2005; Watanabe, 1994). So far research has focused on the sentence-level relationship between translation and comprehension. In most reading situations, however, readers usually comprehend a paragraph as a whole, not as separate sentences. Reading comprehension does not always represent the combined comprehension of each single sentence. Reading comprehension reflects what a writer wishes to convey in a paragraph. Theoretically, translation of a single sentence can reveal only a partial understanding of the paragraph. The primary purpose of this study, therefore, is to find out how much of a paragraph's main idea can be understood through the translation of a single sentence.

According to the *Course of Study for Foreign Languages* (MEXT, n. d.), students should be able to obtain necessary information, make an outline, *summarize the main points* and *understand the writer's intentions* when reading texts (emphasis added). Summarizing main points and understanding the writer's intentions are essential skills in teaching reading at the high school level in Japan. Furthermore, research has indicated that the ability to identify the main idea in an expository text is a fundamental reading comprehension skill (Just & Carpenter, 1992; Tomitch, 1996, 2000). To the researcher's knowledge, no previous studies have been conducted to investigate the relationship between translation and comprehension of the main idea at the paragraph level.

The secondary purpose of this study is to

investigate whether different types of test takers exist. Regarding the shortcomings of translation tests, Aoki (2000), Minowa (2006), Sekine (1994) and Watanabe (1994) identified two types of test takers: those who can comprehend a sentence but cannot translate it into Japanese, and those who can translate the sentence literally but not comprehend its meaning. However, all of these articles investigated the relationship between translation and reading comprehension through a sentence-level interpretation, not a paragraph-level. As Williams (1984) claimed, the ability to identify the main idea in an expository text is a fundamental skill of comprehension. None of these articles used the main idea identificatin as a measure of reading comprehension. In this study, therefore, it is hypothesized that some test takers will fall into these groups: (A) those who can understand the writer's intention but who cannot translate the main idea into Japanese; (B) those who cannot identify the main idea but can provide a literal translation into Japanese.

The present study, therefore, addresses the following research questions (RQs).

RQ1: To what extent can translation tests measure reading comprehension of the main idea at the paragraph level?
RQ2: Do the following two groups really exist?
(A) Those who identify the main idea of a paragraph but cannot translate it into Japanese.
(B) Those who cannot identify the main idea of a paragraph but still provide a literal Japanese translation.

2. Experiment 1

This experiment attempts to investigate the relationship between translation ability and main idea reading comprehension at the paragraph level.

2.1 Method
2.1.1 Participants

A total of 65 Japanese national university students participated in this study. 42 were sophomores, 17 were juniors, and six students were seniors. All were education majors.

2.1.2 Materials

To determine reading comprehension, an expository text from the 2008 National Center Test was used, a passage consisting of seven paragraphs with seven multiple-choice questions. Except for the third question, each question was designed to determine if test takers could identify the main idea of each paragraph. Question three focused on a peripheral detail and was therefore excluded. Answers from the remaining six questions were used for this study (see Appendix A). For a translation test, five sentences from this same text were chosen by five experienced high school English teachers (see Appendix A). Each teacher looked for sentences that seemed to reflect the author's main ideas and intentions.

2.1.3 Procedure, Scoring, and Data Analysis

Study participants first took the reading comprehension test (RC-T). After a week, they were asked to translate five underlined sentences into Japanese as a test of translation ability. In the RC-T, each question was worth one point, for a top score of six. The reliability of the test was .40. Thus caution is needed when interpreting the results.

In scoring the TIJ-T, the researcher and a high school English teacher marked the answers independently using based on model answers (inter-rater = 0.86). The reliability (Cronbach's alpha) of the test was .87. The scores from the two raters were combined and divided by two to calculate the mean score.

To answer RQ1, a Pearson product-moment correlation was calculated between test scores in translation and reading comprehension in order

to examine the validity of translation as a reading comprehension test. To examine RQ2, it was necessary to compare the scores obtained in the two tests. The participants were divided into three groups (using the criterion of $M \pm \frac{1}{2} SD$) as shown in Table 1 and 2: high, middle, and low groups according to the rank order of the scores obtained in each test. The high group in the RC-T was defined as "participants with high reading comprehension" and the lower group as "participants with low reading comprehension". The high group in the TIJ-T was defined as "participants with high translation ability" and the low group as "participants with low translation ability". It was examined whether participants in the middle or low group of the TIJ-T belonged to the high group in the RC-T. In addition, whether the low reading comprehension group fell into the high or middle groups of the TIJ-T was also looked at.

Table 1. *Descriptive Statistics for RC-T groups*

Group	Count	Mean	S.D
High	31	5.32	0.48
Middle	13	4.00	0.00
Low	21	2.48	0.75
Total	65	4.14	1.37

Table 2. *Descriptive Statistics for TIJ-T groups*

Group	Count	Mean	S.D
High	22	18.68	2.11
Middle	23	13.93	1.24
Low	20	7.40	1.80
Total	65	13.53	4.89

2.2 Results and Discussion

Tables 3 and 4 indicate the descriptive statistics and correlation coefficients between the two tests. A Pearson product-moment correlation between translation and RC-T scores yielded a moderate correlation value of $r = .435$. The variation overlap between the two was $r^2 = 0.435^2 = 0.189$. Thus the overlapping variation in the TIJ-T scores accounted for a very low 18.9%.

Table 3. *Descriptive Statistics of the Two Tests (N=65) (Full Score: RCT=6, TIJT=25)*

	Mean	SD
Reading Comprehension (main idea) Test	4.14	1.37
Translation-into- Japanese Test	13.53	4.89

Table 4. *Pearson's Correlation Coefficients Reading Comprehension (main idea) Test*

Translation-into- Japanese Test	.435**
	(65)

Correlation is significant at the .01 (2-tailed)

Accordingly, it can be concluded that translation tests are not valid as a reading comprehension measure. In addition, the ability to translate the main idea of a paragraph is quite different from the ability to simply identify it, a finding in line with Aoki (2000).

The results were not surprising. In the RC-T, test takers seemed to summarize the author's intentions, understand the entire paragraph, and then choose the answer that most closely matched their summaries. In contrast, in the TIJ-T, test takers tried to produce a literal Japanese translation of each given sentence, utilizing their overall knowledge of vocabulary and sentence structures. The ability to identify a main idea was not assessed here. In this fashion, the mental processes involved in taking each test were likely quite different.

As seen in Figure 1 and Table 5, there were four (6.1%) and ten (15.4%) participants who fell into the high comprehension group but were assigned to the low and middle groups of the translation test. In total, 21.5% of the participants with high reading comprehension ability did not translate the main idea into Japanese. In other words, they failed the translation test despite having comprehended the text.

Conversely, there were three (4.6%) and six (9.2%) participants from the low comprehension group who ended up in the middle and high groups of the translation test. In other words, 13.8% of the participants were still able to translate the main ideas of each paragraph despite
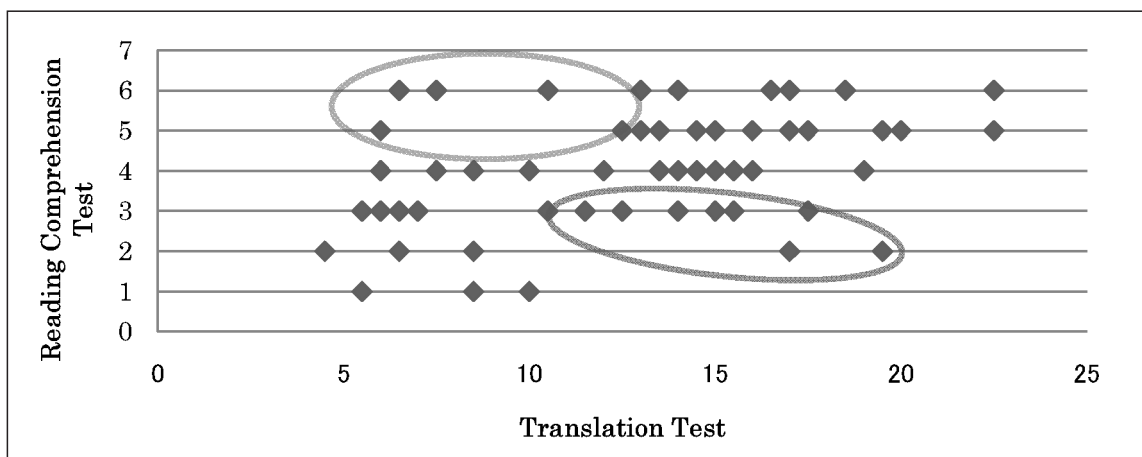
Figure 1. Scatter plot showing the relationship between the two test scores (*r* = .435)

Table 5. *Comparison between RC Test and TIJ Test*

RCT evaluation		TIJ-T evaluation			Total
		high	middle	low	
	high n % in total	17	10	4	31
		26.2%	15.4%	6.1%	47.7%
	middle n % in total	2	7	4	13
		3.1%	10.8%	6.1%	20.0%
	low n % in total	3	6	12	21
		4.6%	9.2%	18.5%	32.3%
	TIJ-T Total % in total	22	23	20	65
		33.9%	35.4%	30.7%	100%

not completely understanding them. These students would have likely passed the translation test despite their lack of comprehension.

3. Experiment 2

In the reading comprehension test in Experiment 1, the test takers seemed to just choose the best answer from the four given choices and did not identify or construct the main idea on their own. In this second experiment, students were asked to summarize what they felt the writer intended to say in each paragraph. Such summarizing tests provide a better evaluation of comprehension ability (Minowa, 2006). Ideally the summaries would be written in English, but this was not always possible due to low English ability of some study participants. Since the aim of this test was to determine if participants could simply articulate themain ideas in their own words, summaries written in Japanese were permitted.

3.1 Method
3.1.1 Participants
A total of 94 first-year Japanese university students with high school level English ability participated in this study. 43 were education majors and 51 majored in science.

3.1.2 Materials
As a test to measure main idea reading comprehension, the final part of the reading section of the 2009 National Center English Test was adopted for this experiment (see Appendix B). The text consisted of eight paragraphs and seven multiple-choice questions. The students were required to construct a main idea statement in Japanese on four paragraphs chosen by the researcher. As a translation test, the same material was used. The test takers were required to translate into Japanese the four underlined

sentences that were chosen to reflect each paragraph's main idea (see Appendix B).

3.1.3 Procedure, Scoring, and Data Analysis

First, the participants took the main idea reading comprehension test (MIC-T). After a week, they were asked to translate four sentences into Japanese.

In the MIC-T, each written statement was worth two points, for a top possible score of eight points. Statements that reflected clear understanding of each paragraph's main idea were given higher marks. The researcher and a high school English teacher scored the statements independently based on the model answers made by the researcher (inter-rater = 0.86). The reliability of the test was .625.

In scoring the translation-into-Japanese test (TIJ-T), the researcher and a high school teachers marked answers independently based on model answers (inter-rater = 0.93). The reliability (Cronbach's alpha) of the test was .837. Each translation question was worth four points, for a total top score of 16.

A Pearson product-moment correlation was calculated between the TIJ-T and MIC-T scores. To answer RQ2, the participants were divided into three distribution groups of high, middle, and low (using the criterion of $M \pm \frac{1}{2} SD$) as shown in Tables 6 and 7.

Table 6. *Descriptive Statistics for MIC-T groups*

Group	Count	Mean	S.D
High	28	5.91	0.96
Middle	34	3.99	0.40
Low	31	2.24	0.74
Total	93	3.98	1.63

Table 7. *Descriptive Statistics for TIJ-T groups*

Group	Count	Mean	S.D
High	28	13.87	0.96
Middle	38	10.91	1.10
Low	27	6.04	2.12
Total	93	10.38	3.37

3.2 Results and Discussion

Tables 8 and 9 indicate the descriptive statistics and correlation coefficients between the two test scores. As shown in Table 7, a Pearson product-moment correlation between translation and reading comprehension test scores yielded a low value (r=.387). This suggests a weak correlation between the two. The value was squared in order to obtain a variance overlap of $r^2=0.387^2=0.150$. This means that only 15.0% of the variation among translation scores can be accounted for by the variation in main idea comprehension. In other words, it cannot be concluded that the two tests measured the same kind of ability. Therefore, a TIJ-T cannot be considered as a valid tool for evaluating main idea comprehension.

Table 8. *Descriptive Statistics of the Two Tests (N=93) (Full Score : MIC-T=8, TIJ-T=16)*

	Mean	SD
Main Idea Comprehension Test	3.98	1.63
Translation-into-Japanese Test	10.38	3.37

Table 9. *Correlation Coefficients Main Idea Comprehension Test*

Translation-into- Japanese Test	.387**
	(93)

Correlation is significant at the .01 (2-tailed)

Figure 2 and Table 10 show the relationship between the two test scores. There were eight (8.6%) and three (3.2%) participants who fell into the high group in MIC-T but who were assigned to the low and middle groups of the translation test. In total, 11.8% of the participants were able to identify a main idea but did not translate it into Japanese. These students would likely fail the translation test despite being able to construct the main idea in a text.

Conversely, there were seven (7.5%) and twelve (12.9%) participants who fell into the low group in the MIC-T but who were assigned to the high and middle groups in the translation test. In total, 20.4% of the participants actually did not comprehend the main idea in the text yet were
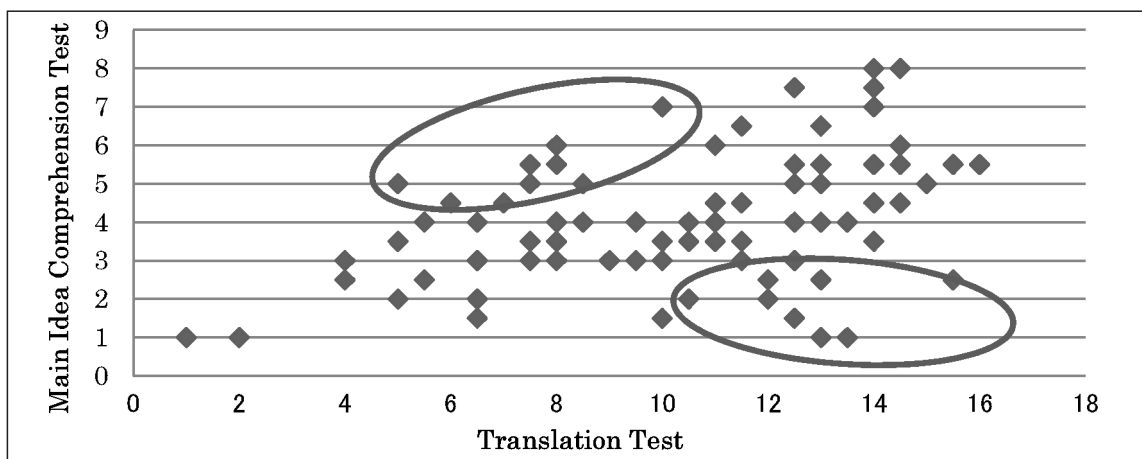
Figure 2. Scatter plot showing the relationship between the two test scores (*r*=.387)

Table 10. *Comparison between MIC-T and TIJ- T*

MIC Test evaluation		TIJ Test evaluation			Total
		high	middle	low	
	high n % in total	17	3	8	28
		18.3%	3.2%	8.6%	30.1%
	middle n % in total	9	16	9	34
		9.7%	17.2%	9.7%	36.6%
	low n % in total	7	12	12	31
		7.5%	12.9%	12.9%	33.3%
	TIJ-T Total % in total	28	38	27	93
		35.5%	33.3%	31.2%	100%

still able to provide a Japanese translation. These students would most likely pass the translation test even though they are unable to comprehend a main idea.

4. Conclusion and Implications

The primary purpose of this study was to investigate the validity of translation tests as a means for measuring comprehension of main ideas in paragraphs. The results of the two experiments suggest that such tests are not valid tools for identifying or constructing the main idea of the paragraph in an expository text. The secondary purpose was to confirm the presence of two types of test takers. The translation test produced the following types of examinees: those who could identify a main idea but do not translate it into Japanese (type A), and those who did not comprehend the main idea but still produced a literal Japanese translation (type B).

This finding suggests that A types may fail the translation test despite understanding the writer's intention. B types may pass the translation test, but they actually do not understand the main idea in a paragraph.

No one can dispute that identifying the main idea of a paragraph is an essential skill when reading an expository text. Reading comprehension tests should assess to what degree the main idea has been understood. However, simple translation tests do not seem to accomplish this task. Instead, they assess whether the examinee can make a literal translation from English to Japanese. In other words, these translation tests measure the ability to comprehend individual words and grammatical patterns but not the deeper meanings the writers are trying to convey. It is nearly impossible to measure whether an examinee can comprehend a paragraph's main idea through translating only one sentence from it. If the purpose of the test is to measure

just whether an examinee has knowledge of basic sentence structures and lexis, it would be sufficient. As it is, it is unclear why long passages are used for reading comprehension tests.

It is not within the scope of the present study to suggest an ideal reading comprehension test. The multiple-choice and open-ended tests employed in this study to extract main ideas in a paragraph also have some limitations as mentioned in this paper. Therefore, further study is urgently needed to help design a test that measures the ability to extract the main idea from a paragraph more accurately.

References

Aoki, Y. (2000). The validity of translation task as a reading comprehension measure. *The Bulletin of the Kanto-Koshin-Etsu English Language Society,14,* 35-42.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 98*, 122-149.

Ministry of Education, Culture, Sports, Science and Technology [MEXT]. (n.d.). The course of study for foreign languages. Retrieved October, 8, 2009, from http://www.mext.go.jp/english/shotou/030301.htm

Minowa, M. (2006). Relationship between Japanese high school students' translations and comprehension: to what extent learners' translation reflects comprehension. *Journal of Education Research in United Graduate School of Education, Tokyo Gakugei University, 13,* 107-123.

Sekine, R. (1994). Testing technique named "translation." *LEO, 23,* 73-97.

Shizuka, T. (2006). Korede iinoka nyuushi eigo mondai: eigo kyouiku oyobi tesuto riron no tachibakara[Is it all right with English entrance examinations for university? From the standpoint of English education and test theory]. *The Rising Generation, 4,* 2-5.

Tomitch, L. M. B. (1996). Individual differences in text organization perception and working memory capacity. *Revista da ANPOLL,* v.2, 71-93.

Tomitch, L. M. B. (2000). Teaching Main Ideas: Are We Really 'Teaching? *Linguagem & Ensino, 3* (1), 45-53.

Ushiro, Y., Hijikata, Y., Shimizu, M., In'nami. Y., Kasahara, K., Shimoda, A., Mizoshita, H., & Sato, R. (2005). Reliability and Validity of Translation Tests as a Measure of Reading Comprehension. *ARELE, 16*, 71-80.

Williams, J. P. (1984). Categorization, Macrostructure, and Finding the Main Idea. *Journal of Educational Psychology,* 76, 5, 874-879.

Watanabe, N. (1994). Validity of translation task as reading comprehension measure. *LEO, 23,* 141-171.

Appendix A: Examples of Translation-into Japanese Test
and Reading Comprehension Test
2008 年大学入試センター試験問題(英語筆記)の第 6 問を一部加工して使用

\<Example of Translation-into Japanese Test\>

下線部を和訳しなさい。(translate the underlined sentences into Japanese)

(1) My niece, Ann, is in her third year at university. She has recently started her job search. When she entered the university, she wanted to be an architect and planned to apply for work in an architectural firm. (A) But as she prepared for her job search, she learned that the way people work has changed a lot in the last few years. She discovered that much of the change has occurred because of what is called the IT Revolution. The "IT (Information Technology) Revolution" refers to the dramatic change in the way information is perceived and used in today's world.

＜省略＞

(4) Another example of this type of arrangement is reliance on overseas call centers, which have become increasingly common.   (D) It has become possible for a telephone operator in India to answer a customer-service call from anywhere in the world, respond directly to the customer and offer a satisfactory solution, at a far lower cost to the company than ever before.   Many large companies now depend on such call centers. Today in Japan, when you call a toll-free number, there is a chance that someone in Chingtao, China will answer the phone in Japanese to help solve your problem.

\<Example of Reading Comprehension Test\>

問 1 Ann realized that people today work in ways unimagined before because
  ┌─────┐
  │ 45 │ .
  └─────┘

  ① the amount of data on the Internet can no longer be dealt with effectively

  ② the demand for architects has risen sharply over the last couple of decades

  ③ they think the impact of the IT Revolution is already a thing of the past

  ④ they deal with information entirely differently than in the past

Appendix B:　Examples of Main Idea Reading Comprehension Test
2009 年大学入試センター試験問題(英語筆記)の第 6 問を一部加工して使用

＊(1)~(5)のパラグラフで<u>筆者の言いたいこと</u>を日本語で簡単にまとめなさい。
　<Main Idea Reading Comprehension Test>

＊下線部(a)~(e)を和訳しなさい。
　<Translation-into Japanese Test>

(1)　When l first entered the university, my aunt, who is a professional translator, gave me a new English dictionary.　I was puzzled to see that it was a monolingual dictionary, which meant that everything was in English.　Although it was a dictionary intended for learners, none of my classmates had one and, to be honest, I found it extremely difficult to use at first. I would look up words in the dictionary and still not fully understand the meanings.　I was used to the familiar bilingual dictionaries, in which the entries are in English and their equivalents are given in Japanese. I really wondered why my aunt decided to make things so difficult for me. (a) <u>Now after studying English at university for three years, I understand that monolingual dictionaries play a crucial role in learning a foreign language.</u>

　<省略>

　(5) Then, if bilingual dictionaries are so useful, why did my aunt give me a monolingual dictionary? As I found out, there is, in fact, often no perfect equivalence between words in one language and those in another. My aunt even goes so far as to claim that a Japanese "equivalent" can never give you the real meaning of a word in English! (e) <u>Therefore she insisted that I read the definition of a word in a monolingual dictionary when I wanted to obtain a better understanding of its meaning.</u> Gradually, I have come to see what she meant.