

High-Stakes Tests and Educational Language Policy in Japan: An Analysis of the National Center English Test from the Viewpoint of Education for International Understanding

Shigeru Ozaki
Takushoku University
Japan

TESOL Journal
Vol. 2, pp. 175-188
©2010
<http://www.tesol-journal.com>



Abstract

Deepening international understanding is one of the most important educational language policies for high schools in Japan. However, the possibility of the negative washback effect of university English entrance examinations on the implementation of the policy has been raised. Therefore, this study analyzed the National Center Tests—the highest-stakes form of university entrance examinations—from the viewpoint of education for international understanding since washback is generated by test content. A careful consideration was given to both interrater reliability and inner reliability throughout the analysis. The results of this study indicated that only a few particular topics and nations appeared in the tests. This result suggests that high school students might be able to develop only a part of their international understanding. However, those who attempt to go on to national or municipal universities need to take individual national/municipal university tests as well. Therefore, the content of these tests also has to be analyzed. Furthermore, the objectives of university English entrance examinations should be analyzed from the viewpoint of education for international understanding

Keywords: high stakes-tests, washback, entrance examinations, test content analysis, educational language policy, education for international understanding

Introduction

The purpose of this study is to analyze university English entrance examinations, which are the highest-stakes foreign language tests in Japan, in terms of topics related to education for international understanding, which is one of the most important educational language policies for secondary schools in Japan.

The possibility that negative washback from university entrance examinations is hindering such policy has been raised (refer to Hosoya, 1999; Kubota, 1996), although the deepening of high school students' international understanding is crucial for secondary school English language education. University entrance examinations in Japan are considered as achievement tests because both the Center for National University Entrance Examinations and individual universities announce that their tests cover some of the high school

English subjects listed in the Course of Study, which is a nationwide educational guideline for high schools. Therefore, international understanding should be one of the most important constructs of university English entrance examinations. This means that if entrance examinations do not measure applicants' international understanding, they lack content validity.

Concerning the importance of entrance examinations for the successful implementation of educational language policy, Kaplan and Baldauf (2003) and Gorsuch (1999) also pointed out such tests are one of the key issues in Japan. However, there has hardly been any content validity research on entrance examinations in terms of education for international understanding with an established analysis scheme. The result of this study will contribute to the improvement of the Center Test and consequently the successful implementation of education for international understanding.

In this study, Japanese terms, such as persons' names, place names, or book titles, are given in Roman letters, and the way of description follows how they are spelled in *Hiragana* (the Japanese cursive syllabary) so that readers will be able to find them easily when they need to refer to the original Japanese words that are mostly spelled in Chinese characters. However, when Japanese proper nouns are given in Roman letters in the original source, the spellings follow the original spellings.

Education for international understanding

Education for international understanding, which was originally advocated by UNESCO, was introduced into Japan in the 1950s under the name of *kokusai rikai kyouiku* in order to implement internationalization, and in 1989 the term *international understanding*, or *kokusai rikai* appeared for the first time in the Course of Study, which is a set of nationwide educational guidelines for both primary and secondary schools. Since then, cultivating and deepening students' international understanding has been one of the most crucial goals of foreign language education in Japan (refer to Monbushou, 1989, 1999; Mochidzuki, Kubota, Iwasaki & Ushiro, 2001; Muranoi, Chiba & Hatanaka, 2001; Niizato, 2000).

In this study, education for international understanding refers to *kokusai rikai kyouiku* rather than the original form of education for international understanding because it has developed in the Japanese context, and consequently it has become very different from the original form. Furthermore, although its definitions have been diverse and ambiguous (refer to Kosaka, 1999; Ozaki, 2009; Satou, 1995; Satou, 2001; Yoneda, Otsu, Tabushi, Fujiwara & Tanaka, 1997), the definition of education for international understanding in this study is simply to cultivate and deepen students' international understanding in order not to confuse readers.

University English entrance examinations in Japan

There are three major types of university in Japan: national, municipal (prefectural or city-run), and private universities. In 2005 there were 709 universities in total: 87 national (624,389 students), 80 municipal (122,864 students), and 542 private (2,062,042 students) universities (Monbukagakusyou, 2005).

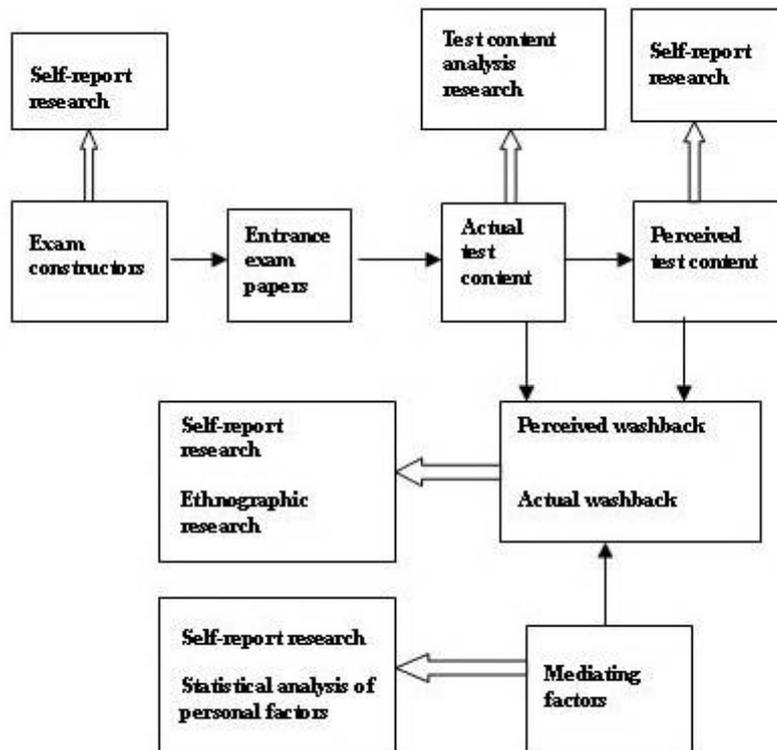
These universities offer different types of entrance examinations: conventional academic entrance examinations on various subjects, recommendation examinations, or admission office (AO) examinations. Recommendation examinations usually consist of essay writing tasks and/or interview tests together with recommendations from high schools. AO examinations have various forms; for example, some universities require their applicants to submit self-recommendation essays and evaluate their achievement in high school club activities and different types of qualifications. Additionally, AO and recommendation examinations are occasionally supplemented by tests in selected school subjects.

The most common entrance examinations can be classified into two types: the National Center Test and tests created by individual universities. The National Center Test is a type of standardized test developed by the National Center for University Entrance Examinations (NCUEE) and is used by national/municipal universities for their first-stage screening process as well as by some private universities. In general, applicants to national and municipal universities are required to take second-stage entrance examinations created and administered by each individual university.

Washback and test content analysis

Washback is defined as the influence of tests on education; when the result of a test is considered important, the test is likely to generate strong washback (Hughes, 2003). The original source of washback is test content (refer to Figure). For example, before Watanabe (1997) conducted ethnographical washback research in terms of translation tasks, he analyzed the content of Japanese university entrance examinations in order to confirm the tests included such tasks. It is essential to analyze the content of target tests before washback research is conducted especially when specific washback is investigated. Specific washback refers to “a type of washback that relates to only one specific aspect of a test or one specific test type” (Watanabe, 2004, p. 20).

Figure 1
Model of Washback Generation Process and Research (Ozaki, 2009, p. 141)



Research Questions

This study was designed to answer the following two research questions:

- (1) To what extent does the content of the National Center Test provide a valid sample reflecting topics related to education for international understanding?
- (2) To what extent does the Center Test cover various areas of the world?

The first question was formed because the Ministry of Education (Monbushou, 1989) emphasizes the importance of selecting appropriate topics for deepening students' international understanding. The second question was developed because some scholars claimed that regions or nations should also be investigated when a textbook (Hosoya, 1999; Kiryuu, Shibata, Tagatani & Wada, 1999) or test (Ozaki, 2009) content analysis is conducted from the viewpoint of education for international understanding.

Although washback is generated by both actual test content and the test content that teachers or students perceive (refer to Ozaki, 2009), this study investigated only the actual content of the National Center Test since the perceived test content is derived from the actual content.

Methodology

Materials for study

The National Center Tests for student intakes of 2002, 2003, and 2004 were selected for this study. The reason why the National Center Test was selected for the study material is that this test is a nationwide entrance examination for national, municipal, and various private universities, and it therefore attracts the largest number of applicants among the various forms of university entrance examination (refer to Section 7.). In other words, the Center Test is the most influential form of entrance examination in Japan, and therefore its content is likely to have a strong washback effect on both teachers and students who are involved in high school English language education. Furthermore, high school education based on the 1989 Course of Study started in 1995; the entrance examinations analyzed in this study were constructed several years after *international understanding* first appeared in the government guidelines. It is assumed that the concept of international understanding had permeated both the English entrance examination system and high school English education by that time.

Content analysts

Neuendorf (2002) stated that “at least two coders” should be used in order to obtain reliable analysis results. Therefore, two analysts were adopted for this study. Experts in the field of language testing who do not construct the target tests are ideal test content analysts for both objective and reliable analysis (Chapelle, 1999; Hughes, 2003; Ozaki, 2008; Ozaki, 2009). One of the analysts had experience in doing research and teaching both undergraduate and postgraduate courses on test construction theory and practice for over ten years. He had also developed various types of university entrance examinations. The other analyst was a PhD student from Japan, who was trained in test content analysis for this study.

Procedures

The development of a codebook and coder training for the establishment of both validity and intercoder reliability is essential (Neuendorf, 2002). Therefore, four people developed an analysis scheme and they participated in an intensive coder training for over three months. Based on the result of these procedures, two of these four people analyzed the 2001 National Center Test as a pilot study. During this pilot-study phase, special attention was paid to both inner reliability and interrater reliability following suggestions made by various scholars (e.g. Brindley, 2000; Clapham, 1996; Muijs, 2004; Neuendorf, 2002). After modifying the analysis scheme to solve problems found through the pilot study, the National Center Tests for the 2002, 2003, and 2004 intakes were analyzed independently by each analyst.

Data analysis

Unlike individual national and municipal university tests, which frequently have essay-writing and translation tasks, the National Center Test consists of only multiple-choice questions. Therefore, only reading passages and sentences for other tasks such as fill-in-the-blank and stress-pattern tasks were analyzed from the viewpoint of topics related to education for international understanding as well as regions and nations.

Test content validity should be examined based on test specifications (Hughes, 2003). However, unfortunately no specifications for the Center Test were available. Furthermore, neither the Course of Study nor its guidebook mentions detailed topics for deepening international understanding. Therefore, I decided to adopt a list of topics related to education for international understanding developed by Ozaki (2009). The final topic analysis criteria (refer to the Appendix) were developed through the results of the pilot study. There were 10 major topics; each of them had various subtopics. In order to make the results of analysis done by each analyst reliable, “the test retest method” (Muijs, 2004, p. 72) was adopted. Each analyst was required to analyze the same materials three times one week after each analysis. Furthermore, interrater agreement was calculated based on the percentage of agreed items between the two analysts and the total number of analyzed items (refer to Neuendorf, 2002).

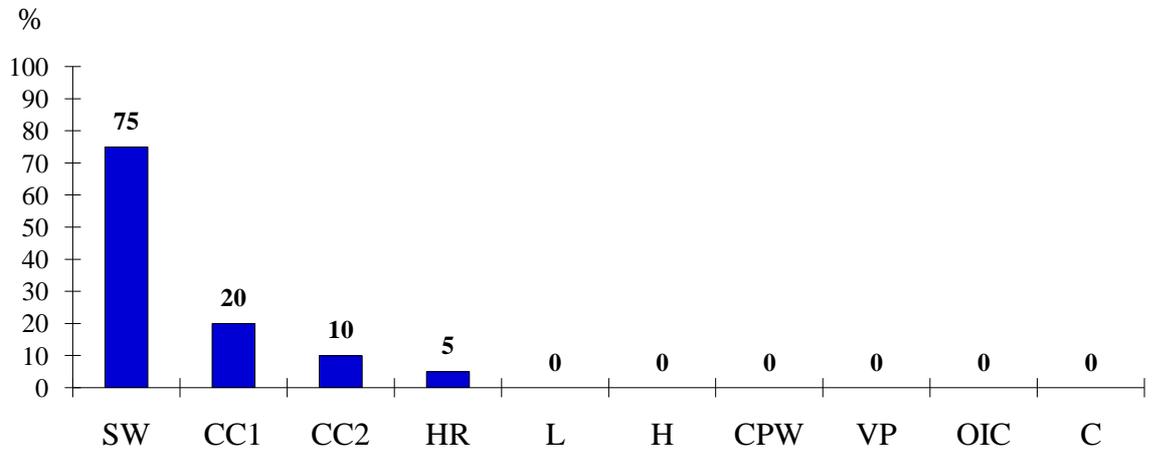
Finally, interrater agreement was calculated by the percentage of agreement, as suggested by Neuendorf (2002).

Results

Topics Related to Education for International Understanding

The results of the analysis of all the three test papers are presented together since all the three sets of tests had the same sections and tasks. The National Center Test had only multiple-choice questions and did not have any translation or writing tasks unlike individual university entrance examinations. Therefore, only passages and sentences for multiple-choice tasks related to reading, stress patterns, word usage, grammar, or conversational skills were analyzed. The sum of analysed passages or sentences was 80; 20 of them had (a) topic(s) related to education for international understanding. A breakdown of major topics is provided in Figure 2. The tests contained only three of the ten major topics (refer to the Appendix) related to education for international understanding, and topics related to *state of the world* (75%) appeared far more frequently than any other topic. Other than *state of the world*, 20% of the topics were classified into *concrete culture* and another 10% into *conceptual culture*. Among the 80 analyzed items, there was disagreement on two of them. Therefore, the agreement rate was 97.5%.

Figure 2
Percentage of Major Topics Included in the Center Tests (N: 20)



SW=State of the world, CC1=Concrete culture, CC2=Conceptual culture, HR=Human rights, L=Life, H=History, CPW=Common problems in the world, VP=Violence and Peace, OIC=Organizations for international cooperation, C=Communication

Furthermore, among the subtopics of *state of the world*, *science/technology* appeared far more frequently than any other topic (refer to Table 1); there were no significant differences among the other subtopics. It should be noted that the sum of topics or subtopics exceeded 20 since more than one topic was found in some passages.

Table 1
Topics and Subtopics in the Center Tests

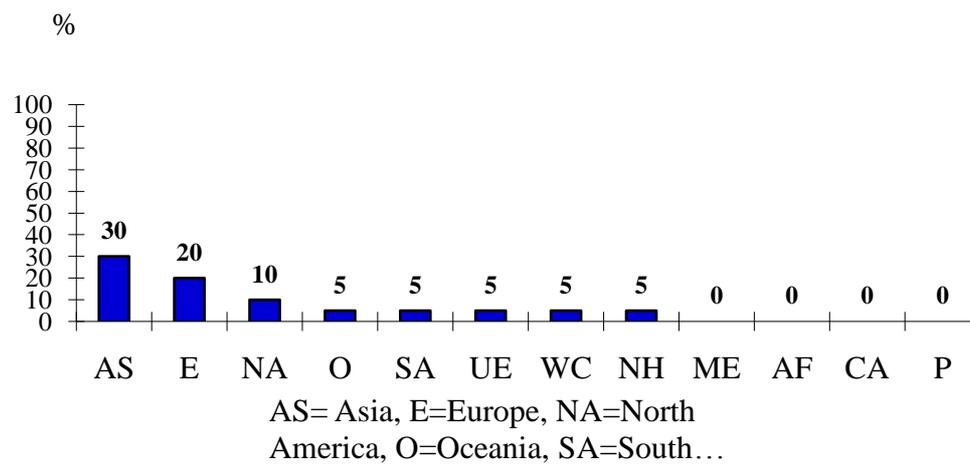
Topics	No. of Passages or Sentences (N: 20)
1. State of the world	15 (75%)
Economy	1 (5%)
Ecology	1 (5%)
Science/Technology	11 (55%)
Languages	1 (5%)
Laws	1 (5%)
2. Concrete culture	4 (20%)
Daily life (school life)	1 (5%)
Daily life (family life)	1 (5%)
Manners and Customs	2 (10%)
3. Conceptual culture	2 (10%)
Diversity of culture	2 (10%)
4. Human rights	1 (5%)
Racial discrimination	1 (5%)

Regions and nations

There were differences between the mention of particular regions (refer to Figure 2). Asia (30%) appeared more frequently than other regions. Following Asia, Europe (20%) and North America (10%) appeared more frequently than other regions, although the actual figures for these regions were not significantly large. On the other hand, the Middle East, Africa, Central America, and the South or North Pole did not appear at all.

Figure 2

Percentage of Regions Included in the Center Tests (N: 20)



There were also differences in the rate of mention of Japan (6 passages or 30%) (refer to Table 2). In Europe and North America, the UK (2 passages or 10%) and the U.S. (2 passages or 10%) appeared more frequently than other countries. There was no disagreement concerning regions or nations between the two analysts, and therefore the interrater agreement rate was 100%.

Table 2 Regions and Nations in the Center Tests

Regions and Nations	No. of Sentences/ Passages (N: 20)
1. Asia	6 (30.00%)
Japan	6 (30.00%)
2. Europe	4 (20.00%)
UK/Britain	2 (10.00%)
Germany	1 (5.00%)
France	1 (5.00%)
3. North America	2 (10.00%)
US	2 (10.00%)
4. South America	1 (5.00%)
Brazil	1 (5.00%)
5. Oceania	1 (5.00%)
Australia	1 (5.00%)
6. Northern Hemisphere	1 (5.00%)
7. English-Speaking Countries (unspecified)	1 (5.00%)

Discussion

The National Center Tests covered only a few of the ten major topics, and there was a significant preference for particular topics such as topics related to state of the world and its subtopic of science/technology. A possible reason for this phenomenon is that examination constructors felt that applicants should be able to comprehend passages about what is happening in the world because an important role of the English language is to collect updated information. Especially science and technology are significantly related to the progress of any society, and a large amount of information about these areas is provided in English. Furthermore, the survival of Japan heavily relies on these two fields. Thus, examination constructors might have believed that it was crucial for young Japanese people to be able to comprehend the information of science and technology in English. These may be the reasons why topics related to them were given a strong preference. Another possible reason is that only three test papers were analyzed, and therefore only limited kinds of topics could be found. If more test papers had been analyzed, more topics and subtopics might have been found.

There was also a clear preference for countries such as Japan, the U.S., and the UK, although various regions appeared in the National Center Tests. Especially Japan appeared far more frequently than any other nation. A possible reason for this phenomenon is that examination constructors believed that applicants should know about their own country, Japan, which would help them to not only express themselves to people from other countries but to also understand those people and their countries. Hashimoto (2000) claimed that internationalization in education means Japanization. Kobayashi (1995) and Tsuchida (2000) also pointed out the tendency to consider promoting international understanding and preserving Japanese tradition or culture to be seen as being identical due to too great an emphasis on the awareness of being a Japanese citizen. Consequently learning about Japanese tradition has come to mean education for internationalization (Tsuchida, 2000).

The U.S. also seemed to appear more frequently than other nations, although the frequency of its appearance is much less than that of the appearance of Japan, and therefore more National Center Test papers should be analyzed to see whether the U.S. really appears more frequently than other nations. A possible reason for the preference for the U.S. is that it has been influencing not only Japan but also the world as a whole in many ways, for example, militarily, economically, and politically. Therefore, it is essential to know what is happening in the U.S. or what that nation is doing with regards to other countries in order to grasp what is happening all over the world.

The frequency of the appearance of the UK was also much less than that of the appearance of Japan, and therefore more National Center Test papers should be analyzed to see whether the UK really appears more frequently than other nations. A possible reason for the preference of the UK over other English-speaking countries is that it is the birthplace of the English language.

The preference for the limited number of topics and nations might cause high school students and teachers to take notice of only those topics and nations, although education for international understanding is meant to include a much wider range of topics and nations. Consequently, students might be able

to deepen only a part of their international understanding. Therefore, both teachers and students need to be aware of the limitation of the Center Test in terms of topics related to education for international understanding and nations covered by the test. Based on such awareness, they should focus on a greater variety of topics and diverse nations when they teach or study English. However, those who attempt to go on to national or municipal universities need to take not only the National Center Test but also individual national/municipal university tests. Therefore, the content of these tests also has to be analyzed, and the results should be compared with the results of this study. The result of such a comparative study would indicate how both the Center Test and individual university entrance examinations should be improved.

It can be concluded that the results of this study are reliable since the interrater agreement rates for both topic and nation/region analysis were very high: 97.5% and 100% respectively. Frey, Botan, and Kreps (2000) stated that 70 % agreement is considered reliable. These high agreement rates were derived from an intensive rater training along with a solid content analysis scheme.

Conclusion

This study has revealed that there was a strong preference for limited types of topics related to education for international understanding and for particular nations in the National Center Tests administered between the years 2002 and 2004. It implies that both high school teachers and students might only take notice of topics and nations that appear in the National Center Test when they prepare for it. Consequently, students may deepen only a limited portion of their international understanding. Therefore, National Center Test constructors need to include a greater variety of topics and nations in order to have the test exercise a beneficial washback effect on high school English language education. Furthermore, an analysis of individual university tests should be conducted to get a fuller picture of the correct washback situation since applicants for national or municipal universities need to take such tests as well as the National Center Test. Objective analysis also needs to be conducted on both the Center Test and individual university tests since clear objectives are the key to the successful implementation of education for international understanding.

It is also crucial to investigate entrance examination content that teachers and/or students perceive since washback is generated by not only actual test content but also by perceived test content. If their perceptions of entrance examination content are different from the actual content, the tests cannot exercise their proper or intended washback effect.

Needless to say, entrance examination washback itself in terms of international understanding needs to be investigated: Washback on various educational aspects, such as classroom teaching, students' self-learning, teaching and learning materials, and curriculum, should be examined. Finally, in order to implement education for international understanding successfully, factors that can mediate the washback need to be specified.

References

- Brindley, G. (2000). Comparing AMEP assessment: A content analysis of three reading assessment procedures. In G. Brindley (Ed.), *Studies in immigrant English language assessment* (Vol. 1, pp. 45-80). Sydney, NSW: Macquarie.
- Chapelle, A. C. (1999). Validity in language assessment. *Annual Review of Applied Linguistics*, 19, 254-272.
- Clapham, C. (1996). *The development of IELTS: A study of the effect of background knowledge on reading comprehension*. Cambridge: Cambridge University Press.
- Frey, L. R., Botan, C. H., & Kreps, G. L. (2000). *Investigating communication: An introduction to research methods* (2nd ed.). Boston: Allyn & Bacon.
- Gorsuch, J. G. (1999). *Exploring the relationship between educational policy and instruction in Japanese high school classrooms*. Ann Arbor, MI: A Bell & Howell Company.
- Hashimoto, K. (2000). 'Internationalization' is 'Japanization': Japan's foreign language education and national identity *Intercultural Studies*, 21(1), 39-51.
- Hosoya, S. (1999). *Japanese foreign language teachers' attitudes toward education for international understanding*. MI: Bell & Howell Information and Learning Company.
- Hughes, A. (2003). *Testing for language teachers*. (2nd ed.). Cambridge: Cambridge University Press.
- Kaplan, R. B., & Baldauf, R. B. (2003). *Language and language-in-education planning in the Pacific Basin*. Dordrecht : Kluwer Academic
- Kiryuu, N., Shibata, T., Tagatani, H., & Wada, T. (1999). Koutougakkou eigo kyoukasho no bunseki [An analysis of senior high school English textbook]. *The Language Teacher*, 23(4), 21-23.
- Kobayashi, T. (1995). *Kokusaika to kyouiku, [Internationalization and education]*. Tokyo: Housoudaigaku Kyouiku Shinkou Kai.
- Kosaka, A. (1999). *Kokusaika in the classroom: Bakhtinian analysis of Japanese teachers' reactions to the "Internationalization of Japanese education"*. MI: Bell & Howell Information and Learning Company.
- Kubota, H. (1996). Gaikokugo kyouiku ni okeru "kokusai rikai" no shiten [The viewpoint of international understanding in foreign language education] [Electronic Version]. *Minami Ashikaga Koutou Gakkou Kiyou [Minami Ashikaga Senior High School Journal]*, 9 from http://www.page.sannet.ne.jp/h_kubota/StudyGlobal.htm.
- Mochizuki, A., Kubota, A., Iwasaki, H., & Ushiro, U. (2001). *Shin gakushuu shidou youryou ni motodzuku eigoka kyouikuhou [Teaching English as a foreign language based on the new course of study]*. Tokyo: Taishuukan.
- Monbushou. (1989). *Koutou gakkou shidou youryou kaisetsu -Gaikokugohen, eigohen - [Course of study for senior high school: Foreign language, English language]*. Tokyo: Kyoiku Shuppan.
- Monbushou. (1999). *Koutou gakkou shidou youryou kaisetsu -gaikokugohen-[Course of study for senior high school: Foreign languages]*. Tokyo: Kairyuudou.
- Muijs, D. (2004). *Doing qualitative research in education with SPSS*. London: Sage Publications Ltd.

- Muranoi, H., Chiba, M., & Hatanaka, T. (2001). *Jissenteki eigoka kyouikuhou: Sougou komyunikeeshon nouryoku o sodateru shidou [Practical English teaching: Improvement of comprehensive communicative ability]*. Tokyo: Seibidou.
- Neuendorf, A. K. (2002). *The content analysis guidebook*. London: Sage Publications Ltd.
- Niizato, M. (2000). Shin gakushuu youryou no zentai zou to shikakuteki kaisetsu [The overview and visual explanation of the new Course of Study]. In M. Niizato (Ed.), *Koutou gakkou gakushuu shidou youryou no tenkai: Gaikokugo (eigo)ka hen [The development of the Course of Study: Foreign languages (English)]* (pp. 25-45). Tokyo: Meiji Tosho.
- Ozaki, S. (2008). *Gengo tesutogaku nyuumon: Tesuto sakusei no kihonrinen to kenkyuuhou [Introduction to language testing: The principles of and research on test construction]*. Okayama, Japan: Daigaku Kyouiku Shuppan.
- Ozaki, S. (2009) *The impact of entrance examinations on education for international understanding*. Nagoya, Japan: Sankeisha.
- Satou, G. (1995). Kokusai rikai kyouiku no jissenjou no kadai: gakkou gakkyuu no kouzou tonon kakawari kara [Issues in the practice of Education for international understanding: From the viewpoint of the relationship with schools and classrooms]. *Kokusai Rikai Kyouiku [International Understanding]*, 1, 42-55.
- Satou, G. (2001). *Kokusai riaki kyouiku: Tabunka kyousei shakai no gakkou dzukuri [Education for international understanding: Creation of school education for coping with multicultural society]*. Tokyo: Akashi Shoten.
- Tsuchida, S. M. (2000). *Internationalization in schools in Japan: Educators' and students' perceptions of efforts toward internationalization and students' intercultural awareness*. MI: Bell & Howell Information and Learning Company.
- Watanabe, Y. (1997). *Washback effects of the Japanese university entrance examination: Classroom-based research*. Unpublished Unpublished doctoral dissertation, Lancaster University, UK.
- Watanabe, Y. (2004). Methodology in washback studies. In L. Cheng, Y. Watanabe & A. Curtis (Eds.), *Washback in language testing* (pp. 19-36). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
- Yoneda, S., Otsu, K., Tabushi, I., Fujiwara, T., & Tanaka, Y. (1997). *Tekisuto kokusai rikai [Textbook for international understanding]*. Tokyo: Kokudoshu.

**Appendix: Topics Related to Education for International Understanding for
Examination Content Analysis**

Major Topics	Subtopics
Life	Origin of human beings
	Birth and evolution of living creatures
	Living creatures and environments
	Structure of the Earth
Human Rights	Respect for others
	Gender equality/sexual discrimination
	Racial discrimination
	Child abuse, Social justice
	Problems of indigenous people
Concrete Culture	Daily life (e.g. food, clothing, shelter, school life, family life)
	Viewpoints and thinking patterns
	Behavioral patterns
	Manners and customs
	Stories (myths, folklores, legends)
Conceptual Culture	Relations between formation of Japanese culture and foreign culture
	Diversity and universality of culture
	Cultural relativism
	Culture shock, Identity
History	History of the Universe
	History of the Earth
	History of nations or regions
	Historical contacts between two or more nations/regions
Common Problems in the World	Poverty
	Hunger
	Diseases
	Energy
	Natural resources
	Expansion of interdependent relationships in the world /among peoples
	North-South problem
	Population
Violence and Peace	Wars
	Terrorism
	Violence
	Peace education
Organizations for International	UN (UNESCO, UNISEF)

Cooperation	NGOs
	NPOs
	ODA
State of the World	Nature
	Geography
	Economy
	Politics
	Ecology
	Technology
	Science
	Tourism/trips
	Religions
	Languages
	Flags
	Environment
	Races
	Laws
	Social structures
Communication	Verbal communication
	Non-verbal communication
	Written communication
	Intercultural communication
	Language learning/teaching