AN ANALYSIS OF CHINESE UNIVERSITY ENGLISH ENTRANCE EXAMINATIONS FROM THE VIEWPOINT OF CULTURE AND COUNTRIES

Shigeru Ozaki
Shuhan Lai

ABSTRACT

Teaching culture is an important policy of high-school English education in China, and the country experiences significant washback effects of university entrance examinations on high-school English education. However, there has been little research on the content of Chinese university English entrance examinations in terms of cultural topics, although the source of washback is test content. Therefore, we analyzed university entrance examinations administered in Beijing and Shanghai from the perspective of cultural topic categories, the source-culture country, and the three concentric circles (Inner, Outer, and Expanding Circles). The results showed that both two tests had the same preferences in favor of particular topics as well as countries: Topics related to society and daily life frequently appeared. The U.S. and UK appeared far more frequently than other countries in both examinations, while Outer- and Expanding-Circle countries rarely appeared. The source-culture country did not appear at all. On the other hand, a few differences were also found in these tests: In the Beijing tests, topics related to social problems often appeared; on the other hand, in the Shanghai tests, career-related topics frequently appeared. This difference might have arisen because the former city is the center of politics, and the latter is the center of commerce. In the Beijing tests, the U.S.A. appeared far more frequently than the UK; the reason might be that the former is politically more important to Beijing, which is the center of politics in China. This study implies that English entrance exam constructors in China should include cultural topics and countries that were not found very often or at all in this study in their tests. This is to ensure that both high-school students and teachers will pay more attention to more diverse cultural topics and countries, which is necessary in learning English as an international language.

Keywords: China, high-school English education, university entrance examinations, washback, culture

Introduction

Teaching culture is one of the most important policies related to English instruction in Chinese high schools (Ministry of Education of China, 2000). However, this policy may not be successfully implemented if university entrance examinations do not include cultural elements. "The effect that tests have on learning and teaching" is described as the washback effect (Hughes, 2003, p. 53); the washback effects of high-stakes language tests may influence the implementation of language education policy (Gorsuch, 1998; Kaplan & Baldauf, 2003; Ozaki, 2012). English instruction in China is no exception; some researchers (Eckstein & Noah, 1993; Li, 1990; Lin, 2009; Li and Qi, 2012; Qi, 2004) have discussed or found strong washback of Chinese university entrance examinations. However, although the original source of the washback effect is a test's content (Ozaki, 2012), there has been little research regarding the content of Chinese university entrance examinations from the perspective of culture. Therefore, this study analyzes them from cultural viewpoints.

LITERATURE REVIEW

Washback in General
Many researchers in educational fields have mentioned the influence of tests on teaching and learning (Vernon, 1956; Davies, 1968; Kellaghan, Madaus & Airasian, 1982; Morrow, 1986; Hughes, 2003; Khaniya, 1990). Language testing is not an exception (Finocchiaro & Sako, 1983; Bachman, 1990). This concept is described as "backwash" or "washback" (Davies, 1990: 1), and the more important a test is considered, the more washback it has (Hughes, 2003; Mehrens and Kaminsky 1989). Due to the significant importance of washback, the validity of a test should be measured based on the degree of beneficial washback it exercises (Alderson & Wall, 1993; Morrow, 1986; Frederiksen & Collins, 1989). Morrow (1986) even coined the term *washback validity*. However, it should not be assumed that washback occurs automatically because the nature of washback is complex (Wall & Alderson, 1993; Blewchamp, 1994; Alderson & Hamp-Lyons, 1996). Watanabe (1997, 2004) classifies the complex characteristics of washback into the following three categories: dimensions, influenced aspects, and mediating factors. Concerning the cause of washback, Ozaki (2012) states that it is test content, although there are sometimes gaps between perceived content and actual content.

Washback of University English Entrance Examinations in China
In China, the existence of the washback of university English entrance examinations is very obvious and crucial; it has been noticed by a large number of researchers from various countries including China (Lin, 2009). Since students there pay very little attention to what entrance examinations do not include (Eckstein & Noah, 1993), the tests' washback is one of the most important keys to successful high-school English language education in China (Lin, 2009). Li (1990) as well as Li and Qi (2012) have indicated the strong washback of university English Entrance examinations in China: The former investigated the washback of the National Matriculation English Test (NMET), the English test for entrance to all universities in China, and found that the test had significant washback on stakeholders, and suggested that a test could be used to enhance and to make a new idea in the

teaching of English. The latter researched into the reading comprehension section of the National Matriculation English Test in Guangdong Province and also found the washback of the test. Qi (2004), on the other hand, found that the NMET had only limited intended washback effects: Teachers still focused heavily on knowledge of language and taught only the skills tested in the NMET.

Cultural Topics for Content Analysis

Although the Course of Study in China emphasizes the importance of culture in high-school English education, it does not refer to the details of culture (Ministry of Education of China, 2000). Therefore, it is crucial to review literature on culture. Culture has frequently appeared in the language education literature (e.g., Brown, 2000; Damen, 1987; Flewelling, 1994; Kramsch, Cain & Murphy-Lejeune, 1996; Lessard-Clouston, 1996; Manto, 1992; Omagio 1993; Ozaki, 1999; Wenying, 2000), and it has been defined in numerous ways (Klopf, 1998; Ashikaga, Fujita, & Ikuta , 2001). For example, there have been over 100 definitions in the social sciences (Ishii, Okabe, & Kume, 1996). Despite its diverse definitions, its general nature is described as automaticity (Lund, 2006), since "culture is passed on to members of a cultural group through socialization and differs across time, place, community, and generation" (Toprak & Aksoyalp, 2014, p. 93).

One of the most common classifications of culture has been Culture and culture (Ashikaga et al., 2001; Yamanaka, 2004), which was originally proposed by Allen and Valette (1972). Culture refers to "the people's achievements and contribution to the civilization as in art, music, technology and architecture" (Ashikaga et al., 2001, p.2), whereas "culture refers to the behavioural patterns of life styles of people" (Ashikaga et al., 2001, p. 2). Ashikaga et al. (2001), citing various classifications of culture (e.g., Allen & Valette, 1972; Robinson, 1988; McGroarty & Galvan, 1985; Moerman, 1988), concluded that culture has two broad aspects: "surface manifestations which can be explicitly observed, and hidden manifestations which cannot be observed from outside" (p. 2).

Further, some researchers (Ashikaga et al, 2001; Liu, 2013; Ozaki, 2012; Yamanaka, 2004) analysed the content of high-school textbooks or English entrance examinations from the viewpoint of culture. Ashikaga et al. (2011) adopted the following two cultural analysis categories: concrete culture and abstract culture, and Ozaki (2012) adopted concrete culture and conceptual culture for his analysis of Japanese university English entrance examinations. On the other hand, Yamanaka (2004) developed more specific cultural categories on the basis of the guidebook of the Course of Study for foreign languages in order to analyze Japanese high-school English textbooks. Her cultural categories included daily life, social affairs, manners and customs, geography, history, science, stories, and other items related to intercultural understanding. She further analyzed the textbooks from the viewpoint of Kachru's three concentric circles: the Inner Circle (countries where English is used as a first or native language), the Outer Circle (countries where English is applied as a second or official language), and the Expanding Circle (countries where English is employed as a foreign language). Her results indicated that approximately half of the cultural topics in both junior and senior high-school textbooks were related to daily life. Yamanaka (2006) also found that the source-culture (McKay, 2002) country of Japan and two Inner-Circle countries (the U.S. and UK) appeared more frequently than other countries. Liu (2013) conducted a cultural analysis of Chinese high-school English textbooks and modified Yamanaka's cultural categories into daily life, society, manners and costumes, geography, history, and science. She found that daily life, geography, and society appeared frequently in the textbooks, and that the U.S., UK, and China appeared much more frequently than other countries. Ozaki (2010), on the basis of an extensive literature review, also developed cultural topics as part of a list of topics for *kokusai rikai kyoiku* (literal translation: education for international understanding), one of the most crucial elements of which is culture, to analyze Japanese university English entrance examinations. He classified culture into two: concrete culture and conceptual culture. The former consists of daily life (e.g., food, clothing, shelter, school life, and family life), viewpoints and thinking patterns, behavioural patterns, manners and customs, stories (myths, folklores, and legends); the latter consists of relations between formation of Japanese culture and foreign cultures, diversity and universality of culture, cultural relativism. Concrete culture appeared more frequently than conceptual culture in the entrance examinations he analyzed. He also found that Japan, the U.S., and UK appeared much more frequently than other countries.

When cultural content is analyzed, it is crucial to examine which country a particular culture belongs to, since culture is common to the majority of members from a certain country. For cultural analysis of English textbooks or tests, Kiryuu, Shibata, Tagatani, and Wada (1999), Yamanaka (2006), and Ozaki (2010) adopted Kachru's (1989) three concentric circles: the Inner Circle, Outer Circle, and Expanding Circle. According to McKay (2002), when English is taught as an international language, culture can be classified into three types: "source culture," referring to "learners' own culture"; "target culture," referring to "the culture of a country where English is spoken as a first language"; and "international target culture," referring to "a great variety of cultures in English and non-English-speaking countries around the world" (p. 88). Liu (2013) and Ozaki (2012) adopted source culture in addition to the three concentric circles for their cultural analyses of Chinese high-school English textbooks and Japanese university English entrance examinations, respectively.

As this section has revealed, there have been studies that analyzed university English entrance examinations in Japan and high-school English textbooks in Japan as well as China from the viewpoint of cultural topics and countries. However, to the best of my knowledge, there has been scarce research on the content of university English entrance examinations in China from the same viewpoint, even though they strongly influence high-school English language education there a crucial element of which is culture learning.

RESEARCH QUESTIONS

This study answers the following research questions, which were formed based on the literature review:

(1) What cultural categories appear in Chinese university English entrance examinations, and how frequently do they appear?

(2) What countries from the three concentric circles appear in the examinations, and how frequently do they appear?
(3) How frequently does the source-culture country appear in the examinations?

METHOD

Analysis Materials
This study analyzed passages from the reading comprehension sections of the university English entrance examinations that were administered in Beijing and Shanghai between 2004 and 2012 for the following reasons:
(1) These two cities are the most populated and cosmopolitan Chinese cities. Therefore, their entrance examinations are assumed to have a significant number of applicants and include cultural topics as well as various countries.
(2) University entrance examinations started reflecting the content of the 2000 Course of Study in 2004.

Analysis Criteria
Brown (2005) suggested that, "In order to investigate content validity, testers must decide whether the test is a representative sample of the content the test was designed to measure" (p. 221). Hughes (2003) stated, "In order to judge whether or not a test has content validity, we need a specification of the skills or structures etc. that it is meant to cover…" (pp. 26–27). However, the Course of Study does not refer to specific cultural topics. Therefore, in this study, the two analysts developed original cultural topic categories (main categories and subcategories, see Appendix) by referring to the topics developed based on literature review (e.g., Ashikaga et al., 2001; Yamanaka, 2004; Ozaki, 2012; Liu, 2013). In addition, following Yamanaka (2004), Ozaki (2012), and Liu (2013), the three concentric circles and the source culture were adopted as analysis criteria (see LITERATURE REVIEW). Finally, the topic categories were revised through a pilot study.

Analysts and Inter-rater Reliability
Test content should be analyzed by those who "are familiar with language teaching and testing but who are not directly concerned with the production of the test in question" (Hughes, 2003, pp. 26–27). In addition, one of the most important issues in content analysis is reliability. Thus, data should be analyzed by more than one person (Neuendorf, 2002). In the calculation of inter-rater reliability, simple percent agreement is the most popular means in social and behavioral sciences (Neuendorf, 2002). In fact, Brindley (2000) adopted this approach for his test content analysis and set the benchmark at 70%. Therefore, this study employed a testing expert and a postgraduate student from China, who were familiar with Chinese university entrance examinations and language testing. The benchmark for inter-rater agreement rate was set at 70%.

Rater Training and a Pilot Study
The two analysts were trained on both intra- and inter-rater reliability, particularly through a pilot study, which was conducted with entrance examinations administered in 2004, 2005, and 2006. The following problems were found and solved accordingly:

(1) Personal experience should be distinguished from culture, which is more general to the members of society.
(2) New subcategories should be added: tools, crime, and accidents.
(3) The list of cultural topic categories should be modified when necessary.
(4) When a location name is presented, analysts should investigate in what country it is located.
(5) It is difficult to avoid misanalysis. Therefore, the results must be discussed and modified when necessary in order to avoid misanalysis and to ensure the validity of the analysis.

PROCEDURES

The two analysts analyzed the entrance examinations individually and reanalyzed the same materials twice to ensure intra-reliability. The second analysis was conducted one week after the first so that the analysts did not remember the first analysis result, yet still remembered and had a good understanding of the analysis procedures and criteria. The outcomes of the individual analyses were compared and discussed in terms of main categories, subcategories, details, and countries. Whenever a discrepancy was found, the cause was immediately identified through discussion to avoid careless mistakes, and a reanalysis was conducted when necessary. When the discrepancy persisted even after the discussion, the items were eliminated from the final analysis results, although the discrepancy was included in the reliability calculation unless the cause was a careless mistake.

RESULTS
Before presenting the actual results, we will briefly explain how they are presented. The Beijing and Shanghai entrance examinations included 42 and 33 passages, respectively, and occasionally one passage contained more than one main category, subcategory, and/or country. The results for these two examination types are first presented separately and then together. The results are not presented for each year since each entrance examination contained only three to five passages. The inter-rater agreement rates for the analysis of Beijing and Shanghai entrance examinations in terms of main categories were 80.8% and 94.4%, respectively. The inter-rater agreement rate for the analysis of both Beijing and Shanghai examinations in terms of countries was 100%. These results indicate that all the agreement rates were higher than the benchmark (70%). The results of the pilot study were included for a better understanding of general tendencies with a larger amount of data, since it did not have any major problems, and its analysis criteria and procedures were exactly the same as the actual study's. Whenever a problem or a new topic category was found in the pilot study, the exam papers were reanalyzed based on the new rules set through it.

Major Findings Regarding Main Categories
Table 1 presents the analysis results in terms of the main categories. The order of the categories is arranged according to the frequency of their appearance so that readers can easily grasp the entire tendency.

Table 1. Main Categories

Main Categories	Beijing (42 passages)	Shanghai (33 passages)
Society	23 (54.8%)	17 (51.5%)
Daily Life	16 (38.1%)	14 (42.4%)
Manners and Customs	5 (11.9%)	2 (6.1%)

The majority of passages (38, 90.5%) in the Beijing examinations included at least one main category. Society and daily life, especially the former, appeared far more frequently than manners and customs. Meanwhile, most of the passages (28, 84.8%) in the Shanghai examinations also included at least one main category. Society and daily life appeared far more frequently than manners and customs. In terms of the main categories, the Beijing and Shanghai examinations were similar.

Major Findings Regarding Subcategories
Of the 37 subcategories, 18 (51.3%) appeared at least once in the Beijing entrance examinations, and 16 (43.2%) appeared in the Shanghai examinations. Table 2 presents the results in terms of the subcategories of daily life. The order of the categories is not arranged according to the frequency of their appearance because the results differed between the Beijing and Shanghai examinations.
  Table 2 presents the results in terms of the subcategories of daily life

Table 2. Subcategories of Daily Life

Subcategories	Beijing (42 passages)	Shanghai (33 passages)
Food	0	0
Clothes	0	0
Shelter	0	0
Sports	0	1 (3.0%)
Entertainment	1 (2.4%)	2 (6.1%)
Languages	3 (7.0%)	1 (3.0%)
Education & School Life	6 (14.3%)	4 (12.1%)
Family Life	2 (4.8%)	1 (3.0%)
Appliances	0	0
Vehicles	0	0
Companies & Organizations	0	0
Career	2 (4.8%)	6 (18.2%)
Animals	2 (4.8%)	1 (3.0%)
Tools	1 (2.4%)	0
Arts	1 (2.4%)	0
Hobbies	0	1 (3.0%)

In the Beijing examinations, education and school life appeared more frequently than the other subcategories. In the Shanghai examinations, career appeared most frequently, and education and school life also appeared frequently compared to the other categories. Conversely, in the Beijing counterparts, career appeared only infrequently. The other categories either did not appear frequently or not at all in both the examinations.

Table 3. Subcategories of Society

Subcategories	Beijing (42 passages)	Shanghai (33 passages)
Social Events	0	1 (3.0%)
Races	1 (2.4%)	0

Religions	0	0
History	3 (7.1%)	1 (3.0%)
Economy	1 (2.4%)	2 (6.1%)
Politics	0	0
Tourism	0	1 (3.0%)
Science	11 (26.2%)	7 (21.2%)
Mass Media	3 (7.1%)	0
Infrastructure	1 (2.4%)	0
Social Problems	13 (30.9%)	4 (12.1%)

In the Beijing examinations, social problems and science appeared far more frequently than the other subcategories. In the Shanghai examinations, science appeared far more frequently than the other subcategories; however, social problems appeared much less frequently than in the Beijing counterparts The other subcategories either did not appear frequently or not at all in both the examinations.

Table 4 presents the results in terms of the subcategories of manners and customs.

Table 4. Subcategories of Manners and Customs

Subcategories	Beijing (42 passages)	Shanghai (33 passages)
Thinking Patterns	2 (4.8%)	2 (6.1%)
Behavioral Patterns	1 (2.4%)	1 (3.0%)
Public Manners	1 (2.4%)	0
Greeting	0	0
Name Order	0	0
Gestures	0	0
Expressions of Love	0	0
Proverbs	1 (2.4%)	0
Superstitions	0	0
Legends	0	0

Four subcategories (thinking patterns, behavioral patterns, public manners, and proverbs) appeared in the Beijing examinations, while only two subcategories appeared in the Shanghai examinations (thinking patterns and behavioral patterns). However, none of these subcategories appeared frequently.

Overall Tendencies Among the Subcategories
Not many of the 37 subcategories appeared very frequently in either the Beijing or Shanghai examinations. In the Beijing examinations, social problems, science, and education and school life appeared more frequently than the other subcategories. In the Shanghai examinations, science, career, and education and school life appeared more frequently than the other subcategories.

Major Findings Regarding Countries
Table 5 presents the results in terms of countries.

Table 5. Countries

Circles & Source	Countries	Beijing (42 passages)	Shanghai (33 passages)
Inner	U.S.	17 (40.5%)	9 (27.3%)
	UK	6 (14.3%)	10 (30.3%)
	Canada	1 (2.4%)	2 (6.1%)
	Australia	0	2 (6.1%)
Outer	India	1 (2.4%)	2 (6.1%)
	Malawi	0	1 (3.0%)

| Expanding | Indonesia | 0 | 1 (3.0%) |
| Source | China | 0 | 0 |

In the Beijing examinations, two Inner-Circle countries, the U.S. and UK, especially the former, appeared far more frequently than other countries, and only one Outer-Circle country appeared: India. In contrast, the Expanding-Circle countries and the source-culture country did not appear at all.

In the Shanghai examinations, two Inner-Circle countries, the U.S. and UK, appeared more frequently than other countries, and two Outer-Circle countries appeared: India and Malawi. Only one Expanding-Circle country appeared in one passage: Indonesia. Although Europe appeared in one passage, the names of the countries were unidentifiable. The source-culture country did not appear at all. A greater number of countries and different types of concentric circles appeared in the Shanghai examinations than in the Beijing counterparts.

DISCUSSION

This study found that only particular cultural topic categories, subcategories, and countries appeared in university entrance examinations, although culture consists of a large number of topics and the teaching of English as an international language should pay attention to the cultures of diverse countries. The bias found in this study may have washback effects on high-school English education, especially as the examination period nears. Thus, both teachers and learners should consciously focus on these categories for well-balanced culture teaching and learning. For example, teachers need to know them and include a wider variety of cultural categories and countries in their teaching materials to develop their students' knowledge and understanding of diverse cultures, which leads to better intercultural communication.

Although most of the findings were common to both Beijing and Shanghai examinations, there were a few differences between them: In the Beijing tests, topics related to social problems appeared much more frequently than in the Shanghai counterparts; on the other hand, in Shanghai tests, career-related topics appeared far more often than in the Beijing counterparts. This difference might have arisen because the former city is the center of politics, and the latter is the center of commerce. In the Beijing tests, the U.S.A. appeared far more frequently than the UK, and the reason might be that the former is politically more important to Beijing, which is the center of politics in China.

The present study and previous studies have similarities and differences. In this study and Ozaki's (2012) analysis of Japanese university English entrance examinations, science topics appeared frequently. This result implies that one goal of English education in both China and Japan is the development of science, which can enable the countries to prosper. Nevertheless, the actual reason is awaiting an investigation of entrance examination constructors. Furthermore, both in this study and in Liu's (2013) analysis of Chinese high-school English textbooks, society and daily life, especially the former, appeared frequently. Conversely, manners and customs did not appear very frequently, although such topic is important for everyday interpersonal communication.

A preference for the U.S. and UK was also found in other studies such as Liu's (2013) analysis of Chinese high-school English textbooks, Ozaki's (2010, 2012) analysis of Japanese university English entrance examinations, and Yamanaka's (2006) analysis of Japanese high-school English textbooks. This tendency in both textbooks and entrance examinations may cause both teachers and learners to focus on only these two countries, although there are other English-speaking countries and communication in English as an international language involves not only Inner-Circle countries but also Outer- and Expanding-Circle countries.

In the present study, the source-culture country did not appear at all, whereas in Liu (2013), it appeared frequently. This remarkable discrepancy between Chinese university entrance examinations and high-school English textbooks might lead to negative washback effects on high-school English education, although these examinations are not necessarily based on the textbooks that Liu analyzed. For example, students might not read textbook chapters about China as seriously as those about the U.S. and UK, even though it is important to understand one's own culture in order to understand other cultures because their understanding tends to be based on a comparison between their own and other cultures.

Finally, the results of this study suggest how university entrance examinations can be improved in terms of their cultural content. They should include more diverse topics, especially ones related to manners and customs, since they are important in interpersonal communication. They should also include more diverse countries from the three concentric circles, especially Outer Circle, Expanding Circle, and the source-culture country; because English is acquired as an intercultural communication tool throughout the world.

CONCLUSION

This study analyzed university English entrance examinations administered in Beijing and Shanghai from the perspective of culture and countries, and it found preferences for particular cultural categories, concentric circles, and countries. These preferences may exert washback effects on high-school English education. Therefore, entrance examination constructors should include more cultural topics and countries in their tests. Further, the results of this study imply the necessity to investigate he content of tests on English as an international language, since they suggest that there is a fear of extreme preference in favor of limited cultural topics and countries in such tests.

This research has several limitations: First, it only focused on two particular examinations, although a few other types of entrance examinations exist in China. Therefore, the findings may not represent nationwide tendencies. Second, it analyzed only reading passages. However, it is desirable to investigate other sections, such as essay-writing and listening sections, in order to

examine the tests more thoroughly from the perspective of various skills. Third, it did not investigate the process of actual entrance examination construction. Thus, it is necessary to investigate this process to examine the causes of the tendencies found in this study. For example, it would be effective to interview entrance examination constructors. Finally, it is necessary to investigate both actual and perceived washback effects on high-school English education by classroom observation, study journals, interviews, and questionnaires.

In concluding this article, we would like to emphasize the following point: To achieve fruitful and effective high-school English education, it is crucial to improve not only everyday teaching and learning, including materials and curriculum, but also high-stakes tests such as university entrance examinations, since they tend to exert washback effects on education.

REFERENCES

Alderson, J. C., & Hamp-Lyons, L. (1996). TOEFL preparation courses: A study of washback. *Language Testing, 13,* 280–297.

Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics, 14,* 115–129.

Allen, E. D., & Valette, R. M. (1972). *Modern language classroom techniques: A handbook*. New York: Harcourt Brace Jovanovich.

Ashikaga, T., Fujita, R., & Ikuta, Y. (2001). A study of cultural aspects in Japanese EFL communication textbooks. *JACET Bulletin, 33*, 1–10.

Bachman, F. L. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.

Blewchamp, P. (1994). *Washback in TOEFL classrooms: An explanatory investigation into the influence of the TOEFL test on teaching content and methodology.* Unpublished MA Thesis, Lancaster University, UK.

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, Australia: Macquarie University.

Brown, H. D. (2000). *Principles of language learning and teaching* (4th ed.). White Plains, NY: Pearson Education.

Brown, J. D. (2005). *Testing in language programs: A comprehensive guide to English language assessment*. New York: McGraw-Hill.

Damen, L. (1987). *Culture learning: The fifth dimension in the language classroom*. Reading: Addison-Wesley.

Davies, A. (1968). *Language testing symposium: A psycholinguistics approach*. Oxford: Oxford University Press.

Davies, A. (1990). *Principles of language testing.* Oxford: Basil Blackwell.

Eckstein, M. A., & Noah, H. J. (1993). *Secondary school examinations: International perspectives on policies and practice*. New Haven: Yale University Press.

Finocchiaro, M., & Sako, S. (1983). *Foreign language testing: A practical approach.* New York: Regents.

Flewelling, L. J. (1994). The teaching of culture: Guidelines from the national core French study of Canada. *Foreign Language Annals, 27*(2), 103–108.

Fredriksen, R. J., & Collins, A. (1989). A systems approach to educational testing. *Educational Researcher, 18*(9), 27–32.

Gorsuch, J. G. (1999). *Exploring the relationship between educational policy and instruction in Japanese high school classrooms*. Ann Arbor, MI: Bell & Howell.

Hughes, A. (2003). *Testing for language teachers* (2nd ed.). Cambridge: Cambridge University Press.

Ishii, S., Okabe, R., & Kume, T. (1996). *Ibunnka komyunikeeshon* [Intercultural communication]. Tokyo: Yuuikaku.

Kachru, B. (1989). Teaching world Englishes. *Cross Currents: An International Journal of Language Teaching and Cross-Cultural Communication*, *16* (1), 15–21.

Kaplan, R. B., & Baldauf, R. B., Jr. (2003). *Language and language-in education planning in the Pacific basin.* Dordrecht, The Netherlands: Kluwer Academic Publishers.

Kellaghan, T., Madaus, G. F., & Airasian, P. W. (1982). *The effects of standardized testing.* London: Kluwen, Nijholf.

Khaniya, T. R. (1990). The washback effect of a textbook-based test. *Edinburgh Working Papers in Applied Linguistics, 1,* 48–58.

Kiryuu, N., Shibata, T., Tagatani, H., & Wada, T. (1999). Koutougakkou eigo kyoukasho no bunseki [An analysis of senior high school English textbooks]. *The Language Teacher, 23*(4), 21–23.

Klopf, D. W. (1998). *Intercultural encounters: The fundamentals of intercultural communication*. Englewood, Colorado: Morton.

Kramsch, C., Cain, A., & Murphy-Lejeune, E. (1996). Why should language teachers teach culture? *Language, Culture and Curriculum, 9*(1), 99–107.

Lessard-Clouston, M. (1996). Chinese teachers' views of culture in their learning and teaching. *Language, Culture and Curriculum, 9*(3), 197–242.

Li, X. J. (1990). How powerful can a language test be? The Met in China. *Journal of Multilingual and Multicultural Development*, *11*(5), 393–404.

Li, J., & Qi, L. (2012). Guangdong gaokao yingyuyuedu xintixingde fanboxiaoyingyanjiu. [Washback research on a new type of reading comprehension task in the National Matriculation English Test Guangdong Version]. *Yingyu Jiaoshi*, *10*, 2–8 and15.

Lin, J. (2009). *Washback of English examination on English teaching and learning in China*. Unpublished paper. University of Wisconsin-Platteville.

Liu, N. (2013). *An analysis of Chinese high school English textbooks from the viewpoint of culture*. Unpublished master's thesis. Takushoku University, Tokyo, Japan.

Lund, R. (2006). *Questions of culture and context in English language textbooks. A study of textbooks for the teaching of English in Norway*, Unpublished Doctoral Dissertation, University of Bergen, Norway.

Manto, K. (1992). *Kokusaika to eigo kyouiku* [Internationalization and English language education]. Tokyo: Taishukan.

McGroarty, M., & Galvan, J. L. (1985). Culture as an issue in second language teaching. In M. Celce-Murcia (Ed.), *Beyond basics: Issues and research in TESOL* (pp. 81–95). Cambridge, MA: Newbury House.

McKay, L. S. (2002). *Teaching English as an international language: Rethinking goals and approaches.* Oxford: Oxford University Press.

Mehrens, A. W., & Kaminsky, J. (1989). Methods for improving standardized test scores: fruitful, fruitless or fraudulent? *Educational measurement: Issues and practice, 8*(1), 14–22.

Ministry of Education of China. (2000). *Education guidelines in China.* Beijing: People Education Public. Retrieved from http://wenku.baidu.com/view/d9b88431b90d6c85ec3ac6af.html

Moerman, M. (1988). *Talking culture: Ethnography and conversation analysis.* Philadelphia: University of Pennsylvania Press.

Morrow, K. (1986). The evaluation of tests of communicative performance. In M. Portal (Ed.), *Innovations in language testing.* Windsor: NFER/Nelson.

Neuendorf, A. K. (2002). *The content analysis guidebook.* London: Sage.

Omagio, A. C. (1993). *Language teaching in context: Proficiency oriented instruction.* Boston: Heinle & Heinle.

Ozaki, S. (1999). The role of Nihonjijo: Based on a model of interpersonal communication. *The Language Teacher, 23*(9), 34–37.

Ozaki, S. (2010). High-stakes tests and language education policy in Japan: An analysis of the National Center English Test from the viewpoint of education for international understanding. *TESOL Journal, 2,* June, 175–188.

Ozaki, S. (2012). *Content and perceived washback of Japanese university English entrance examinations in terms of international understanding*. Unpublished doctoral dissertation, The University of Queensland, St. Lucia, Australia.

Qi, L. (2004). Has a high-stakes test produced the intended changes? In L. Cheng, Y. Watanabe &

A. Curtis (Eds.), *Washback in language testing: Research contexts and methods.* (pp. 171–190). Mahwah, NJ: Lawrence Erlbaum Associates.

Robinson, G. (1988). *Cross-cultural understanding.* New York: Prentice Hall.

Toprak, T. E. & Aksoyalp, Y. (2015). The question of re-presentation in EFL course books: Are learners of English taught about New Zealand? International Journal of Society, *Culture & Language, 3*(1), 91–104. Retrieved from http://www.ijscl.net/article_9933_a30435ab90128f93943fa828b7b4c518.pdf

Vernon, P. E. (1956). *The measurement of abilities.* London: University of London Press.

Wada, K. (1999). *Eigo ni okeru kokusai rikai kyouiku* [Education for international understanding in English language education]. Tokyo: Taishukan.

Wall, D., & Alderson, J. C. (1993). Examining washback: The Sri Lankan impact study. *Language Testing, 10,* 41–69.

Watanabe, Y. (1997). *Washback effects of the Japanese university entrance examination: Classroom-based research.* 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.

Wenying, J. (2000). The relationship between culture and language. *ELT Journal, 54*(4), 328–344.

Yamanaka, N. (2004). An evaluation of English textbooks from the viewpoint of culture based on the 2003 Ministry of Education's course of study guidelines. *JACET Bulletin*, *39*, 87–107.

Yamanaka, N. (2006). An evaluation of English textbooks in Japan from the viewpoint of nations in the inner, outer, and expanding circles. *JALT Journal, 28*(1), 57–76.

APPENDIX

Cultural Categories

Main Categories	Subcategories
Daily Life	Food
	Clothes
	Shelter
	Sports
	Entertainment
	Languages
	Education and school life
	Family life
	Appliances
	Vehicles
	Companies and Organizations
	Career
	Animals
	Tools

	Arts
	Hobbies
Society	Social events
	Races
	Religions
	History
	Economy
	Politics
	Tourism
	Science
	Mass media
	Infrastructure
	Social problems
Manners and Customs	Thinking patterns
	Behavioral patterns
	Public manners
	Greetings
	Name order
	Gestures
	Expressions of love
	Proverbs
	Superstitions
	Legends

Shigeru Ozaki
Faculty of Foreign Languages
Takushoku University, Tokyo Japan
Email: shigeruozaki@hotmail.com

Shuhan Lai
Graduate School of Language Education
Takushoku University, Tokyo Japan