

RESEARCH

Open Access



Application of diffusion of innovation theory to educational accountability: the case of EFL education in Japan

Miyuki Sasaki

Correspondence:
sasaki@hum.nagoya-cu.ac.jp
Graduate School of Humanities and
Social Sciences, Nagoya City
University, Yamanohata 1,
Mizuho-cho, Mizuho-ku, Nagoya
467-8501, Japan

Abstract

Background: This study investigates how the criterion-referenced (CR) approach has impacted relationships between goals, classroom practices, and student achievement in English education in Japan from 1994 to the present, a period covering three government-mandated administrations of a national curriculum (Course of Study). No study has investigated such relationships longitudinally as evidence of accountability of these curriculum policies, and this is a first step.

Methods: Study 1 compares their alignment from two periods (1994–2002 and 2003–2013) based on the legally-binding goals set by the Government, nation-wide CR tests based on these goals, and English teachers' answers to a questionnaire investigating their classroom practices. Study 2 explores how current goals relate to the results of a new set of CR tests and a new survey of classroom practices. The study contributes to the field in two significant ways.

Results: Using Roger's (Rogers, 2003) Diffusion of Innovation Theory made it possible to analyze the implementing processes for new educational policies as multi-faceted and susceptible to influences from stakeholders' societal value systems.

Conclusions: Results of tests and surveys collected from large samples truly reflect the targeted populations and provide empirical evidence supporting the widely-recognized narrative that high-stakes university entrance examinations strongly affect Japanese EFL education.

Background

This study investigates how relationships between goals, classroom practices, and student achievement in English as a foreign language (EFL) education have changed in Japan from 1994 to the present, covering three administrations of government-mandated curriculum policies since the introduction of criterion-referenced (CR) assessment.¹ These three administrations represent the most recent attempts by the Japanese government to better adjust EFL education to accelerating globalization, with learning outcomes (both positive and negative) expected to provide useful information for improving similar educational systems, especially in Asian contexts. In this study, I focus on the three variables of goals, classroom practices, and student achievement because these are “key aspects of curriculum policy” (Cumming, 2001, p. 213), especially when the policy utilizes CR assessment as its tool for evaluating learning success. In

other words, the alignment of the three key aspects provides important evidence for the “accountability” (Cumming, 2009, p. 90) of the educational policy in question.

Such alignment has been investigated in various contexts, including that of the implementation of the Common European Framework of Reference for Languages (CEFR; e.g., Alderson, 2005), which transcends second language education across European countries, as well as the implementation of the 2001 No Child Left Behind Act in the US (e.g., Harper, Platt, Naranjo, & Boynton, 2007). Although the introduction of CR measurement in educational policies has a long history (e.g., Bachman, 1990), few studies have investigated how these policies have impacted classroom practices, especially on a longitudinal basis. Turning to recent EFL educational policies in Japan, although a number of studies have focused on specific aspects of these policies (e.g., communicative language teaching: Tahira, 2012; EFL education in the elementary school: Butler, 2015), no study (to my knowledge) has addressed the issue of the nationwide impact of CR types of governmental goals for teaching and learning in EFL classrooms using data from rigorously selected samples, not to mention the longitudinal consequences of such policies. This study therefore covers new ground.

Methodologically, the study is unique in its use of Roger’s (2003) Diffusion of Innovation Theory. I selected this framework because the three targeted administrations of curriculum policies stood out through “the introduction of the government’s new educational policies” (Sasaki, 2008, p. 73) in the 160-year history of Japanese EFL education. Roger’s theory is especially appropriate because it has proved successful in providing tools for revealing how, why, and how fast an innovation achieves (or fails to achieve) its intended goals, including in the fields of education (e.g., Lee, Hsieh, & Hsu, 2011) and educational policy (e.g., Dingfelder & Mandell, 2011). Furthermore, the theory’s underlying assumption that innovation is accepted through communication over time as a result of stakeholders’ values and beliefs is also relevant when I examine the alignment of the three targeted variables, which inevitably involve different types of stakeholders. Yet, despite its potential, few studies have adopted this theory to explain longitudinal changes in one country’s language policies. Finally, in terms of the targeted population, I focus on EFL education for Grade 12 (18 years old) because over 95% of Japanese students proceed to senior high school (Grades 10–12) after the compulsory Grades 1–9, and most receive English education from Grade 5 (since 2012) or Grade 7 (between 1945 and 2011). Furthermore, because only about 50% proceed to tertiary education, Grade 12 English proficiency is seen as the end-product of Japanese formal English education.

This study consists of two parts: Study 1 targets the two periods (1994–1998 and 1999–2012) covered by two Courses of Studies (CoS), or sets of legally-binding curriculum guidelines, while Study 2 targets 2013 to the present. Comparing these three periods nationwide is valid because Japan has had a centrally controlled educational system since CoSs became legally binding in 1958 (Imura, 2003).

Study 1

In Study 1, I asked the following questions for each targeted period:

- (1) How well did goals, classroom practices (i.e., teachers’ activities and students’ understanding of the class content), and student achievement align for EFL education Grade 12 in Japan?

- (2) If the three key aspects (goals, classroom practices, student achievement) aligned, what might be possible reasons for this match?
- (3) If the three key aspects did not align, what might be possible reasons for this mismatch?
- (4) Does Roger's (2003) Diffusion of Innovation Theory help us better understand the reasons discussed in (2) and (3)?
- (5) How do answers to Research Questions 1–4 for 1994–2002 compare with those for 2003–2012?

Methods

Data

To answer Questions 1–3, I chose the data described in A below as the most appropriate evidence for goals, those in B for student achievement, and those in C for classroom practices. Justification for their selection for this study is discussed in their respective sections.

A Goals: English I – 1994–2002 and 2003–2012 Courses of Studies

As explained earlier, a CoS is a set of curriculum guidelines promulgated by the Japanese Government (through the Ministry of Education, Culture, Sports, Science, and Technology – MEXT). Although a CoS is created for each level from kindergarten (age 3–6), elementary school (Grades 1–6), junior high school (Grades 7–9), and senior high school (Grades 10–12), those promulgated in the same year (thus forming a set of CoSs even though some may become effective one or two years apart) were created under the same set of governmental goals and policies. Since first promulgated in 1947, CoSs have been revised seven times at approximately ten-year intervals to accommodate sociocultural changes.

B 2002 and 2005 Senior High School Test of English Proficiency and Student Surveys

In November 2002 and 2005, the National Institute for Educational Policy Research (NIER) administered CR tests and surveys to Grade 12 students and their teachers to ascertain how far the basic goals set by the CoS for those periods (1994–2002 for the 2002 test and 2003–2012 for the 2005 test) had been achieved in seven subjects (e.g., Japanese, mathematics, physics) in 2002 and ten subjects (e.g., geometry, history, and civics in addition to the seven subjects tested in 2002) by Grade 12 students who had studied under each CoS for the full three years of senior high school. The 2002 and 2005 test items were written and verified to measure the main points aimed at in the CoS by committees of experts in each subject (NIER, 2004, 2007).² These tests therefore represented the first attempts to ascertain the accountability of Japan's EFL educational policy from a CR perspective. Furthermore, the results represent the proficiency of Grade 12 students in 2002 since 8% of them (105,000) and 13% (150,537) of 2005 Grade 12 test takers were randomly selected from all public and private senior high schools throughout Japan. For English, the selected subject was *English I*, which all test takers ($n = 31,189$ for 2002 and $n = 29,880$ for 2005) took by the time they reached Grade 12. No other test of a similar kind was conducted on such a scale.

C 2002 and 2005 Senior High School Teacher Surveys

When the two tests described above were administered, all teachers of participating students (English subtest: $n = 891$ for 2002; $n = 887$ for 2005) responded to questionnaire items, some of which asked to what extent they had conducted the activities specified for English I in the CoS for the period. No other survey was conducted on such a scale.

Results

Goals for English I: 1994–2002 and 2003–2012

The two targeted CoSs shared three general characteristics that differed from previous ones: redefinition of academic ability, introduction of CR measurement, and further advancement of a liberal, flexible, and comfortable school life (the *yutori* orientation). First, academic abilities to be achieved by Grade 12 were redefined as “motivation and attitude toward learning and the ability to solve problems as an autonomous individual responding to societal changes” (Kariya, 2002, p. 56, author’s translation). This was a drastic change from the earlier CoSs, which tended to see academic abilities in terms of knowledge and skills (Abiko, 1996). For English, the term “communication” was introduced as part of “attitude” for the 1994–2002 CoS and added to the abilities to be developed under the 2003–2012 CoS. The second innovation of the two CoSs was related to the first. For the first time in its history, the 1994–2001 CoS required teachers to use CR measurement in classrooms to further promote the cultivation of the above-mentioned newly defined academic abilities (Sasaki, 2008). These changes reflected a response to nationwide criticism of the excessive emphasis on knowledge cramming, which was widely blamed for education-related problems such as dropping out (Mizuhara, 2010). This orientation dated back to the 1982–1993 CoS, under which class hours were cut, but the cuts that took place under the 2003–2012 CoS were even more drastic, reducing curriculum content throughout elementary and secondary education by 30%. This was intended to create more room in which to cultivate the “liberal, flexible, and comfortable school life” aimed for by the CoS by making the content to be learned easier.

Table 1 indicates that *English I*, the only English subject required for high school graduation in the 1994–2002 and 2003–2012 CoSs, shared similar objectives and suggested classroom activities. The only differences were that the 2003–2012 objectives and suggested activities were more specific and more oriented toward skills integration than in 1994–2002. Below I focus on the linguistic aspects of this objective because data related to students’ attitude are unavailable for classroom activities or outcomes.

Teaching practices: English I

Table 1 shows the percentage of English teachers who responded to the 2002 and 2005 surveys and reported having conducted the communicative activities suggested by the relevant CoS. Although the questions in the 2002 survey were less specific than those in the 2005 survey, they asked whether participating teachers conducted the four-skills activities listed in the 1994–2002 CoS. As Table 1 shows, percentages for all these activities are high (92.9–100%) compared to the corresponding ones for 2003–2012. The relatively low percentages (16.7 to 51.2%) of teachers who conducted the suggested activities for 2003–2012 may be due to the more specific wording of the four questions,

Table 1 Objectives, Suggested Communicative Activities, and Actual Classroom Practices for the 1994–2002 and 2003–2012 Courses of Study*

Course of Study	General objectives	Suggested communicative activities	% of teachers who (often or sometimes) conducted the suggested activities (2002 Survey: $n = 836-846$ ** 2005 Survey: $n = 887$)	% of students who understood the suggested activities (2002 Survey: $n = 25,355$)
1994–2002	To develop students' basic abilities such as understanding what they hear and read, expressing their own ideas, and fostering a positive attitude toward communication through English	A. Listening 1. Understanding what they hear 2. Understanding the gist of a coherent text 3. Summarizing their ideas based on what they hear	A-1: 96.0 for helping students understand what they hear	29.0
		B. Reading 1. Reading silently or aloud and understanding the meaning of the text 2. Understanding the gist of a coherent text 3. Reading a coherent text as fast as possible 4. Reading aloud to express the meaning of the text more efficiently	B-1: 100.0 for helping students understand what they read	40.3
		C. Speaking 1. Asking and answering appropriately based on what they hear in the given situation and for the given purpose 2. Speaking their ideas about what they read 3. Speaking the gist of what they understand	C in general: 92.9 for helping students say what they want to say	20.4
		D. Writing 1. Writing what they hear 2. Writing the gist of what they hear 3. Writing the gist of what they read 4. Writing the gist of what they want to express	D-4: 94.1 for helping students write what they want to express	22.3
2003–2012	To develop students' basic communicative abilities such as understanding what they hear and read, speaking and writing the information gained along with their own ideas, and fostering a positive attitude toward communication through English	A. Listening for understanding the given information and the speaker's intention or grasping the main points	38.2	Not Surveyed (N/S)
		B. Reading for understanding the given information and the speaker's intention or grasping the main points	51.2	N/S

Table 1 Objectives, Suggested Communicative Activities, and Actual Classroom Practices for the 1994–2002 and 2003–2012 Courses of Study* (*Continued*)

Course of Study	General objectives	Suggested communicative activities	% of teachers who (often or sometimes) conducted the suggested activities (2002 Survey: <i>n</i> = 836–846,** 2005 Survey: <i>n</i> = 887)	% of students who understood the suggested activities (2002 Survey: <i>n</i> = 25,355)
		C. Discussing and exchanging their ideas or the information gained through listening and reading	16.7	N/S
		D. Writing by sorting their ideas or the information gained through listening and reading	18.8	N/S

* All statements cited in this paper were translated by the author

** Numbers vary depending on the number of teachers who responded to some or all of the items

which closely reflected the activities emphasizing the four-skills integration recommended by that CoS. Although many of the 2002 teachers had their students listen, read, speak, and write in their classrooms, they may not have conducted such integrative activities as often. Interestingly, the ranking based on the percentage of students who understood the content of these activities, which are only available for the 2002 survey, were similar to those for 2003–2012.

Outcomes of classroom activities represented in test results

Table 2 presents brief specifications, percentages of those who answered each item correctly, and expected accuracy rates (i.e., if the teacher spent the standard amount of time covering activities suggested by the current CoS for cultivating the skill and knowledge measured by the item; NIER, 2004, 2007)³ for the 26 items for English I in the 2002 and 2004 senior high school tests. Both the 2002 and 2005 tests contained 10 items measuring listening skills, 9 measured reading skills, and 7 measured writing skills. The listening and reading items were in multiple choice format, and the writing items were descriptive (requiring test takers to write out answers). However, how the answers were rated is not revealed. There was no speaking item. Lastly, the difficulty level of the 2002 and 2005 tests can be compared because they shared 21 of the 52 items.

The information revealed in NIER documentation (2004; 2007) in addition to that presented in Table 2 can be summarized as follows:

- (1) In both tests, the reading section had the highest accuracy rate, the listening section the second highest, and the writing section the lowest.
- (2) In both tests, the reading section had the highest number of items whose accuracy rates were higher than expected rates (14/18 for 2002 and 12/18 for 2005), the listening section the second highest (8/20 for both tests), and the writing section the lowest (6/14 for 2002 and 4/14 for 2005).
- (3) Students were especially weak at responding to Item 20, which required them to write a coherent text consisting of more than three sentences. Only about 20% of test takers could write such a text (compared to the other writing items, whose mean scores were about 50%).

Table 2 Specifications, accuracy rates, and other information regarding the English items in the 2002 and 2005 senior high school tests*

Item no.	Main skill to be measured	Ability to be tested	Item Type	2002 Test: 1994–2002 Course of Study (n = 31,189)		2005 Test: 2003–2012 Course of Study (n = 29,880)	
				Mean % of those who answered the item correctly	Expected mean accuracy rate	Mean % of those who answered the item correctly	Expected mean accuracy rate
1–3	Listening	Ability to respond when spoken to	MC**	64.4	67.5	61.9	65.8
4–6		Ability to comprehend the details of a spoken text	MC	58.6	65	58.6	66.7
7–10		Ability to comprehend the main points of a spoken text	MC	62.1	62.3	65.2	63.2
11–12	Reading	Ability to comprehend the details of a written text	MC	64.9	62.5	66.5	63.7
13–15		Ability to comprehend the main points of a written text	MC	69.5	60.0	66.9	61.7
16–17		Ability to understand the logical development of a written text	MC	55.9	57.5	65.5	60
18–19		Ability to understand the writer's intention	MC	70.9	62.5	60.6	65.0
20	Writing	Ability to write a coherent text of more than three sentences	Open-ended	22.8	45.0	23.4	45
21–23		Ability to fill in blanks with appropriate words or phrases	Open-ended	52.9	55.0	47.8	55.0
24–26		Ability to write a sentence using scrambled words	Open-ended	51.4	55.8	50.2	60.8

* Means were calculated based on data presented in NIER (2004, 2007), with number of students taken into consideration

** MC: Multiple choice

- (4) Among the 21 overlapping items in the 2002 and 2005 tests, four, all of them listening items, showed significantly higher accuracy in 2005 (NIER, 2007).

Discussion of study 1

The results of Study 1, which investigated the alignment of goals, classroom practices, and student achievement in the two administrations of EFL education policy in Japan (1994–2002 and 2003–2012) reveal the following:

- A Although the CoSs valued the four skills equally, alignment between the three focal aspects was satisfactory only for reading abilities;
- B Alignment was better for reading and listening abilities than for writing and speaking abilities and improved significantly only for listening abilities;

- C Alignment for writing and speaking abilities was well below expectations;
- D Overall student achievement did not deteriorate in 2003–2012 even though the content for English was cut by 30%.

To answer Research Questions 1 to 3, I now discuss these results based on the five perceived characteristics of innovations, which Rogers (2003) argues are most likely to influence “different rates of adoptions” (p. 15) through innovation diffusion.

Relative advantage: Whether the innovation is perceived as better than its predecessor in terms of “economic terms, social prestige, convenience, and psychological satisfaction” (Rogers, 2003, p. 15)

We saw that the teachers in the two CoS administrations conducted writing and speaking activities much less often than reading and listening activities. Although students could study writing and speaking in English outside of the classroom (Cumming, 2009), they did not seem to have done so judging from the results of the 2004 Student Survey (Table 1) and the 2004 and 2007 Tests (Table 2). As many studies (e.g., Butler, 2015; Hu & McKay, 2012) have pointed out, this can be explained by the lack of immediate need for the students to use English outside of the classroom in a country such as Japan, where English is not used for general communicative purposes. O’Ki (2015) reported that in his needs analysis of 580 Grades 10–12 Japanese students, the three most popular reasons for studying English were: (1) high school graduation (44.8%); (2) university entrance exams (43.3%); and (3) Japan’s internationalization (35.7%). These results are especially important in analyzing those of this study because 56.6% of the 580 participants answered that they would not need English for university entrance exams. This suggests that many Japanese students study English mainly to obtain a high school diploma but not with university entrance exams in mind. This is in sharp contrast with East Asian countries such as China or South Korea, where many parents spent extravagant sums on cultivating their children’s communicative (mainly oral) abilities outside of formal education (Butler, 2015; Hu & McKay, 2012). Perhaps Japanese people believe that advanced learning of a foreign language may lead to a loss of identity as Japanese (Butler & Iino, 2005), which is closely related to the next characteristic of innovation, namely *compatibility* with societal values.

Compatibility: Is the innovation perceived as being consistent with existing values, past experiences, and needs?

Two findings of this study can be explained by compatibility between the 1994–2002 and 2003–2012 CoSs. First, the finding that goals, teaching practices, and student achievement aligned best with reading abilities throughout the two administrations was consistent with the fact that grammar-translation, which is exclusively based on reading, has been long cherished in Japan (Sasaki, 2008). After English became a virtually compulsory subject for Grades 7–9 in 1947, the method remained widely used in English classrooms in Japan while the more current 2003–2012 CoS was being implemented (Tahira, 2012). Teachers probably used the method based on reading activities because they were familiar with it.

A second finding that can be explained by compatibility was that the students’ achievement did not significantly deteriorate (Table 1) even after content was cut by

30% in 2003–2012. This may be because this policy had been severely criticized by almost all stakeholders (i.e., teachers, parents, and students; Imura, 2003), and MEXT consequently revised the 2003–2012 CoS as early as 2003, steering it once again in a more meritocratic direction (Butler & Iino, 2005). Because many of these changes were revivals of the 1994–2002 CoS, it may have been easier for teachers to readopt them. Yet the question remains as to why this substantial improvement was concentrated in listening abilities and not in other skills.

Complexity: Is the innovation difficult to understand or use?

As mentioned above, two of the most innovative aspects of the 1994–2002 and 2003–2012 CoSs were their promotion of communicative language teaching and the introduction of CR tests. As regards the former, Tahira (2012) reported that it was neither understood nor implemented in EFL classrooms over the previous 20 years mainly due to the nature of “communicative approaches,” which allowed for “many interpretations” (Brown, 2007, p. 45). The additional requirement of the 2003–2012 CoS that the four skills be taught in an integrated manner seems to have further added to teachers’ confusion judging from drastic falls in the percentage of self-reported teaching practices that satisfied the requirement in 1994–2002 relative to 2003–2012 (Table 1). Yet the fact that the teachers were accustomed to devising reading activities (the *compatibility* factor) must have helped them implement their integrated reading and listening activities more than other skill activities (51.2% and 38.2%, respectively), as indicated in Table 1. The other innovative aspect, namely the use of CR tests, is also known to have caused confusion and even shock when first introduced because the teachers were accustomed to traditional norm-referenced measurement (Sasaki, 2008).

Trialability: Can the innovation be tried on a limited basis (i.e., part by part)?

We saw that alignment between the three targeted aspects of EFL curriculum policy in Japan was fairly successful for reading abilities for both 1994–2002 and 2003–2012 and for listening abilities for 2003–2012. This partial success can be attributed to the fact that the activities suggested by both CoSs were set up according to each of the four skills (i.e., part by part).

Observability: Are the results of the innovation visible to others?

One of the most observable results in English education is student achievement. In a country such as Japan, which values meritocracy highly (Butler, 2015; Kariya, 2002), university entrance exam results are critical. This tendency accelerated after 1979, when the government created a common exam all applicants to public universities had to take 2.5 months before high school graduation. As the numbers of universities that used the results of this exam (now called the “Center Exam”) for admission increased (i.e., all public university and 86% of private universities in 2016), its results are regarded as highly reliable for ranking participating universities ($n = 777$, 89% of all universities in 2016; MEXT, 2016), and become a recurrent topic in the mass media. Consequently, the number of graduates who went to more highly-ranked universities is considered key evidence of accountability and educational quality for all stakeholders involved in each high school. Although some universities require yet another test (mostly one or two months later than the Center Exam), many (e.g., Hirai, Fujita, &

Oki, 2013; Taniguchi, Nishigaki, Murakoshi, & Watanabe, 2014) argue that the Center Exam is the most influential university entrance exam in present-day Japan.

Given such highly valued social observability, we can confidently argue that the exclusive improvement in listening scores in the 2005 Senior High School Test was mainly due to the introduction of the listening section of the English subtest for the first time in its history. Before 2005, the Center English subtests measured applicants' knowledge in written form only. Sasaki's (2016) detailed analysis of the 2006 Center Exam revealed that "18% of the 50 main English test items appear to measure speaking-related ability (but indirectly through written texts), while 32% measure grammatical knowledge, and the rest (50%) measure reading-related ability" (p. 102). In contrast, in the newly introduced listening test, the "items mainly measured the participants' ability to listen to English." If participants in the 2005 Senior High School test hoped to enter university after graduating in March 2006, most had to take the Center Exam's English test in January 2006. Although the listening section was optional, 99% of test takers ($n = 492,555$) did so.

Because the University Entrance Exam Center, Japan (2013) announced the introduction of a listening section in the 2006 Center Exam as early as March 2000, applicants for the 2006 exam had plenty of time to prepare for that subtest, which probably raised the 2005 Senior High School Test takers' listening abilities significantly relative to their other English skills. This speculation is supported by the fact that additional exams some of the students had to take for admission to their preferred university also rarely measured their ability to write or speak (e.g., Nakano, 2009). We can therefore claim with some confidence that the content of university entrance exams mirrors the results of the 2005 Survey and the 2005 High School Test (Tables 1 and 2, respectively). No other sociocultural or academically-related event surrounding these students during 2002–2005 explain the significant improvement in student scores in the 2005 High School Test (Negishi, Matsuzawa, Sato, Toyota, & Nakano, 2010).

Study 2

The most recent CoS for senior high schools became effective in 2013. However, nationwide tests such as the 2005 Senior High School Test, which ascertained how closely this CoS was being implemented, were not available at the time of writing (March, 2017). The purpose of Study 2 is therefore to present the most recent – albeit incomplete – state of current EFL educational policy by investigating alignment between goals, classroom practices, and student proficiency using evidence available at the time of writing. Study 2 thus employed current goals as data along with classroom practices and their outcomes as reported in one large nationwide CR test (the 2014 Test; see below for details) and a Survey of the 2014 test takers' teachers (the 2014 Survey; see below) conducted from July to September 2014 "for the purpose of improving EFL education in Japan" (MEXT, 2015). However, unlike those reported in Study 1, the classroom practices and their outcomes examined in Study 2 should not align with the 2013 CoS because the 2014 Test and Survey were answered by Grade 12 students and their teachers in public schools only, all of whom were influenced by the previous 2003–2013 CoS since it became effective only for those who entered senior high school in 2013). Nonetheless, since the 2013 CoS was promulgated in 2009 and many other influential political steps related to EFL education were publicized and implemented by

2014 (see below), we may see the influence of such sociopolitical expectations on the outcomes of EFL classroom practices and their outcomes in 2014. I therefore present them as evidence of possible influence of the most recent goals set by the 2013 CoS.

A. Goals: English I – 2013 CoS

The 2013 CoS aims to achieve three overall goals: 1) To respond to public criticism that the reduction in curricular content in the past two CoSs led to a deterioration in the academic competence of Japanese students as exemplified by the results of global measures such as PISA (Mizuhara, 2010). In response, curricular content and class hours were restored to their 1973–1982 level. The second major goal is to further expand the academic ability redefined in the last two CoSs as “motivation and attitude toward learning and the ability to solve problems as an autonomous individual responding to societal changes,” as first suggested by the OECD in 2003. The final major goal is related to the second in that it emphasizes the cultivation of communicative abilities, especially foreign languages (mostly English). To achieve this goal, a total of 70 class hours called “Foreign Language Activities” (which are not graded) were introduced in Grades 5 and 6 for the first time in EFL education history in public schools in Japan. Curiously, however, the suggested classroom activities designed to achieve that objective are similar to those suggested in the 2003–2012 CoS (see the next section).

B. Teaching practices for achieving the goals set for English communication I

Fortunately, the 2014 Survey of teachers ($n = 2,493$ from 477 public senior high schools randomly selected throughout Japan) asked questions about their classroom activities that were similar to those asked in the 2005 Survey because they generally share similarly integrated activities for the 1994–2002 and 2013–2013 CoSs presented in Study 1. Table 3 presents the percentage of teachers who conducted such activities.

Although the ranking of the four skills in responses to the 2014 Survey is exactly the same as in 2005 (with reading showing the highest percentage and speaking the lowest), Table 3 shows that the percentages of teachers who started conducting activities as suggested by the 2003–2012 CoS increased for all four skills. Most noticeable is the

Table 3 Suggested Communicative Activities and Classroom Practices for the 2003–2012 and 2013–Present Courses of Study

Suggested Communicative Activities	% of teachers who (often or sometimes) conducted the suggested activities ($n = 887$): 2005 Survey	% of teachers who (often or sometimes) conducted the suggested activities ($n = 2,493$): 2014 Survey
A. Listening for understanding the given information and the speaker’s intention or grasping the main points	38.2	65.1
B. Reading for understanding the given information and the speaker’s intention or grasping the main points	51.2	89.2
C. Discussing and exchanging one’s ideas or the information gained through listening and reading	16.7	33.0
D. Writing by sorting one’s ideas or the information gained through listening and reading	18.8	39.7

doubling in the percentages for speaking and writing activities. The changes indicated in Table 3 suggest that teaching practices can be gradually geared toward closer alignment with goals under certain conditions.

C. Students' achievement as represented by test results

A total of 68,054 Grade 12 students took the reading and listening subtests, 69,052 took the writing subtests, and 16,583 also took the speaking subtest of the nationwide 2014 Test. The reading subtest consisted of 43 multiple-choice items (45 min), the listening part of 36 multiple-choice items (23 min), the writing part of two items, one of which required summarizing abilities and the other required the ability to write a convincing opinion composition (27 min), and the speaking part of three items, one of which required reading a text aloud, another the ability to answer questions, and the third the ability to logically explain one's opinions (10 min). Test takers are ranked as A1 (Beginner), A2 (Elementary), B1 (Intermediate), and B2 (Upper Intermediate) following CEFR (Common European Framework of Reference for Languages) levels.

Compared with the results of the 2014 Teacher Survey, the results of the 2014 test were disappointing as the students' reading and listening abilities clustered around upper A1 (Lower basic) to lower A2 (Upper basic), their speaking abilities at lower A1, and their writing abilities at the lowest A1 level (see also Council for Cultural Cooperation Education Committee Modern Languages Division, 2001). Furthermore, 13.3% of test takers did not say anything for the speaking items, and 29.2% of them did not write anything for the writing items.

Discussion of study 2

I discuss the results of Study 2 in terms of Roger's (2003) five perceived characteristics of innovations. The 2014 test takers' relatively low scores for all four skills measured by the 2014 Test may be due to the fact that they were measured against CEFR specifications, which included the ability to think, judge, and express oneself in addition to purely linguistic proficiency (MEXT, 2015), which had not been required by the 2003–2012 CoS. Although the current CoS emphasizes such abilities, as we saw in Study 1, the diffusion of an innovative policy is slow when its content is complex. Because the current CoS is based on a further redefinition of academic abilities, it might have made it more difficult for the English teachers to understand its benefits (relative advantages) and how they could put it into practice (complexity) than the 2003–2012 version, which had already proved difficult to understand (Tahira, 2012).

That said, the results of the writing part of the 2014 Test still resonate with the results of the 2005 Senior High School Test, in which the writing score for Item 20 requiring Grade 12 test takers to write a coherent text longer than three sentences, was much lower than those for the other descriptive writing items, not to mention the listening and reading items. Furthermore, based on the 2014 Test results, we can easily imagine that Grade 12 students in 2015 and 2014 lacked the ability to speak a coherent text as well. Because the latest 2017 Center Exam still does not measure any performative types of writing or speaking abilities, as speculated in Study 1, this lack of alignment between governmental goals and student proficiency can again be attributed mainly to the ways in which the high-stakes university entrance exams are conducted (compatibility and observability).

The only solid evidence of alignment between goals and teaching practices in the current CoS is the finding that more teachers conducted communicative writing and speaking activities for participants in the 2014 Test than in 2005. Although such efforts to improve particular aspects of the students' English proficiency did not result in improvements in the short term, they may be reflected in their achievement if measured later. Nevertheless, these results are in stark contrast with the fact that the 2005 test takers' listening scores were significantly higher than their 2002 counterparts only three years later and despite the fact that the content of the English curriculum was cut by 30% during those three years. Compared to the slow improvement in writing and speaking in Japanese EFL education over the past ten years, this further emphasizes the power of high-stakes tests when a strong meritocratic social discourse embraces them (Shohamy, 2001).

In sum, the diffusion rates of the new EFL educational policies in Japan over the past 20 years seem to have been affected to some degree by all five perceived characteristics of innovations, as predicted by Innovation Diffusion Theory (Rogers, 2003). Among the five, however, *Relative advantage* and *Observability* related to what Butler (2015) called the two "societal factors" (p. 305) characterizing Japan and other east Asian EFL countries were far more powerful than the other three, namely: (1) English is not used for authentic communicative purposes outside classrooms (related to *Relative Advantage*); and (2) success in high-stakes exams is believed to lead to social success and even happiness (related to *Observability*).

Conclusions

This study revealed how well the goals, classroom practices, and student achievement aligned for EFL education in Japan over three administrative generations starting in 1994. Rogers' (2003) Diffusion of Innovation theory was useful in analyzing the acceptance rate of each of the three administrations' new policy along with the five perceived characteristics of innovations (relative advantage, comparability, complexity, trialability, and observability) as its tools. In sum, all three targeted aspects of curriculum policies seem to have been affected to some degree by all five characteristics of the newly introduced curriculum guidelines even though *relative advantages* (English does not have high social advantages for Japanese students) and *observability* (university entrance exams best motivate the students to study English) were more powerful than the others.

Based on these findings, this section makes suggestions for EFL education in Japan as well as for future studies. Even though the suggestions for improving EFL policies are limited to the Japanese system, some may be useful in other contexts with similar backgrounds. First, alignment between curriculum guidelines and teaching practices could be greatly improved if new guidelines were to become more adaptable in terms of Rogers' (2003) five perceived characteristics of innovations. For example, a new goal will be more easily accepted and implemented if its promoters provide the adopters with convincing reasons (i.e., a clear *relative advantage*) for its adoption accompanied by necessary training (involving less *complexity*). The fact that more teachers started conducting communicative activities for all skills (especially speaking and writing) in the 2014 Survey compared to 2005 (Table 3) may be a slow but steady consequence of a series of ad hoc governmental policies such as the *Action Plan to Cultivate Japanese with English Abilities* that started in 2003. This plan included sending at least 300 English teachers abroad each year as well as providing one-month in-service training

for at least 2000 English teachers in 2002–2009 (see Butler & Iino, 2005 for details). Given such public aid, faster adoption of curriculum guidelines is possible especially because the guidelines can be characterized as highly *trialable*; that is, writing and speaking abilities, which need much more improvement than the other skills, can be emphasized in classrooms.

Second, I turn to the most influential aspect of Japan's EFL educational policy, namely its *observability*, as realized through high-stakes university entrance exams. In fact, MEXT may well have concurred, because in December 2014, the Central Council of Education (CCE, 2014), a subdivision of MEXT, suddenly announced the introduction of new university entrance exams starting in 2020. As regards English admission tests for universities, MEXT also encourages universities to employ commercial tests (e.g., TOEFLIBT) that measure all four skills in their own entrance exams (MEXT, 2013). If implemented, this should positively influence the students' writing and speaking skills. Of course, we should be aware of potential downsides to such a revolutionary change. For example, commercial cram schools may respond most quickly and develop strategies designed to achieve high scores on such items, thus widening the gap between those who can afford to attend such schools and those who cannot. The government should therefore prepare for such an eventuality by financially supporting those who cannot afford to attend such schools.

Finally, this study is limited in various ways, which need to be complemented by future studies. I focus here on the targeted populations, theoretical framework, and methodology. One of the limitations of the study is its exclusive focus on one country's educational policies. Unless we replicate this study using changes in other countries' EFL educational policies, we cannot be certain that the implications of this study are applicable to other contexts. Although sociocultural and historical differences (e.g., the degree of centralization of educational policies) should be carefully considered, comparison such as that conducted by Butler (2015), who compared EFL education for young learners across four East Asian countries is a promising future direction. In addition, since the current CoS introduced "foreign language activities" in Grades 5 and 6, the impact of such earlier exposure to foreign languages (mainly English) should also be investigated and compared with the results of previous CoSs, such as those targeted in this study.

Another limitation of the study is its methodology since it only used as data the results of publicly reported results of tests and questionnaires conducted by the government. Although the data have their own merits, including generalizability (e.g., the data come from randomly selected schools from among all public and private high schools in Japan), such quantitative data should be supplemented by emic qualitative data. For example, a study can utilize data probing how Grade 12 students feel about their EFL educational histories and how their teachers perceive the newly introduced policies. Inclusion of the students' English learning experiences outside classrooms is also necessary if we are to understand their entire learning experiences and their consequences. Finally, the impacts of new policies should be investigated through ecologically valid theories. In this study, Rogers's (2003) Diffusion of Innovation theory was helpful in analyzing the impacts of the three consecutive educational policies, and the theory should be adopted by more studies of different populations across different situations. Alternatively, other

useful theories such as Kaplan and Baldauf's (2005) "model for language-in-education policy planning" employed in Butler (2015, p. 306), or a more recent view of *English as a lingua franca* (e.g., Kirkpatrick, 2012) should be employed to evaluate not only policy implementation but also the policies themselves as frameworks for future studies.

Endnotes

¹CR-type measurements gauge a person against standards or benchmarks, as opposed to norm-referenced measurement, which gauges a person against any others who happen to take the same test on that occasion (Bachman, 1990).

²NIER (2004) is a composite of five official reports containing explanations and the results of the 2002 Senior High School Test, and NIER (2007) combines six similar reports. Regrettably, the data revealed in these two composite documents are not sufficient (e.g., they did not include SD for the means) for the author to conduct any statistical procedures.

³The term "accurate" is defined as the answer being exactly as expected or judged to be correct when considering what the item is intended to measure in light of the goals and content of the 2003–2012 CoS (NIER, 2007).

Acknowledgements

I would like to thank Paul Bruthiaux for his valuable comments and suggestions.

Funding

Preparation of this article was aided by an Abe Fellowship 2016–2018 granted through the Social Science Research Council and the Japan Foundation Center for Global Partnership.

Competing interests

There are no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 22 September 2017 Accepted: 5 December 2017

Published online: 16 January 2018

References

- Abiko, T. (1996). *Shin gakuryokukan to kisogakuryoku: Nani ga towareteiruka [New perspectives on basic academic ability: What is being examined?]*. Tokyo: Meijitoshō.
- Alderson, J.C. (2005). *Diagnosing foreign language proficiency*. London: Continuum.
- Bachman, L.F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Brown, J.D. (2007). *Teaching by principles: An interactive approach to language pedagogy*, (3rd ed.,). New York: Pearson Education.
- Butler, Y.G. (2015). English language education among young learners in East Asia: A review of current research (2004–2014). *Language Teaching*, 48(3), 303–342.
- Butler, Y.G., & Iino, M. (2005). Current Japanese reforms in English language education: The 2003 action plans. *Language Policy*, 4, 25–45.
- CCE. (2014). Atarashii jidaini fusawashii koudaisetzokuno jitsugennimuketa koutougakkou, daigakuyouiku, daigakunyugakusha no itkaiteki kaikakunitsuite [a reform designed to unify education in senior high schools and universities and university admissions procedures and implement appropriate connections between senior high school and university education for a new era]. Retrieved from: http://www.mext.go.jp/b_menu/shingi/chukyo/chukyo0/toushin/_jcsFiles/afiedfile/2015/01/14/1354191.pdf
- Council for Cultural Co-operation Educational Committee Modern Language Division (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.
- Cumming, A. (2001). The difficulty of standards, for example in L2 writing. In T. Silva, & P.K. Matsuda (Eds.), *On second language writing*, (pp. 209–229). Mahwah, NJ: Lawrence Erlbaum.
- Cumming, A. (2009). Language assessment in education: Tests, curricula, and teaching. *Annual Review of Applied Linguistics*, 29, 90–100.
- Dingfelder, H.E., & Mandell, D.S. (2011). Bridging the research-to-practice gap in autism intervention: An application of diffusion of innovation theory. *Journal of Autism and Developmental Disorders*, 41(5), 597–609.

- Harper, C, Platt, E, Naranjo, C, Boynton, S. (2007). Marching in unison: Florida ESL teachers and no child left behind. *TESOL Quarterly*, 41(3), 642–651.
- Hirai, A, Fujita, R, Oki, T. (2013). Senter lisuning ga motarasu lisingu gakushuuiyokueno eikyou: Daigakushubetsu, nyuushikeitai, senkougotono bunsekini motozuku kousatsu [The influence of the center listening test on listening learning motivation: An analysis focusing on university type, admission type, and major]. *JACET Journal*, 57, 59–81.
- Hu, G, & McKay, SL. (2012). English language education in East Asia: Some recent developments. *Journal of Multilingual and Multicultural Development*, 33(4), 345–362.
- Imura, M (2003). *Nihon no eigokyouiku nihyakunen [200 years of English language education in Japan]*. Tokyo: Taishuukan.
- Kaplan, RB, & Baldauf, RB (2005). Language-in-education policy and planning. In E Hinkel (Ed.), *Handbook of research in second language teaching and learning*, (pp. 1013–1034). Mahwah, NJ: Lawrence Erlbaum.
- Kariya, T (2002). *Kyouiku kaikaku no gensou* [Disillusion over educational reforms]. Tokyo: Chikumashobou.
- Kirkpatrick, A. (2012). English as an Asian lingua Franca: The “lingua Franca approach” and implications for language education policy. *Journal of English as a Lingua Franca*, 1(1), 121–139.
- Lee, AY, Hsieh, Y, Hsu, C. (2011). Adding innovation diffusion theory to the technology acceptance model: Supporting employees’ intentions to use E-learning systems. *Educational Technology & Society*, 14(4), 124–137.
- MEXT. (2013). Guroobaru jinzai ikusei ni fansuru seisaku hyouka [Program evaluation regarding workforce cultivation for a globalized world]. Retrieved from: http://www.mext.go.jp/a_menu/keikaku/detail/_icsFiles/afieldfile/2013/06/14/1336379_02_1.pdf
- MEXT. (2015). Heisei 26 nendo eigokyouiku kaizen notameno eigoryokuchousa jigyouhoukoku [Results of the English test conducted to improve English education in Japan in 2015]. Retrieved from: http://www.mext.go.jp/a_menu/kokusai/gaikokugo/1358258.htm
- MEXT (2016). *Heisei 28 nen gakkou kihon chousa [2016 basic school survey in Japan]*. Tokyo.
- Mizuhara, K (2010). *Gakushuushidouyouryou wa kokuminkeisei no sekkeisho* [The courses of study are blueprints for forming desired citizens]. Sendai: Touhoku Daigaku Shuppankai.
- Nakano, T (2009). Bunryohen: Dokkai, eisakubun, lisuning [The issue of quantity: Reading, writing, and listening items]. In K Kanatani (Ed.), *Kyokasho dakede daigakunyuuishi wa toppa dekiru* [Passing university entrance exams with English textbooks only], (pp. 99–168). Tokyo: Taishuukan.
- Negishi, M., Matsuzawa, S., Sato, R., Toyota, Y., & Nakano, T. (2010). Daigaku nyuushi ga kawareba eigokyouiku mo kawarunoka: Zadankai [Will English instruction in Japan change if university entrance exams change: A discussion]. *The English Teachers’ Magazine*, August, 10–19.
- NIER. (2004). Heisei juuyonendo koutougakkou kyoiuku katei jisshi joukyou chousa [Results of the 2002 survey of the implementation of the senior high school curriculum]. Retrieved from: http://www.nier.go.jp/kaihatsu/katei_h14
- NIER. (2007). Heisei juunanendo koutougakkou kyoiuku katei jisshi joukyou chousa [Results of the 2005 survey of the implementation of the senior high school curriculum]. Retrieved from: http://www.nier.go.jp/kaihatsu/katei_h17_h
- O’Ki, T. (2015). Tekisuto mainingu o mochiita koukousei eigogakushuushano niizu bunseki: Diagakujkenyoteisha tohiyoteisha no hikaku [Needs analysis of high school English learners using text mining: A comparison between university examinees and non-examinees]. *Hakuo University Ronshuu*, 29(1,2), 193–216.
- Rogers, EM (2003). *Diffusion of innovations*, (5th ed.,). New York: Free Press.
- Sasaki, M. (2008). The 150-year history of English language assessment in Japanese education. *Language Testing*, 25, 63–83.
- Sasaki, M (2016). English writing instruction in senior high schools: A historical ecological approach. In T Silva, J Wang, J Paiz, C Zhang (Eds.), *L2 writing in the global context: Represented, underrepresented, and unrepresented voices*, (pp. 84–109). Beijing: Foreign Language Teaching and Research Press.
- Shohamy, E (2001). *The power of tests: A critical perspective on the uses of language tests*. London: Pearson Education.
- Tahira, M. (2012). Behind MEXT’s new course of study guidelines. *The Language Teacher*, 36(3), 3–8.
- Taniguchi T., Nishigaki, C., Murakoshi, R., & Watanabe, M. (2014). Ankeeto kara mietekurumono: Shougakkou, chuugakkou, koukou, daigaku no tachibakara [What can be seen from the results of the questionnaire: From the perspectives of elementary schools, junior and senior high schools, and universities]. *The English Teachers’ Magazine*, October Special Issue, 6–11.
- University Entrance Exam Center, Japan. (2013). Rijiicho aisatsu [Greetings from the chair]. Retrieved from: <http://www.dnc.ac.jp/corporation/goaisatsu.html>

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com