

An investigation of factors involved in Japanese students' English learning behavior during test preparation

Takanori Sato
Sophia University, Japan

Japan has recently been promoting university entrance examination reform with the goal of positively influencing students' English learning, but the extent to which entrance examinations themselves affect English learning is not known. The promotion of better learning requires changing the factors that affect learning behavior, rather than merely modifying existing examinations or introducing new ones. This study investigated the factors determining Japanese students' English learning while they prepared for high-stakes university entrance examinations, aiming to construct a model that explicates how test-related and test-independent factors are intertwined. Semi-structured interviews were conducted with 14 first-year university students asking how they had prepared for their examinations and why they had chosen particular preparation methods. After thematic analysis, four main factors in student learning behavior were identified (examination, student views, school, and examination-independent factors) and their relationships explored. The study findings provide useful insights for policymakers in English as a foreign language (EFL) educational contexts, where English tests are used as part of language education policies. Furthermore, the proposed model is theoretically important as it explains the complex washback mechanism and deepens our understanding of why intended washback effects on learning are not necessarily achieved.

Key words: washback, entrance examination, English learning, test preparation, EFL

Introduction

High-stakes examinations influence students' learning and teachers' pedagogy (Tsagari & Cheng, 2017) and are thus often used as part of language education policies (Menken, 2017). In Japan, high-stakes English examinations have been used as a driving force in promoting English use and enhancing students' communication skills (e.g., Watanabe, 2013), as English is one of the core school subjects and students' English examination scores influence their admission to upper-level school (Vongpumivitch, 2014). University entrance examinations, which include a section on English proficiency, are particularly important for students. Numerous public and private high schools offer examination preparation classes, and there are even special schools that offer instruction in test-taking techniques. Furthermore, examinations in junior and senior high schools are often designed to align with these university entrance examinations (Vongpumivitch, 2014).

However, the current university entrance examinations have been criticized for creating a negative *washback* effect (the influence of examinations on teaching and learning) and blamed for low overall English proficiency. For example, the English section of a nationwide examination used for university admission purposes, the National Center Test for University Admissions (hereafter the Center Test), only measures language knowledge, reading comprehension, and listening skills through multiple-choice type questions (Watanabe, 2013). Furthermore, in-house English examinations developed and administered by individual universities for admission rarely measure applicants' writing and speaking skills. Researchers argue that this causes teachers and students to mostly focus on receptive skills rather than productive skills (Green, 2014). To improve this situation, the Ministry of Education, Culture, Sports, Science and Technology in Japan (MEXT) has recently been promoting university entrance examination reform, involving the implementation of a new test to replace the Center Test in 2020 and the use of external English tests such as TOEFL or IELTS for university admission, with the importance of assessing all four language skills (reading, listening, writing, and speaking) being highlighted (MEXT, 2017). It is believed that these efforts will lead to a positive change in English education in Japan. The examination reform is expected to make teachers and students focus more on productive language use and learn the four language skill areas in a

more balanced way.

Currently, however, the extent to which the examinations themselves affect Japanese students' English learning activities outside the classroom remains unknown due to a dearth of empirical studies on the washback of testing in Japan. If students' learning methods are not directly and strongly affected by the examinations, the reform is not likely to directly change how students study English. To promote better learning, it is necessary to address factors that affect student learning behavior, rather than arbitrarily modifying existing examinations or introducing new tests.

This study investigated the factors determining Japanese students' English learning while they prepared for high-stakes university entrance examinations, aiming to construct a model that explicates how various intertwined test-related and test-independent factors influence students' learning. The findings of this study may provide useful insights for policymakers who implement English tests as part of language education policies in Japan. Furthermore, the present study is theoretically important as it addresses the complex washback mechanism of examinations on learning behavior, examining the factors that influence students' English learning and the relationships between those factors (Tsagari & Cheng, 2017).

Literature Review

University Entrance Examinations in Japan

Applying to Japanese universities consists of several procedures, including interviews, essays, teachers' recommendations, and admission office examinations (Sasaki, 2008). However, as the nationwide university entrance examinations widely prevail, they are the most influential (Vongpumivitch, 2014). As Japanese high schools do not administer school-leaving examinations, these entrance examinations are crucial for those who aim to enter university. All Japanese national and local public universities (approximately 20% of Japanese universities) require students to take two examinations: (a) the Center Test, administered in January, which is taken by more than 500,000 students each year, and (b) an in-house examination, developed and administered by each university in February or March, before the start of the Japanese school year in April. On the

other hand, private universities (approximately 80% of Japanese universities) administer their own examinations in February or March without necessarily requiring candidates to take the Center Test, although more than 80% of private universities in Japan currently employ the Center Test. While some private universities rely only on the results of the Center Test or their in-house examination to determine whether students are granted admission, others use the results from both tests to make their decisions. Because the examinations offered by each university are administered on different dates, students can take more than one examination. For example, if a student wishes to attend University X (a private university using both the Center Test and in-house examination results) or University Y (a private university using only in-house examination results), he or she needs to take the Center Test, apply for admission into the universities, and take the in-house examinations of Universities X and Y. In fact, it is common for students to take examinations for several universities that they wish to attend and to prepare for multiple examinations.

The English section of the Center Test consists of a written component, which measures candidates' language knowledge and reading comprehension skills (for a total of 200 points), and a listening component (for 50 points). More specifically, it assesses knowledge of pronunciation, vocabulary, and grammar, as well as the comprehension of conversations, lectures, stories, and expository passages. It is constructed based on the study guidelines for high schools prescribed by MEXT, as the test is designed to measure students' knowledge of the content they learned at high school. At the same time, it serves as a certification test to judge if students have the ability required to take the second-stage examination that each university administers (Watanabe, 2013). The Center Test's content is thus covered by authorized textbooks used at high schools. In comparison, the format, content, and difficulty of the second-stage, in-house English examinations vary considerably among different universities. While many in-house examinations adopt selected response tasks, others use limited or extended production tasks. Furthermore, almost all the in-house examinations assess reading comprehension as well as grammar and vocabulary knowledge. At the same time, as educators acknowledge the importance of English communicative ability, the in-house examinations assess students' listening skills and, less frequently, writing skills. Students' speaking skills are still rarely measured (Vongpumivitch, 2014).

As part of the examination reform movement in Japan, the Test of English for Academic Purposes (TEAP) was developed by Sophia University and the Eiken Foundation of Japan in collaboration with the Centre for Research in English Language Learning and Assessment at the University of Bedfordshire (Eiken Foundation of Japan, n.d.-b) and was implemented in 2014. Currently, approximately 70 Japanese national and private universities use the TEAP as their English examination. The TEAP is different from the Center Test and conventional university entrance examinations in that it measures the academic English proficiency required for students to learn and conduct research at Japanese universities, testing all four language skills (Eiken Foundation of Japan, n.d.-a; In' nami, Koizumi, & Nakamura, 2016). The reading and listening subtests assess knowledge of vocabulary and grammar as well as comprehension of on-campus conversations, lectures, graphs and charts, e-mails, and research articles through multiple-choice format. The writing section employs integrated tasks in which students produce a 70-word summary of an academic passage and write a 200-word essay using multiple information sources that are provided. The speaking test is an approximately 10-minute face-to-face interview with an examiner in which students are required to answer questions about themselves, ask the examiner questions and give a speech. The difficulty of the TEAP is set from A2 to B2 of the Common European Framework of Reference (CEFR), and the content is based on the study guidelines prescribed by MEXT. It is administered three times a year in 12 cities in Japan; second- and third-year senior high school students may take the TEAP as many times as possible to reach the cut-off points set by each university. The test developer expects students to focus on the abilities measured by the test and engage in language tasks to develop these abilities (Green, 2014).

Few studies have been conducted on the degree to which the abovementioned high-stakes English examinations influence Japanese students' learning outside the classroom. Furthermore, it is also not known what other factors, if any, influence their learning during the examination preparation period.

Washback of Testing on Learning

The number of studies on how English language tests affect students' learning has gradually grown over the past two decades. Empirical studies conducted thus far have demonstrated that the most popular language learning activity during test preparation is practicing past or mock examination papers (Allen,

2016; Mickan & Motteram, 2009; Shih, 2007; Stoneman, 2005; Zhan & Wan, 2016). Students choose this learning method with the goal of successfully completing tests and obtaining high marks. Moreover, as students perceive test-oriented activities as more effective than communication-oriented tasks (Pan, 2016), they frequently focus on the specific items or tasks that appear in the given test. Thus, while test developers often wish to influence students' learning behavior by featuring more communicative test tasks (see Qi, 2007), students do not necessarily adopt communication-oriented learning activities to prepare for the test. For example, Pan (2014) compared the English learning behaviors of students who were required to take standardized university exit tests and those who were not, finding that the frequency of doing non-test-related language skill-building activities (such as reading English magazines and browsing English-language websites) was not significantly different between the two groups, although the former practiced mock tests and accessed online test preparation programs more frequently than the latter. This finding indicates that the washback effect of tests on student learning may be more limited than test developers generally expect (see also Green, 2007; Pan & Newfields, 2012).

Studies have also demonstrated that students engage in various English learning activities other than practicing past or mock examination papers when preparing for tests. Relatively popular activities are memorizing phrases and studying grammar and vocabulary (Damankesh & Babaii, 2015; Mickan & Motteram, 2009; Zhan & Andrews, 2014). One factor influencing their choice of activities appears to be the content of the given test; for example, when a test includes vocabulary questions, students focus on the vocabulary that appears on that particular test. In addition, students may believe that learning vocabulary and grammar is effective in improving their English proficiency (Pan, 2016). Still, not everyone studies vocabulary and grammar (Zhan & Andrews, 2014), and vocabulary being part of the test content does not necessarily lead to intensive vocabulary studying (Pan & Newfields, 2012). In addition to the abovementioned learning activities, students appear to engage in various activities that are not directly relevant to given test content or tasks, albeit to a small degree, such as listening to English news programs, reading newspapers, speaking English online, and gaining world knowledge (Allen, 2016; Damankesh & Babaii, 2015; Mickan & Motteram, 2009; Pan & Newfields, 2012; Zhan & Andrews, 2014). These findings suggest that the content of the test is not the sole factor influencing student learning behavior.

Several studies have identified a range of such factors with many not considered to be part of the washback effect itself, but rather factors mediating the washback. For example, students' learning appears to be affected by their perception of the given test's difficulty, importance, and construct (Allen, 2016; Green, 2007; Xie & Andrews, 2012; Zhan & Andrews, 2014; Zhan & Wan, 2016). This means that examinations influence students' perceptions, which in turn influence their learning activities. Green (2007) argues that washback on learning is intensified when examinations are perceived as important and moderately challenging. In addition, Zhan and Wan (2016) found that students' beliefs about what the test measures, which is not necessarily identical to what test developers intend to measure, influenced their choice of learning methods. This suggests that, while a change in test content and difficulty is likely to influence students' learning behavior, it may not do so because of the mediating effect of students' perception. Other mediating factors are closely related to the students' views of themselves as learners and their beliefs about effective learning methods (Allen, 2016; Xie & Andrews, 2012; Zhan & Andrews, 2014). Thus, students focus on their weaker skills and employ learning methods that may maximize their score. Xie (2015) claims that cramming and drilling may be inevitable as these methods boost test scores, regardless of test design.

Taking advice from peers and teachers was also identified as a mediating factor in washback on learning (Allen, 2016; Shih, 2007; Stoneman, 2005; Xie & Andrews, 2012; Zhan & Andrews, 2014; Zhan & Wan, 2016). Zhan and Wan (2016) found that teachers' coaching and instruction for a high-stakes listening/speaking test were reflected in students' preparation. Furthermore, previous studies showed that students receive advice from their teachers and peers, which influences their learning. Such advice from teachers is often more relevant to test preparation than English skill development and can disrupt the flow of washback (e.g., Allen, 2016; Shih, 2007). These factors thus indicate that teachers may play an important role in students' examination preparation behavior in certain educational contexts.

Other factors in the washback on learning found by previous studies are restrictions on resources and students' interests. Allen (2016) and Shih (2007) reported that students avoid preparing for speaking tests when they lack a partner with whom they can practice speaking. Moreover, students' choice of learning methods can be influenced by their interests (Allen, 2016; Zhan &

Andrews, 2014), which suggests that learning activities are not always influenced by exams. In summary, these mediating factors are considered to be relatively impervious to changes in test content, demonstrating the complexity of the mechanism of washback effects on learning.

Although previous studies have shed light on the mechanism of washback by exploring the factors influencing students' English learning behavior (Wei, 2017), few have addressed the relationships between the factors. This is a serious limitation because theories and models must explain the relationships between concepts rather than merely list them (Corbin & Strauss, 2008). A notable exception is the work of Xie and Andrews (2012), which tested a washback hypothesis based on expectancy-value motivation theory using Structural Equation Modeling (SEM). They statistically analyzed the relationship between students' knowledge of test demands, perceived test uses, expectation of success, perceived test importance, and test preparation practices. The data were self-reported by university students taking the College English Test (CET). It was found that students' self-efficacy and perceived test value acted as mediators between their knowledge of the test and their preparation practices. More specifically, greater knowledge of test demands led to a higher expectation of success and perceived test importance, which were associated with more engagement in test preparation. However, this study did not include many of the influential factors identified in the previous studies, including students' intrinsic factors and learning environment. Further hypotheses on the relationships between the relevant factors should be constructed based on exploratory studies and tested through confirmatory studies.

The Present Study

The present study investigated the factors determining Japanese students' English learning while they prepared for university entrance examinations, aiming to construct a model that explicates the relationships between the factors. The research questions (RQs) are:

1. What are the factors that determine Japanese students' university entrance examination preparation activities?
2. What are the relationships between these factors?

Method

This study forms part of a broader investigation of the washback of a recently implemented examination, the TEAP, to reveal whether and how it influenced student English learning practices. The data collection consisted of a large-scale online questionnaire survey on test preparation practices and interviews investigating the factors that influence these practices (see Sato, 2018). This paper reports on the findings of the latter phase.

Participants

The participants in this study were 14 first-year students who enrolled in a Japanese university in April 2016. This university contains nine faculties and is one of the prestigious private universities that emphasize foreign language education (Table 1). These participants were drawn from the 218 students who responded to the online questionnaire survey and indicated their willingness to participate in the interview study. The researcher selected students from different faculties who took a wide variety of examinations, as Japanese students need to study for different examinations depending on the departments they apply for. For their admission, the students had the choice of submitting their TEAP scores, taking an English examination administered by each department in February, or doing both. Regarding other subjects, all students took examinations developed by the department they wished to attend in February. While some departments required applicants to submit 4-skill TEAP scores, others required only TEAP scores for reading and listening; in other words, the 2-skill TEAP examination. Twelve participants also took the examinations from other universities and the Center Test, which indicates that the majority of participants prepared for and took multiple examinations. The university that these students attended did not require applicants to take the Center Test. In addition to their high schools, 10 participants went to cram schools (called *yobiko* or *juku*), where they prepared for the examinations and obtained various examination-related information based on meticulous analyses of past examinations (Watanabe, 2004).

Table 1. Information About Participants (N=14)

	Gender	Faculty	High School	Cram School	Center Test	TEAP	Exam of Other Univ.
A	Female	Human sciences	Public	Yes	Yes	4-skill	Yes
B	Male	Humanities	Private	Yes	Yes	4-skill	Yes

C	Female	Humanities	Private	No	Yes	4-skill	Yes
D	Male	Science and tech.	Private	Yes	Yes	No	Yes
E	Female	Human sciences	N/A	Yes	No	4-skill	No
F	Female	Foreign studies	Public	Yes	Yes	No	Yes
G	Female	Economics	Private	Yes	Yes	No	Yes
H	Male	Human sciences	Private	Yes	Yes	No	Yes
I	Female	Global studies	Private	No	Yes	4-skill	No
J	Female	Foreign studies	Private	No	Yes	2-skill	Yes
K	Male	Global studies	Public	No	Yes	4-skill	Yes
L	Female	Law	Private	Yes	Yes	4-skill	Yes
M	Female	Humanities	Private	Yes	Yes	2-skill	Yes
N	Female	Theology	Private	Yes	No	2-skill	Yes

For logistical reasons, this study elicited data from university students. It must be acknowledged that the participants are not representative of all high school students who study for entrance examinations. Important differences from the general population of high school students are that the participants in this study were more successful English learners with high English proficiency and were more motivated to learn English. The findings of this study may not be transferable to less proficient students' English learning and test preparation behavior. The data should therefore be carefully interpreted, taking into account this sampling limitation.

Interviews and Procedure

Individual semi-structured interviews were conducted from July to September 2016. Prior to the interviews, the participants were informed that the interview concerned their English learning behavior while they were preparing for the university entrance examination. They were also asked to bring any materials they used to study. At the start of each interview, the researcher explained the purpose of the research and asked the participant to sign a consent form. Following this, the researcher confirmed the participant's demographic information, including whether they went to a public or a private high school, took the Center Test, went to cram school, and took the examination of other universities. Subsequently, participants were asked to (a) describe how they had prepared for their entrance examinations and (b) explain why they had chosen those methods. In addition to these questions, the participants were asked to elaborate on their English learning activities, show the materials they had used, and answer other questions. The researcher also used the participants' responses to the online questionnaire on learning activities, asking about their reasons for doing or not doing particular activities. The interviews were conducted in

Japanese and audio-recorded. They lasted approximately 52 minutes on average, ranging from 34 minutes (Interviewee A) to 74 minutes (Interviewee D). At the end of the interviews, the participants received a ¥2,000 pre-paid card as an honorarium.

Data Analysis

The interview data were transcribed, and thematic analysis was conducted to identify the factors in student learning behavior during the entrance examination preparation period (Braun & Clarke, 2006). This study focused on what influenced the participants' learning activities and the reasons for adopting these activities. Themes were explored inductively without using pre-existing coding frameworks or the researcher's preconceptions. In this sense, the theoretical concepts and relationships among them were derived from a qualitative analysis of the data (Corbin & Strauss, 2008). The washback literature was used to verify the findings and offer alternative explanations.

The researcher followed the steps of thematic analysis proposed by Braun and Clarke (2006). First, all the transcriptions were read, and initial ideas were noted regarding reasons for students' choices of learning activities. Second, a coding scheme was developed based on the initial ideas in an inductive manner. The initial coding scheme had categories such as entrance examinations, student perception of examinations, school, student strengths in English, effectiveness of study methods, restriction of resources, and examination-independent factors. Each category included two to five subcategories. The entire data set was coded using NVivo 11. Third, the codes were collated into potential themes, and the themes were reviewed to confirm if the data within themes were coherent. For example, three main categories—student perception of examinations, student strengths in English, and effectiveness of study methods—were collated into one main category because they were all related to students' views or perspectives. At this stage, the researcher examined the relationships between the themes and developed a thematic map showing how the themes and subthemes were related. Since the researcher asked in-depth questions regarding the reasons for students' choices of learning activities, it was possible to explore the relationships among some of the categories and subcategories. For example, the following explanation about an examination preparation method reveals that students' poor performance on mock or past tests influenced their views on weaknesses:

The examination of University A contains grammatical error identification questions. I was bad at these questions. ... After I studied the Red Book [a workbook including past examination papers], I noticed I cannot solve error identification questions well. So, as a task for myself, I also started to study for error identification in winter. (Interviewee H)

This relationship was noted as entrance examination → views on weaknesses (where an arrow denotes an influence). Finally, each theme was further refined by merging themes, creating overarching themes, naming them, and generating a clear definition of each. Accordingly, as the data analysis procedure was based on an inductive approach similar to grounded theory, the identified themes were not derived from pre-existing frameworks or theories, but were strongly linked to the data themselves.

Results and Discussion

Factors in Student English Learning

The factors in student English learning during the examination preparation period were explored within the interview data. Table 2 presents four main factors and 14 sub-factors.

Table 2. Factors in Student English Learning

Factor	Sub-factor
1. Examination	1.1 Tasks and items
	1.2 Skill areas measured by examinations
	1.3 Weight of scores or items
	1.4 Commonality between different examinations
2. Student Views	2.1 Views on examinations
	2.2 Views on strengths and weaknesses
	2.3 Views on learning methods
3. School	3.1 Classroom activities
	3.2 Tests at school
	3.3 Advice from teachers or friends
4. Examination-independent Factors	4.1 Restriction on resources
	4.2 Eiken (an English proficiency test)
	4.3 Interests in English and communication
	4.4 Others (Habits, Relaxation, Personality)

Several aspects of the examinations influenced the content and focus of the

participants' English learning activities, with the most considerable aspect being the tasks and items included in an examination. All the participants intensively engaged in practicing mock or past examination papers, especially when the examination date approached. This suggests that students were keen on studying the tasks and items of the examinations they would take in the future. This finding is consistent with that of previous studies conducted in different educational contexts (Allen, 2016; Mickan & Motteram, 2009; Shih, 2007; Stoneman, 2005; Zhan & Wan, 2016).

The participants also seemed to strategically prioritize activities based on the language skills measured by the examinations (i.e., reading, writing, listening, or speaking), score weights, and the commonality between different examinations. They tended to focus primarily on reading because it is tested in almost all the examinations and is heavily weighted. This suggests that score weighting might be one of the examination features that influence students' preparation (Xie, 2015; Zhan & Andrews, 2014; Zhan & Wan, 2016), and students might spend less time preparing for test sections to which few marks are allotted. In contrast to reading, the other skill areas were less attended to or completely ignored, as they were less frequently assessed by the entrance examinations. For example, Interviewee H, who took the Center Test as well as examinations administered by several universities, stated the following reason for not practicing listening and speaking.

Excerpt 1: Since listening was not necessary for other universities [private universities], I only practiced listening in supplementary classes at school and didn't study it by myself. ... I did nothing for speaking because there was no speaking test in the examinations I took.

The oft-mentioned criticism of high-stakes examinations—restricting educational experience (Cheng & Curtis, 2004)—may be applicable to student learning.

This study's educational context was unique because of the commonalities between different examinations. In the Japanese education system, as students usually take several examinations offered by different universities, their learning is influenced by the tasks and items that commonly appear in different examinations. These findings suggest that, while changes to test tasks and score weights are likely to impact student learning, changing a single university entrance examination may only have a small impact.

The participants' English learning activities were also affected by their views on the examinations, including the perceived importance, difficulty, and construct of each examination, which were identified as influential factors by previous studies (Allen, 2016; Green, 2007; Zhan & Andrews, 2014; Zhan & Wan, 2016; Xie & Andrews, 2012). They prioritized the English examinations of their first-choice university and did not study for the TEAP due to the greater perceived importance of the former. In addition, the participants intensively studied reading comprehension because the reading section of the examination was perceived to be difficult and require their attention.

The perceived construct also influenced their learning in that they were engaged in activities to develop the knowledge, skills, and abilities (KSAs) considered necessary for the examination. Grammar and vocabulary were often regarded as the basic KSAs for the whole examination, rather than just being studied for the grammar and vocabulary items. Although previous studies demonstrated that students often focus on learning grammar and vocabulary while preparing for exams, they either did not specifically address the factors underlying this behavior (Mickan & Motteram, 2009; Zhan & Andrews, 2014) or claimed that it was due to the inclusion of grammar and vocabulary items (Damankesh & Babaii, 2015). In the present study, regarding grammar and vocabulary as basic KSAs for examinations led students to focus on these language aspects and engage in various activities to develop them, rather than simply practicing with mock or past examination papers. Excerpt 2 presents Interviewee H's view on vocabulary and fast reading skills.

Excerpt 2: First, because vocabulary is definitely indispensable, though there were vocab quizzes at cram school, of course I studied it by myself as well. Also, the exams of University A and University B² included long and difficult reading passages. Because you need to read a large amount of text very quickly, you need to be able to read quickly and carefully.

Furthermore, learning behavior was influenced by participants' views on their own strengths and weaknesses as well as on learning methods. They spent a great deal of time on the KSAs or examination sections at which they were weak, but did not devote the same amount of effort to those at which they were strong.

² University A is the one that the participants attended, and University B is another prestigious Japanese private university.

While this finding may be at odds with that of Xie and Andrews (2012), it supports the claim that 'supreme self-efficacy in conjunction with perception of easy tasks may discourage them [test-takers]' (p. 63). In contrast, students' perception of their own weaknesses may help them recognize the necessity of studying to overcome these weaknesses (Zhan & Andrews, 2014). Moreover, some participants expressed specific views on certain learning methods, which suggests that students select English learning activities that they consider effective and useful. Excerpt 3 shows Interviewee N's view on the effectiveness of a vocabulary learning technique.

Excerpt 3: I also studied vocabulary items by writing them down. But reading vocabulary aloud was the most effective for me. I was not the type of person who writes vocabulary a lot.

The participants appeared to consider which learning methods would work to both learn English and maximize their scores. Accordingly, the perception of the effectiveness and efficiency may influence students' choice of learning methods (Allen, 2016; Xie, 2015) and be one of the reasons learning methods differ among students.

Participants' English learning activities were also influenced by classroom activities, tests administered at their schools, and advice from their teachers and peers. The participants mentioned a wide variety of learning activities conducted in classrooms, including practicing past examination papers, translating English into Japanese, translating Japanese into English, reading aloud, reading newspapers, shadowing, writing essays, speaking English with teachers, and watching TED talks. While this study focused on students' learning outside the classroom, the participants frequently mentioned English learning activities conducted in class, possibly because students normally study for the entrance examinations in school and are often required to prepare for classes and review relevant content. The participants also followed the advice given by their teachers or peers. Excerpt 4 demonstrates that Interviewee N practiced translation (from English to Japanese) very frequently because it was included in the materials used at school.

Excerpt 4: Translation practice was frequently conducted in test preparation classes at school. The examination of University A was in multiple-choice format, and translation was not very necessary. But translation questions

appeared a lot in the material used in class. [The researcher: Was the material used at school or cram school?] Both at school and cram school.

These findings indicate that, in the Japanese educational context, student learning during the examination preparation period is inseparable from the school environment, so the influence of school environment may be stronger than other contexts, such as preparation for the IELTS (Allen, 2016; Mickan & Motteram, 2009) or university exit tests (Shih, 2007; Zhan & Andrews, 2014).

Finally, participant learning behavior was influenced by factors unrelated to the examinations, including restrictions on resources, an English proficiency test (Eiken), and interest in English and communication. The participants mentioned a lack of resources, including time to study and partners with whom to practice English, which explained why some participants did not spend much time preparing for the TEAP's speaking section. A lack of speaking partner has been reported as a reason why students spend less time preparing for speaking sections of examinations (Allen, 2016; Shih, 2007). Thus, fostering a learning environment for test preparation is important (Shih, 2007), in addition to introducing spoken English tests to induce intended washback.

Activities unrelated to the entrance examinations that participants engaged in included speaking English, reading aloud, watching TV dramas or TED talks, listening to the radio, keeping a diary, reading newspapers, and sending English email messages. Interviewee D stated the following when he was asked the reason for exchanging English e-mails or online messages very frequently.

Excerpt 5: It was because I was interested in foreign countries and wanted to exchange messages. I did it with people I got to know through the Internet or smartphone applications.

The participants' interest in English and communication explains why they engaged in activities unrelated to the entrance examinations (Allen, 2016; Zhan & Andrews, 2014).

In addition to these factors, this study determined that an English proficiency test (Eiken) influenced participants' learning activities. Five participants prepared for the Eiken test, which assesses the four language skills, by studying vocabulary, keeping a diary, and practicing oral interviews. This suggests that a single examination does not determine learning behavior if students need to take other

tests as well.

Relationships Between the Factors

The content of the examinations directly affected participants' learning because they practiced the examination tasks through mock or past examination papers. However, they did not spend their time on all examination items because of mediating factors. The examination tasks and items had a strong impact on participants' perception of task difficulty and their own weaknesses. When the participants identified challenging examination tasks and their relevant weaknesses, they spent much of their time on those particular areas. In contrast, they did not pay much attention to tasks that they perceived to be easy (e.g., the speaking test of the TEAP). These two factors—participants' perception of task difficulty and views on their own weaknesses—were often intertwined, as tasks that the participants felt were difficult were thought to be their own weakness. In other words, the participants realized their weaknesses when they encountered tasks or items that they were not able to complete well in mock or past examination papers. Excerpt 6 shows why Interviewee C used a particular textbook for reading comprehension.

Excerpt 6: I was a bit weak at comprehending long passages and could not read quickly, so I wanted to train myself. By practicing with this textbook and measuring the time, I wanted to be able to read quickly. That's why I bought this book. ... [Researcher: Why did you think that you are weak at reading comprehension?] The score for reading comprehension on a mock examination was a bit bad, so I wanted to take measures.

Xie and Andrews (2012) claim that 'Contingent with perceived task difficulty or easiness, self-efficacy may exercise positive or negative effects on learning effort' (p. 63). Similarly, students' perception of task difficulty may be connected to their self-efficacy as well as the perception of their own weaknesses and strengths, which impact their learning. These perceptions may depend on students' English proficiency levels, as high-proficiency students may regard tasks as easy and thus enjoy relatively high self-efficacy compared to low-proficiency students.

The content of examinations seemed to influence participants' views on the test construct and made them consider the necessary KSAs, as shown in Excerpt 2 (vocabulary and speed reading skills). At the same time, the perceived difficulty

of the test and their own weaknesses made participants consider the KSAs that were necessary for particular tasks. For example, Excerpt 2 suggests that Interviewee H believed that speed reading skills were needed for tackling long and difficult reading passages. Additionally, Excerpt 7 shows why Interviewee I practiced identifying the main points of paragraphs and their relationships.

Excerpt 7: I thought the connection between paragraphs was the key when I thought about the reason for making a mistake on certain questions. The reading passages in the examination of University A were long, and I didn't think I could deal with them easily. So, I thought identifying the gist was the most important for reading. I tried to be able to analyze continuous passages in a simple manner.

This finding suggests that students may analyze examination questions to identify test task demands and appropriate test-taking skills (Xie & Andrews, 2012). Thus, if students misinterpret the KSAs that test developers intend to measure (see Sato & Ikeda, 2015; Zhan & Wan, 2016), an intended washback effect on learning is not likely to be achieved. The findings from this study thus support the idea that 'The assumption that enhanced test validity would discourage cramming and drilling is clearly unwarranted' (Xie, 2015, p. 65). Instead, test-takers' perception of the ability measured by the test (or face validity) needs to be addressed when test developers aim to bring about an intended washback effect (Sato & Ikeda, 2015).

It was found that when the participants identified the relevant KSAs, they engaged in a variety of activities to improve them rather than continuing to practice mock or past examination papers, including memorizing vocabulary or studying grammar. For example, two participants mentioned that familiarity with English was necessary for the examinations. Interviewee N stated the following reason for listening to English songs and making her iPod indication English.

Excerpt 8: Nobody told me to do so. I can't learn natural English if I only read long passages. So, I thought I needed to be familiar with English used in a daily life. Also, according to the information from the Internet, a lack of knowledge of real English or useful English is problematic if you will take the examination of University A. So, I thought that it was better to be exposed to English on a regular basis.

This indicates that the same test could bring about positive washback (engagement in communicative activities) and negative washback (drilling and cramming) depending on how students perceive the required KSAs.

It was also found that while the participants experienced a variety of activities in classrooms, they did not automatically incorporate them into their self-study. Instead, they only employed the methods that they thought useful and effective. Excerpt 9 presents Interviewee L's views on dictation practice used at school.

Excerpt 9: I did dictation exercises for a semester and also had opportunities to do it at school. But I started to think that these exercises are useless for me and a waste of time. So, I quit doing so.

Similarly, the participants seemed to critically consider the effectiveness of teachers' advice and adopt it only when they were convinced by it. These findings suggest that students do not automatically practice the learning activities that are used in classrooms and advised by teachers; their views on particular learning methods mediate their decision whether to utilize those methods. Thus, the relationship between school factors and student view factors may be more complex than previous studies have reported (Allen, 2016; Zhan & Andrews, 2014; Zhan & Wan, 2016).

Their views on learning methods and the necessary KSAs were partly formed based on their teachers' advice on English learning. In particular, the idea that grammar and vocabulary are the basics for the entrance examinations came from their teachers. They often advised participants to study grammar and vocabulary before practicing mock or past examination papers. Interviewee D expressed his views on learning methods that were influenced by his teachers.

Excerpt 10: Teachers normally tell their students to study vocabulary, grammar, and collocation first, and then practice reading comprehension. I think that people who don't like English should follow this way.

This indicates that teachers play a pivotal role in students' learning, as Zhan and Wan (2016) argue, although the importance depends on the educational context. As with classroom activities, however, teachers' advice may not impact learning unless students perceive the effectiveness and usefulness of suggested activities or methods. The participants' views on English learning methods were also formed by their own learning experience. If they felt that a particular learning

activity was useful, they tended to continue engaging in the activity. Excerpt 11 demonstrates why Interviewee M continued to read English passages aloud.

Excerpt 11: I gradually noticed that I am able to understand the passage if I am able to pronounce it at a glance. I read English passages aloud twice or three times a week when I was a second-year student, but I practiced every day after I noticed it.

Positive outcomes from particular learning methods might intensify students' views on the effectiveness of these methods, and lead to their continued use during the examination preparation period.

Summary

Figure 1 illustrates the factors identified by this study and their relationships. First, each main factor—examination, student views, school, and examination-independent factors—affected participants' English learning behavior, including learning activities, focus, and materials (Arrows 1 to 4). Furthermore, some sub-factors were found to affect and mediate other sub-factors in the following ways:

- Examination tasks and items influence students' perceptions of task difficulty, their own weaknesses, and the required KSAs. (Arrow 5)
- Classroom activities as well as advice from teachers and friends are mediated by students' views on learning methods, and influence test preparation behavior. (Arrows 6 and 2)
- Advice from teachers and friends as well as learning experience form views on learning methods. (Arrows 6 and 7)
- Advice from teachers and friends influences student views on the required KSAs (views on examination). (Arrow 8)
- Perceived difficulty (views on examination) and views on strengths and weaknesses are intertwined. (Line 9)

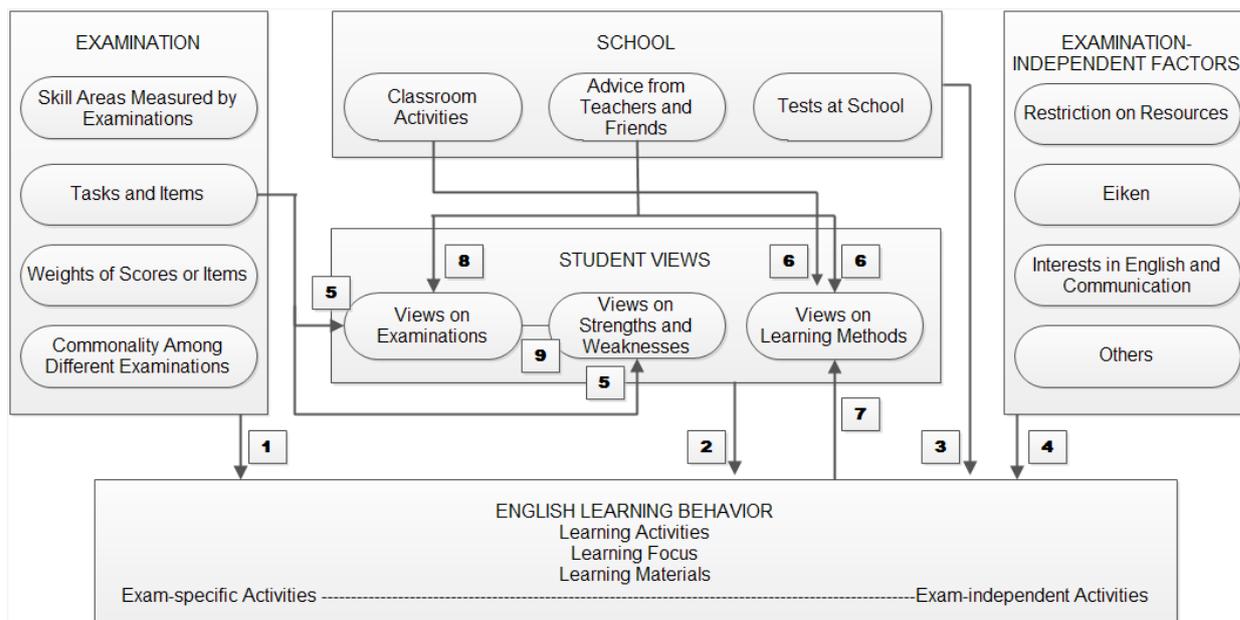


Figure 1. A model of Japanese students' learning behavior for the university entrance examinations.

Conclusions and Implications

This study (a) identified the factors that determined Japanese students' English learning activities while they prepared for their university entrance examinations and (b) constructed a model that explicates the relationships between these factors. This study has demonstrated that students' learning might be influenced not only by the examinations they will take but also by other factors that are complexly intertwined. Examination reform in Japan that introduces new examinations might change students' learning behavior, but it is necessary to address other factors to induce an intended positive washback effect, in order for students to focus more on productive language use and learn the four language skills in a more balanced way.

The findings of this study also make a contribution to the theory of washback on learning behavior. More specifically, in addition to students' learning behavior and the factors that influence it, the study attempted to explain how examinations and other factors influence each other (see Figure 1). The model presented by this study includes various washback factors recognized by previous studies (Allen, 2016; Shih, 2007; Zhan & Andrews, 2014; Zhan & Wan, 2016), thereby taking into account the complexity of the washback mechanism.

The following phenomena, which have not been identified by previous studies, may be added to the existing theories of washback on learning as hypotheses to be empirically tested.

1. When students need to prepare for multiple examinations, their learning is influenced by which test items commonly appear in more than one examination.
2. Students consider grammatical and lexical knowledge basic KSAs knowledge for reading comprehension, and these are studied even when there are no grammar and vocabulary test items.
3. Students recognize their weaknesses based on which test tasks are more difficult for them and analyze what KSAs are required based on their weaknesses. They engage in various activities to tackle their weaknesses, including but not limited to reviewing their past examinations.
4. Classroom activities and advice from teachers influence students' learning behavior but are mediated by their views on learning methods. If students do not recognize the effectiveness of particular learning activities, they do not employ them. The perceived effectiveness of learning methods is formed by their own learning experiences.

There are several limitations for future studies to address. First, as all the participants passed their examination and entered a university, the sample excluded those who prepared for the entrance examinations but did not pass. In particular, this study's participants were likely to be more proficient and interested in English than general Japanese high school students. As Pan (2014) suggests, high-proficiency students may engage in more learning activities and have more positive views on examinations than low-proficiency students. Thus, different factors might be found if research included low-proficiency and less successful English learners. Another limitation is that the study did not elicit all possible factors in students' learning. For example, students' own individualized image of how they wish to and ought to act was not found to be a factor influencing learning behavior in this study as it was in others (such as Zhan & Andrews, 2014). This might be due to the difference in data collection methods.

While some studies investigated students' learning activities, materials, and feelings through their diary entries or study journals (Mickan & Motteram, 2009; Zhan & Andrews, 2014), this study investigated students' learning activities and their reasons for engaging in them through semi-structured interviews. The conclusion drawn from this study should be interpreted with care because of these limitations.

Furthermore, inter-coder reliability was not obtained because data analysis did not involve quantification and inferential statistics. While this may be a limitation of the study as other researchers may devise alternative themes, Spencer, Ritchie, Ormston, O'Connor, and Barnard (2014) claim, 'for qualitative analytic approaches where labelling is done to manage data rather than facilitate enumeration, there is not a "right" or "wrong" way of labelling the data and the aim is not to produce a perfectly consistent coded set' (p. 278). A more serious limitation may be that the study failed to validate qualitative data by triangulation and member checks. It must be acknowledged that these limitations potentially undermine the conclusions drawn from the study.

Further studies are recommended to construct further hypotheses regarding washback on learning and statistically test those hypotheses. With regard to the first recommendation, more exploratory studies should be done to investigate the factors involved in English learning and their relationships in various educational contexts. In particular, longitudinal studies using multiple data collection methods (e.g., interviews, diary studies) can deepen our understanding of the washback of testing and other external factors that influence students' learning behavior. There are still few studies that statistically examine the relationships between factors that contribute to the washback on learning (Xie & Andrews, 2012). The hypotheses derived from this study and future exploratory studies will contribute to statistics-based confirmatory studies, which will in turn provide insights into the complex phenomenon of washback on students' English learning.

Acknowledgements

I am grateful to Prof. Yoshinori Watanabe for his constructive comments on an early draft of the paper. My thanks also go to the editors and the anonymous

reviewers of PLTA for their insightful comments and suggestions.

References

- Allen, D. (2016). Investigating washback to the learner from the IELTS test in the Japanese tertiary context. *Language Testing in Asia*, 6(7), 1-20.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101.
- Cheng, L., & Curtis, A. (2004). Washback or backwash: A review of the impact of testing on teaching and learning. In L. Cheng, Y. Watanabe & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods* (pp. 3-17). Mahwah: Lawrence Erlbaum Associates.
- Corbin, J., & Strauss, A. (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (3rd ed.). Thousand Oaks: Sage Publications.
- Damankesh, M., & Babaii, E. (2015). The washback effect of Iranian high school final examinations on students' test-taking and test-preparation strategies. *Studies in Educational Evaluation*, 45, 62-69.
- Eiken Foundation of Japan. (n.d.-a). *Mondai kousei mihon mondai* [Sections and sample questions of the TEAP]. Retrieved from <http://www.eiken.or.jp/teap/construct/>
- Eiken Foundation of Japan. (n.d.-b). *TEAP*. Retrieved from <http://www.eiken.or.jp/teap/>
- Green, A. (2007). *IELTS washback in context: Preparation for academic writing in higher education*. Cambridge: Cambridge University Press.
- Green, A. (2014). *The Test of English for Academic Purposes (TEAP) impact study: Report 1 - Preliminary questionnaires to Japanese high school students and teachers*. Retrieved from http://www.eiken.or.jp/teap/group/pdf/teap_washback_study.pdf
- In' nami, Y., Koizumi, R., & Nakamura, K. (2016). Factor structure of the Test of English for Academic Purposes (TEAP®) test in relation to the TOEFL iBT® test. *Language Testing in Asia*, 6(3), 1-23.
- Menken, K. (2017). High-stakes tests as de facto language education policies. In E. Shohamy, I. G. Or, & S. May (Eds.), *Language testing and assessment* (3rd ed., pp. 385-396). Cham, Switzerland: Springer International Publishing

AG.

- MEXT. (2017). *Daigaku nyugaku kyotsu tesuto jisshi hoshin* [Implementation policy of the University Entrance Common Test]. Retrieved from http://www.mext.go.jp/component/a_menu/education/micro_detail/_ics/Files/afieldfile/2017/10/24/1397731_001.pdf
- Mickan, P., & Motteram, J. (2009). *The preparation practices of IELTS candidates: Case studies*. Retrieved from Canberra: IELTS Australia.
- Pan, Y.-C. (2014). Learner washback variability in standardized exit tests. *TESL-EJ*, 18(2), 1-30.
- Pan, Y.-C. (2016). Traditional and non-traditional test preparation practices: Learner performance and perspectives. *Electronic Journal of Foreign Language Teaching*, 13(2), 170-183.
- Pan, Y.-C., & Newfields, T. (2012). Tertiary EFL proficiency graduation requirements in Taiwan: A study of washback on learning. *Electronic Journal of Foreign Language Teaching*, 9(1), 108-122.
- Qi, L. (2007). Is testing an efficient agent for pedagogical change? Examining the intended washback of the writing task in a high-stakes English test in China. *Assessment in Education: Principles, Policy & Practice*, 14(1), 51-74.
- Sasaki, M. (2008). The 150-year history of English language assessment in Japanese education. *Language Testing*, 25(1), 63-83.
- Sato, T. (2018). The impact of the Test of English for Academic Purposes (TEAP) on Japanese students' English learning. *JACET Journal*, 62, 89-107.
- Sato, T., & Ikeda, N. (2015). Test-taker perception of what test items measure: A potential impact of face validity on student learning. *Language Testing in Asia*, 5(10), 1-16.
- Shih, C.-M. (2007). A new washback model of students' learning. *The Canadian Modern Language Review*, 64(1), 135-162.
- Spencer, L., Ritchie, J., Ormston, R., O'Connor, W., & Barnard, M. (2014). Analysis: Principles and processes. In J. Ritchie, J. Lewis, C. M. Nicholls, & R. Ormston (Eds.), *Qualitative research practice* (2nd ed., pp. 269-293). London: Sage Publications.
- Stoneman, B. W. H. (2005). *The impact of an exit English test on Hong Kong undergraduates: A study investigating the effects of test status on students' test preparation behaviours*. (Unpublished PhD thesis), The Hong Kong Polytechnic University, Hong Kong.

- Tsagari, D., & Cheng, L. (2017). Washback, impact, and consequences revisited. In E. Shohamy, I. G. Or, & S. May (Eds.), *Language Testing and Assessment* (3rd ed., pp. 359-372). Cham, Switzerland: Springer International Publishing AG.
- Vongpumivitch, V. (2014). Assessing English in east Asia. In A. J. Kunnan (Ed.), *The companion to language assessment* (pp. 1671-1680). Malden, MA: John Wiley & Sons.
- Watanabe, Y. (2004). Teacher factors mediating washback. In L. Cheng, Y. Watanabe & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods* (pp. 129-146). Mahwah, NJ: Lawrence Erlbaum Associates.
- Watanabe, Y. (2013). The National Center Test for University Admissions. *Language Testing*, 30(4), 565-573.
- Wei, W. (2017). A critical review of washback studies: Hypothesis and evidence. In R. Al-Mahrooqi, C. Coombe, F. Al-Maamari, & V. Thakur (Eds.), *Revisiting EFL assessment: Critical perspectives* (pp. 49-67). Cham, Switzerland: Springer International Publishing AG.
- Xie, Q. (2015). Do component weighting and testing method affect time management and approaches to test preparation? A study on the washback mechanism. *System*, 50, 56-68.
- Xie, Q., & Andrews, S. (2012). Do test design and use influence test preparation? Testing a model of washback with structural equation modeling. *Language Testing*, 30(1), 49-70.
- Zhan, Y., & Andrews, S. (2014). Washback effects from a high-stakes examination on out-of-class English learning: Insights from possible self theories. *Assessment in Education: Principles, Policy & Practice*, 21(1), 71-89.
- Zhan, Y., & Wan, Z. H. (2016). Test takers' beliefs and experiences of a high-stakes computer-based English listening and speaking test. *RELC Journal*, 47(3), 363-376.