



Castledown

OPEN ACCESS

Language Education & Assessment

ISSN 2209-3591

<https://www.castledown.com/journals/lea/>

Language Education & Assessment, 2 (2), 96-109 (2019)
<https://doi.org/10.29140/lea.v2n2.150>

The Washback of Listening Tests for Entrance Exams on EFL Instruction in Taiwanese Junior High Schools



HUEI-CHUN TENG ^a
CHUN-WEI FU ^b

^a National Taiwan University of Science & Technology,
TAIWAN
Email: tenghc@mail.ntust.edu.tw

^b Taipei Municipal Hong Dao Junior High School,
TAIWAN
Email: qpourrr@hotmail.com

Abstract

The study aims to explore the washback of incorporating English listening test in the senior high entrance exam (Comprehensive Assessment Program: CAP) on EFL instruction in junior high schools in Taiwan. The participants were thirty English teachers and 298 students from three junior high schools in Taiwan. The instruments consisted of questionnaires and interview guide. The questionnaires were conducted to investigate the participants' perceptions of the influence of CAP English listening test on EFL teachers' instructions, students' learning activities and test preparation. A semi-structured interview was also adopted to probe specific information of participants' perceptions of the CAP listening test, their needs and suggestions. Results showed that teachers and students had positive perceptions of the incorporation of CAP listening assessment. Moreover, the CAP reform generated positive washback effects to the instructors. Most of them were willing to change their teaching approaches, arrange some teaching activities, and provide diverse learning materials as long as class time allows. In addition, the results showed that students' English learning motivation increased due to the CAP listening test. It is also indicated that students' stress and test anxiety for senior high entrance exam may increase. As for the suggestions, both teachers and students indicated that they cared about the time for practicing listening and speaking, listening assessment of different English accents, and better audio equipment. Moreover, because the inequality of educational resources and bimodal distribution of students' EFL proficiency still exist in Taiwan, most of the participants considered that the difficulty level and score weighting of the CAP listening test should not be increased. Implications for enhancing EFL listening instruction in junior high schools were further proposed.

Keywords: washback, listening test, EFL, entrance exam

Copyright: © 2019 Huei-Chun Teng & Chun-Wei Fu. This is an open access article distributed under the terms of the Creative Commons Attribution Non-Commercial 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within this paper.

Introduction

With the trend of globalization, English has become the lingua franca to communicate with people from different countries. For many Taiwanese, the ultimate goal of English learning is to use the language for authentic communication. Yet enhancing Taiwanese students' English proficiency is not only memorizing vocabularies, grammar or sentence patterns but knowing how to communicate to others with English. When it comes to communication, listening skill plays the leading role in that conversation would keep fluently if there is mutual understanding between interlocutors (Capan & Karaca, 2013).

According to Grade 1-9 Curriculum Guidelines in Taiwan (MOE, 2008), five capacity descriptors of English listening are identified for English curriculum in junior high school: (a) to know short English songs, rhymes, phonology and its content; (b) to distinguish different meaning with the intonation in sentences; (c) to understand daily dialogues and easy stories; (d) to identify the contexts and main ideas of dialogues or information; and (e) to know about most of the content of simple videos or short plays. These five listening objectives guide the instruction of English listening in primary and secondary schools. Since listening competence is a fundamental skill in English learning, it is crucial to let learners and instructors lay much emphasis on English listening skill. Therefore, Chang (2006) indicated the necessity and feasibility of incorporating a listening test into the Basic Competence English Test (BCET), the public entrance examinations for junior high graduates in Taiwan from 2001 to 2013. However, the test type and format of BCET was still the same until 2013. In 2014, a new assessment, the Comprehensive Assessment Program (CAP) for Junior High School Students, replaced the original testing system—the Basic Competence Test. The new assessment affects students' choices of entering senior high schools, students' learning outcomes, and teachers' teaching effects (Comprehensive Assessment Program, 2017). In addition, the CAP unprecedentedly incorporated English listening test in the CAP English comprehensive assessment. The design of the CAP English listening assessment is based on the competence indicators derived from the above objectives as well. For these reasons, the incorporation of English listening test in the CAP English comprehensive assessment becomes a central issue among students, teachers, parents and administrators.

Based on the language testing and assessment literature (e.g., Cheng, 1998; Chou, 2017), the reform of high-stakes examinations would influence teachers' instruction and students' learning behavior. Such influence is often referred to washback with the ultimate goal of using the influence of tests on teaching and learning. A widely accepted definition of washback was provided by Hughes (2003), who considered this construct as “the effect of testing on teaching and learning” (p. 1). Specific to an L2 classroom, washback was defined by Brown (2001) as “the benefit that tests offer to learning” (p. 410). Both definitions put emphasis on the tests' influence on teaching and learning. However, Green (2013) pointed out not only is the result of the test vital but probing the result of learning as well as exploring learners' motivation that might encourage participants to react to tests in a certain way and enhance their language development. In addition, according to Alderson and Wall (1993), the test also has the potential to influence specific areas (e.g. teaching methodology, content of teaching, ways of evaluation), positive or negative directions, and other assumed impact.

In light of the scope of the present study, washback will be considered in terms of positive or negative response for learners' and instructors' responses on language learning and teaching through language testing. This study aims to examine the general perceptions and the washback effects of CAP listening test on EFL instruction in Taiwanese junior high schools. Four key research questions are hence proposed in this study:

1. What are EFL teachers' perceptions of the CAP English listening test?

2. In what ways do the incorporation of CAP listening test influence EFL teachers' instruction in terms of test preparation?
3. What are students' perceptions of the CAP English listening test?
4. In what ways does the incorporation of CAP English listening test influence the students' English learning?

Literature Review

Overview of Washback

“Washback” is a neutral term which involves either positive or negative impact on teaching and learning (Alderson & Wall, 1993; Bailey, 1996; Hamp-Lyons, 1997; Hughes, 2003). For instance, if the use of dictation in a proficiency test encourages learners to practice listening skills, then the beneficial effect can be seen as positive washback. Yet, if the listening skill is tested only by obtaining the input without any comprehension and invoking interactive skills, and then the harmful influence on learning can be viewed as negative washback.

Positive washback effects have been regarded as an essential criterion in expanding and evaluating language assessment in current years, and there are a number of widely accepted arguments of positive washback effects. For instance, Wall and Alderson's study (1993) addressed that tests lead learners to focus on particular subjects and skills, and that learners are capable to participate in their learning actively as well as to discover their insufficiencies and problems of learning. In addition, Bailey (1996) pointed out four areas that promote positive washback including the integration of language learning goals, authenticity, learner autonomy and self-assessment, and detailed score reporting.

Messick (1996) further suggests that positive washback and validity can be achieved by improving test tasks and content so that they adequately represent test constructs. To promote the beneficial washback, he also pointed out that assessments have to involve “authenticity” and “directness”. That means designing assessments for the realistic settings or close simulations to the real world. Khodabakhshzadeh, Zardkanloo, and Alipoor's study (2017) found that mock tests can facilitate Iranian IELTS candidates in the test preparation. That is, mock tests have positive washback effects for the candidates. Similarly, Turner (2001) indicated that the standard to rate high-stakes tests is to consider the potential of washback effects at different times of the assessment. A crucial inference of this is that intended progresses towards positive washback which needs to include test design as the ways of achieving washback. Furthermore, a study carried out by Saif (2006) investigates efforts to achieve positive washback can be enriched by executing washback investigations that includes the test development process. She pointed out that the teacher's increasing awareness of the test caused by her participation in the test administration process, interaction with other evaluators, comprehension of the evaluating process, and the capability mechanism of the evaluating instrument. The results of her research are in accordance with Shohamy's (1992) exploration of positive washback. Shohamy (1992) suggested that the involvement of different media in the test development process would increase the beneficial washback on the instruction.

The effects of testing which impede the accomplishment of educational goals on teaching and learning may be referred to as negative washback. Harmful influence of testing has produced potential problems in language education for a long time. Over the past decades, there have been a few studies which examined the damaging washback effects. For instance, negative washback effects arouse anxious feelings among teachers and students (Ferman, 2004; Shohamy et al., 1996). They also restrict the ranges of the curriculum and educational process (Alderson & Hamp-Lyons, 1996; Cheng, 2005; Madaus, 1988; Shohamy, 2001; Qi, 2005; Spolsky, 1995; Wall & Alderson, 1993). Furthermore, Wall

and Alderson (1993) suggested that some important language proficiency skills (e.g. listening or speaking) would be neglected in teachers' instruction in that these skills are not included in the examination and preparing them is inefficient for students. Therefore, negative washback effects also decrease emphasis on skills that require sophisticated thinking or problem-solving (Alderson & Hamp-Lyons, 1996; Bailey, 1996; Lin, 2010; Qi, 2007). Furthermore, further test-preparation classes or tutorials increase (Alderson & Hamp-Lyons, 1996; Bailey, 1996; Cheng, 1998; Shohamy et al., 1996).

There were other instances of negative washback effects related to students' learning and instructors' teaching attitudes. Cheng (1997) indicated that the exams made students work hard to attain high scores but at the same time some of them recognized the high scores could not represent all aspects of their learning outcomes accurately. The high scores gain may be a result of rote memorization of teaching materials, instead of acquiring the real language knowledge. This may fail to construct learners' learning motivation and also affect the interpretations of test scores. In addition, Barnes (2016) signified that the textbooks have the capability to affect teacher behaviour and therefore have a negative washback effect on how teachers teach. The study showed that the mass of class time for teachers' instruction was spent on explaining the tasks or activities in the textbooks and then making the students complete the related exercises of the textbooks during the class.

According to Wall (1997), the impact of testing on teaching and learning includes any of the effects that an assessment might have on individuals, policies, the classroom, the school, the educational system or even the society. A complex interaction of innumerable contextual factors may be embraced in the process of producing washback. Wall and Alderson (1993) noted that tests can be powerful determiners, both positively and negatively of what happens in classrooms" (p.41). Thus, washback appears with the complicated interaction among testing, teaching and learning. As Green (2013) indicated that washback research has implicated that issues educators face are not the influence of tests on teaching and learning, but the interaction among tests, teaching and learning. Messick (1996) also argued "a poor test may be associated with positive effects and a good test with negative effects because of other things that are done or not done in the educational system" (p.242).

As indicated by Cheng, Watanabe and Curtis (2004), the effects of assessment on teaching and learning depend on 5W (who, where, when, why and how) factors. The 5W factors involve "who" truly implements the investigation within a specific education context, and "where" the assessment exists. Besides, "when" refers to the time and length of using such assessment practices, as well as "why" to apply the rationale, and "how" to demonstrate the assessment within the particular context by different participants. The 5W factors infer the significance of disentangling the contextual factors when the researcher attempts to define and explain washback effects.

Related Studies on CAP

The CAP has been implemented in Taiwan since May, 2014. The changes of educational policy and assessment system make a great influence on students, instructors and parents. For instance, Wang (2015) found that students recognized learning English is crucial and the CAP assessment increases their learning motivation. The students also wanted to improve their English communicative abilities and learned more knowledge of cross-cultural customs due to being aware of the importance of learning English. On the other hand, Liu (2016) suggested the differences in scoring procedures led to confusion on students and parents. The BCT applies the percentile rank as a rating basis; however, the CAP employs three standard levels (i.e. excellent, fair and improvement needed) in assessing the results of the assessment. Finally, according to Udn News, Peng (2017) revealed the results of the CAP affect students' choices of their high schools in the future. Moreover, teachers and parents can utilize the results to provide students with the information and advice about choosing their next academic

stage. Due to the change in the entrance exam, the possibility of the washback may be engendered in the incorporation of English listening assessment.

Two washback studies on CAP English listening test are found to date. One is conducted by Chang (2015), finding that the CAP reform had positive effects for most of the teachers. But, there are few teachers indicating that the reform made them focus more on teaching test-taking strategy than listening comprehension. Furthermore, students expressed that they understood more about the importance of listening comprehension and it is necessary to improve their own listening proficiency due to the CAP reform. However, a number of students questioned the fairness of incorporating English listening assessment. They considered the unbalanced distribution of educational resources would cause the unfairness for students in the rural area. The other study (Wang, 2015) on CAP washback investigated junior high students' perceptions towards the changes of their learning activities, their learning motivation and test anxiety, teaching activities in the classroom.

Methodology

Participants

The participants (see Table 1) consisted of 298 students and 30 English teachers from three junior high schools in northern Taiwan. Both groups were from the same three public junior high schools in Keelung City, Taipei City, and New Taipei City. They were recruited to participate in the study based on convenience sampling. Moreover, ten students and five teachers were randomly selected from the voluntary participants who were willing to help with the interviews.

Table 1 *Background Information of Participants*

School	Location	Teacher	Student
A	Taipei City	11	129
B	New Taipei City	11	102
C	Keelung City	8	67
Total		30	298

Instruments

In this study, the questionnaires with a 4-point Likert scale (from 4: strongly agree to 1: strongly disagree) included 29 items for English teachers and 30 items for students. There were three parts in the questionnaires for both teachers (see Table 2) and students (see Table 3). The questionnaires were adapted from Chang (2015) and Her (2006). The Cronbach's coefficient is 0.76 for teacher part, and 0.77 for student part.

Moreover, a semi-structure interview was adopted to triangulate the questionnaire responses. The interview included five questions focusing on the effects of CAP listening test on instruction and learning, as well as interviewees' needs and suggestions.

Procedures

A pilot study was conducted to two teachers and 25 students for obtaining useful information to make sure the survey questions were appropriate and clear. Based on the pilot study responses, one item of teacher part was deleted because of redundancy. After that, the questionnaires were distributed in paper form to the English teachers who volunteered to administer the questionnaires in their class. With the

English teachers' generous assistance, 298 students' questionnaires and 30 teachers' questionnaires were received in two weeks. Finally, semi-structured interviews were scheduled and held individually with five teachers and ten students in their spare time.

Table 2 *Structure of Teachers' Questionnaires*

Part	Category	Item
One	Background information	1-5
Two	Perceptions of the CAP reform	6-15
	Fairness of the English listening assessment	16-21
Three	Instruction and influence	22-29

Table 3 *Structure of Students' Questionnaire*

Structure	Category	Item
One	Background information	1-6
Two	Perceptions of the CAP reform	7-16
	Fairness of the English listening assessment	17-22
Three	Learning activities and test preparation	23-30

Data Analysis

SPSS version 20 was used for analysing the questionnaire responses by English teachers and students. Descriptive statistics were calculated for all the questionnaire items to uncover general patterns emerging from the data. Besides, the interview responses were analysed qualitatively in order to enquire teachers' and students' needs and acquire their suggestions of the CAP reform. First, the audiotape recordings of interview answers were transcribed verbatim. The transcripts were then summarized in the form of matrices (Marshall & Rossman, 2016) to systematically record the categories on the basis of main focusing questions. Content analysis involving coding and categorizing the prime samples in the data was then employed to analyse interview data.

Results & Discussion

This section presents the research findings and discussion with reference to each of the four research questions.

(1) What are EFL teachers' perceptions of the CAP English listening test?

As revealed by the results of the teachers' questionnaire (see Table 4), most teachers indicated that they agreed with the incorporation of the listening assessment into the CAP. Most (93.1%) of the teachers agreed that the CAP listening assessment is in accordance with the students' life experiences. They suggested that the CAP listening assessment should include more different accents. The finding is congruent with Hsu's (2015) study indicating that EFL teachers can help students become more familiar with different accents to raise students' learning interests and promote English listening on TELC. Moreover, a number of previous studies (e.g., Heeren, 2012; Matsuda, 2012) on global English also indicated that the importance of listening comprehension of various English accents.

Table 4 Responses to Teachers' Questionnaires

No.	Item	Mean	SD	Agree	Disagree
6.	The English listening assessment with the incorporation into the CAP.	3.45	0.63	93.1	6.9
7.	The CAP listening assessment should be implemented in students' own junior high schools.	2.07	0.65	17.2	82.8
8.	The schools have to strengthen or purchase more listening equipment.	3.10	0.49	93.1	6.9
9.	The listening assessment should lay more emphasis on different accents.	3.10	0.49	93.1	6.9
10.	The schools should increase the number of times of the listening assessment.	2.93	0.59	79.3	20.7
11.	After the reform, I lay more emphasis on teaching listening.	3.03	0.50	89.7	10.3
12.	The CAP listening assessment is in accordance with the students' life experience.	3.10	0.49	93.1	6.9
13.	The CAP listening assessment assists the students with assessing students' basic communicative ability.	3.00	0.46	89.6	10.3
14.	The scores on the CAP listening assessment can evaluate teachers' teaching efficacy.	2.03	0.63	20.7	79.3
15.	If the students pass the CAP listening assessment, I think it helps them enter into their next academic stages.	2.72	0.59	72.4	27.5
16.	The students who live in cities have benefits of the incorporation of the listening assessment.	3.48	0.51	100	0.0
17.	The students who go to cram schools have benefits of the incorporation of the listening assessment.	2.86	0.74	65.5	34.5
18.	The students' schools with latest and newest equipment have benefits of the listening assessment.	2.79	0.73	62.0	37.9
19.	The students may be influenced by the equipment of the examination rooms and the test surroundings.	3.28	0.53	96.5	3.4
20.	The listening assessment is fair to different groups of students.	2.17	0.60	27.6	72.4
21.	The listening assessment is able to assess the students' learning outcomes and proficiency.	2.52	0.69	62.1	37.9
22.	There is little influence on my teaching ways due to the CAP reform.	2.69	0.47	69.0	31.0
23.	I provide my students with other contents from other textbooks or teaching materials due to the CAP reform.	2.90	0.49	82.8	17.2
24.	I am willing to change my teaching content and ways because of the incorporation of the listening assessment.	2.93	0.49	93.1	13.8
25.	I am willing to arrange some teaching activities to accord with the incorporation of the listening assessment.	3.03	0.33	96.6	3.4
26.	There is just a little influence on my teaching content due to the CAP reform.	2.76	0.64	65.5	34.5

27.	There is just a little influence in my teaching schedule due to the CAP reform.	2.79	0.56	72.4	27.6
28.	I change my ways of evaluation to the students due to the CAP reform.	2.86	0.44	82.7	17.2
29.	I teach the students some listening skills and strategies especially when it closes to the date of the CAP assessment.	3.14	0.52	93.1	6.9

Agree: percentage of SA (strongly agree) & A (agree)

Disagree: percentage of SD (strongly disagree) & D (disagree)

Besides, 89.6% of the teachers agreed that the CAP listening test can assess students' basic communicative ability, but 72.4% of them disagreed the assessment is fair to different groups of students. Because most of them think that the students who live in cities or have more resources to learn English listening have benefits of the incorporation of the listening test. As mentioned by Her (2006), since urban-rural gaps still exist, differences in equipment quality and audio effects concern teachers and students the most. Hsu (2015) also pointed out that the government should put more efforts and funding on schools in the rural area, helping them to catch up with other schools that enjoy richer resources.

With regard to teaching efficacy, 79.3% of the teachers disagreed that the students' scores on the CAP listening assessment can evaluate teachers' teaching efficacy. Chang (2015) suggested that the CAP listening test generate negative washback effects on teachers' teaching efficacy. Fu (2007) also found students' BCT scores increase teachers' pressure and can be used to evaluate teachers' teaching efficacy.

(2) In what ways do the incorporation of CAP listening test influence EFL teachers' instruction in terms of test preparation?

Many (69.0%) of the participating teachers agreed that there is little influence on their teaching ways due to the CAP reform. Moreover, all of the interviewees responded that the reform did not strongly influence their instructions. However, it is still interesting to note that the incorporation of the CAP listening assessment generated positive washback effects for more senior EFL teachers. For example, interviewee T4 who have taught English for more than 20 years indicated that the listening exercises of the textbook and workbook must be completed by the students after the CAP reform. Her statement is presented below:

“The main influence is the listening exercises in the textbook and workbook must be completed by the students. At the BCT period, I focus on grammar teaching and ignore listening because of the insufficient time. Now, instead of ignoring listening, I told my students that the time of practicing listening is precious. Besides, I got used to reading out the texts by myself before. Now, I use e-books to let my students listen to the native speakers' pronunciation.”

Interviewee T1 who have taught English less than 10 years, stated:

“The reform doesn't influence my teaching because I always integrate the four skills while teaching English. However, I notice that some teachers increase the times of using e-books or CD players. These teachers do these because they want the students to get familiar with the foreign accents.”

In addition, more than half of the teachers agreed that there is little influence on their teaching content (65.5%), and teaching schedule (72.4%) due to the CAP reform. Most of the participating teachers always integrate the four skills while teaching English, so the reform did not strongly affect their

instruction. The result also echoes Hsu's (2015) study, finding that the incorporation of the Test for English Listening Competence (TELC) into the General Scholastic Ability Test (GSAT) and the Department Required Test (DRT) will not bring much change on teachers' instruction since her school already emphasizes English listening ability.

In response to some changes in teachers' instruction, most of the teachers are willing to change their teaching approaches or arrange some teaching activities because of the incorporation of the listening assessment into the CAP. Based on the questionnaire results, 86.2% of the participants agreed that they are willing to change their teaching approaches. Nearly all (96.6%) of them agreed that they are willing to arrange some teaching activities. Most (82.8% and 82.7%) of them are not only willing to change their ways of evaluation but also provide their students with other contents from other textbooks or teaching materials due to the CAP reform. Most of the participating teachers expressed that they are willing to do some changes as long as their class time is sufficient. Some of the interviewees indicated the insufficient class time is the major difficulty they encounter while implementing listening exercises in class. For example, Interviewee T3 pointed out that the sufficient class time is crucial:

"It is difficult for me to find out each student's understanding of the listening test items. Because the class time is limited and I have many students, it is hard to monitor students' learning situation. I can only know their learning outcomes by their listening scores."

Another interviewee T4 also stated:

"The difficulty is the time. I don't have enough time to let students doing listening exercises in the classroom."

However, the result is inconsistent with Chou's (2017) research, which found that not all teachers were willing to modify their teaching approaches although nearly all teachers agreed that the implementation of the TELC matched the senior high school English curriculum and teaching objectives. It is probably due to the fact that in the current study, the CAP listening assessment is a compulsory test. The junior high English teachers might encounter the pressure from students' learning outcomes. However, the TELC is an additional test, separate from the GSAT and AST; therefore, some senior high teachers still focused on the training of reading, writing, and Chinese-English translation skills.

In terms of test preparation, the findings revealed that there was no difference after the CAP reform. Some of the interviewees also expressed that they did not change their teaching approaches since they have already taught English with the four skills. Moreover, most of the interviewees indicated that the class time was not enough, so they did not adopt the listening mock tests or handouts for the CAP listening test. They only encouraged their students to subscribe to English magazines or handouts to enhance their English listening proficiency by themselves. The result supports the findings of Hsu (2015), who found that the limited teaching hours made it impossible for English teachers to spend too much time on teaching listening comprehension. Although the results showed that the insufficient class time bothers some teachers, most (93.1%) of the teachers agreed that they still try to teach their students some listening skills and strategies especially when it is close to the date of entrance exam.

(3) What are students' perceptions of the CAP listening test?

As shown in the students' questionnaire responses (see Table 5), over half (60.4%) of the students supported the incorporation of the English listening assessment into the CAP. This is consistent with previous findings (e.g., Chang, 2015; Her, 2006; Hsu, 2015) that it is necessary to incorporate listening

assessment into the English test of entrance exams. Most (75%) of the students also considered that the listening test can assess their basic communicative ability. Moreover, 71.9% of the students agreed that the schools should increase the number of English listening classes. Many (78.8%) of the participants also suggested that the schools' listening assessment should lay more emphasis on different accents. As for the equipment, most (82.7%) of the students reported that the schools have to strengthen or purchase better listening equipment.

Table 5 Responses to Students' Questionnaires

No.	Item	Mean	SD	Agree	Disagree
7.	The English listening assessment with the incorporation into the CAP.	2.55	0.94	61.4	38.5
8.	The CAP listening assessment should be implemented in students' own junior high schools.	2.72	0.87	66.6	33.4
9.	The schools have to strengthen or purchase more listening equipment.	3.11	0.84	82.7	17.4
10.	The schools' listening assessment should lay more emphasis on different accents.	3.06	0.87	78.8	21.1
11.	The schools should increase the number of English listening classes.	2.85	0.89	71.9	28.1
12.	After the reform, I lay more emphasis on learning listening.	2.64	0.93	60.1	39.9
13.	The CAP listening assessment is in accordance with my life experience.	2.49	0.93	53.1	46.9
14.	The CAP listening assessment assists me with assessing my basic communicative ability.	2.89	0.89	75.0	25.0
15.	The scores on the CAP listening assessment can evaluate my learning efficacy.	2.53	0.92	69.1	44.8
16.	If I pass the CAP listening assessment, I think it helps me enter into my next academic stages.	2.77	0.93	68.4	31.6
17.	The students who live in cities have benefits of the incorporation of the listening assessment.	2.61	0.99	57.3	42.7
18.	The students who go to cram schools have benefits of the incorporation of the listening assessment.	2.62	0.96	56.6	43.4
19.	The students' schools with latest and newest equipment have benefits of the listening assessment.	2.85	0.94	67.7	32.3
20.	I may be influenced by the equipment of the examination rooms and the test surroundings.	3.10	0.85	79.2	20.8
21.	The listening assessment is fair to different groups of students.	2.54	0.94	54.5	45.5
22.	The listening assessment is able to assess the students' learning outcomes and proficiency.	2.55	0.99	55.9	44.1
23.	My time at the cram schools may increase due to the CAP reform.	2.58	0.91	55.9	44.1
24.	My high-school-entrance stress may increase due to the CAP reform.	2.91	0.91	70.8	29.2
25.	My learning motivation increases due to the CAP reform.	2.51	0.87	53.5	46.5
26.	The CAP reform is no help with the students' English learning; instead, it increases my test anxiety.	2.65	0.96	55.9	44.1
27.	There is just a little influence on my English learning due to the CAP reform.	2.60	0.87	54.5	45.5
28.	Preparing the listening assessment is different	2.96	0.82	77.4	22.6

	from preparing traditional English assessment, so the students have to change the ways to learn English.				
29.	It is easier to prepare a listening assessment than a reading assessment.	2.66	1.00	58.0	42.0
30.	I do some listening exercises of other textbooks and additional learning materials.	2.46	1.00	51.7	48.2

Agree: percentage of SA (strongly agree) & A (agree)

Disagree: percentage of SD (strongly disagree) & D (disagree)

(4) In what ways does the incorporation of CAP listening test influence the students' English learning?

The study results show that most of the changes for students' English learning are found to be positive. For instance, 60.1% of the participants agreed that they laid more emphasis on learning listening after the CAP reform. In addition, as found in the previous studies (e.g., Chang, 2015; Her, 2006), 51.4% of the participants expressed that there were some changes in their learning activities due to the CAP reform, such as doing some listening exercises of other textbooks and additional learning materials. Some interviewees also indicated that they have enhanced their listening proficiency by listening to English songs or communicating with foreigners. They would also use some listening strategies to prepare for the listening test. For example, interviewee S5 stated:

"My learning strategy is communicating with foreign people when I play the online games. I can practice English listening and speaking while talking to them."

Interviewee S6 stated:

"Listening to English songs is my learning strategy. I try to understand the meaning of lyrics."

The finding is consistent with Hsu's (2015) finding that students adopted new learning methods (e.g., listening to the music or watching movies) after the implementation of TELC.

Although the study has yielded findings that have contributions to enlarging the knowledge of washback effects of the CAP listening test, there are still some limitations in terms of experimental design. First, sampling is one of the limitations. The participants only covered the three cities, i.e. Taipei city, New Taipei city and Keelung city. Because the small number of the participants, the study results cannot present the general perceptions of all the English teachers and students in Taiwan. Hence, the findings cannot indicate the urban-rural gaps among northern, central, southern, and eastern Taiwan. Besides, the interviewees were randomly chosen to take part in the interviews, so it is suggested that interviewees may be recruited based on their questionnaires' results in the future research. Therefore, the interview responses may be diverse according to the participants who have different opinions on the questionnaires.

Second, as for the data collection, although the study adopted both quantitative and qualitative methods to triangulate the participants' perceptions about the CAP reform, due to time constraints, the researcher did not implement the classroom observation in the current study. In the study by Cheng (1997), she adopted classroom observations in sampled schools in Hong Kong to discuss the observed washback effect of the Hong Kong Certificate of Education Examination (HKCEE) in English. For the purpose of verifying the changes and frequency with which types of behaviors occurred in the classroom after the CAP reform, classroom observation can be taken into considerations in the further research.

Finally, the research only focused on exploring students' perceptions about the CAP reform, the influence of learning, and their needs and suggestions, but it did not indicate the different background students' views. In the current study, the questionnaire included students' background information. However, the data analysis of the questionnaire may not be able to present a whole picture of different background students' perceptions toward the CAP reform. Therefore, for future studies, researchers can analyze the individual differences among students with different backgrounds to obtain diverse opinions so that the findings could be more objective.

Conclusion

To sum up, the participating teachers and students in the study mostly give positive perceptions of the incorporation of CAP listening assessment. Both of them think that the listening test can assess students' basic EFL communicative ability and will help their future learning. The CAP listening test also generates positive washback effects to EFL teachers. If the class time allows, most of the teachers are willing to change their teaching approaches, conduct more listening activities, and provide diverse listening materials.

Based on the interview responses, there are some suggestions of the CAP listening test for teachers, schools, and Ministry of Education (MOE). For teachers, the students hoped that EFL teachers can have more interaction in English with them because they will have more chances to practice listening comprehension. Echoing the previous studies (e.g., Chang, 2015; Huang, 2004), it is also suggested that EFL teachers provide students with more diverse listening materials and increase the frequency of listening assessment. The suggestions to schools and MOE are similar to Her's (2006) research indicating that differences in equipment quality, noise interference, and audio effects due to seat arrangement concern teachers and students the most.

The present study is expected to contribute to the understanding of the washback of listening assessment in entrance exams. It can provide pedagogical implication for English teachers to assist the junior high students in facilitating their EFL listening comprehension. The study results can also enable the authorities to understand students' and instructors' needs on the CAP listening test so that they can provide the more suitable and better curriculum plans and learning environment.

References

- Alderson, J. C., & Hamp-Lyons, L. (1996). TOEFL preparation courses: A study of washback. *Language Testing, 13*, 280-297.
- Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics, 14*(2), 115-129.
- Bailey, K. M. (1996). Working for washback: A review of the washback concept in language testing. *Language Testing, 13*, 257-279.
- Barnes, M. (2016). The washback of the TOEFL iBT in Vietnam. *Australian Journal of Teacher Education, 41*(7), 158-174.
- Brown, J. D. (2001). *Using surveys in language programs*. Cambridge University Press.
- Capan, S. A., & Karaca, M. (2013). A comparative study of listening anxiety and reading anxiety. *Procedia-Social and Behavioral Sciences, 70*, 1360-1373.
- Chang, C. M. (2015). *An investigation of incorporating English listening in the comprehensive assessment program for junior high schools* (Unpublished master's thesis). Kang Ning University, Taipei, Taiwan.
- Chang, W. (2006). English language education in Taiwan: A comprehensive survey. *Educational Resources and Research, 69*, 129-144.
- Cheng, L. (1997). How does washback influence teaching? Implications for Hong Kong. *Language*

- and Education, 11(1), 38-54.
- Cheng, L. (1998). Impact of a public English examination change on students' perceptions and attitudes toward their English learning. *Studies in Educational Evaluation*, 24(3), 279-301.
- Cheng, L., Watanabe, Y., & Curtis, A. (2004). *Washback in language testing: Research contexts and methods*. Mahwah, NJ: Erlbaum.
- Cheng, L. (2005). *Changing language teaching through language testing: A washback study*. Cambridge: Cambridge University Press.
- Chou, M. H. (2017). Impacts of the test of English listening comprehension (TELC) on teachers and teaching in Taiwan. *Asian-Pacific Journal of Second and Foreign Language Education*, 2, 1-14. <https://www.doi.org/10.1186/s40862-017-0028-9>
- Comprehensive Assessment Program. (2017). *The background of the comprehensive assessment program*. Retrieved from <https://cap.nace.edu.tw/background.html>
- Ferman, I. (2004). The washback of an EFL national oral matriculation test to teaching and learning. In L. Cheng & Y. Watanabe (with A. Curtis) (Eds.), *Washback in Language testing: Research Contexts and Methods* (pp. 191-210). Mahwah, NJ: Erlbaum.
- Fu, C. Y. (2007). *The study of the attitude and improvements on the basic competency test for junior high school students for students, parents and educators* (Unpublished master's thesis). National Sun Yat-sen University, Kaohsiung, Taiwan.
- Green, A. (2013). Washback in language assessment. *IJES*, 13(2), 39-51.
- Hamp-Lyons, L. (1997). Washback, impact, and validity: Ethical concerns. *Language Testing*, 14, 295-303
- Her, T. C. (2006). *Incorporating a listening component into the basic competence English test: A questionnaire survey* (Unpublished master's thesis). National Taiwan Normal University, Taipei, Taiwan.
- Heeren, K. D. (2012). *The effect of an unfamiliar accent on Taiwanese students' listening comprehension* (Unpublished doctor's thesis). National Kaohsiung Normal University, Kaohsiung, Taiwan.
- Hsu, Y. T. (2015). *An exploration of Taiwanese senior high school students' and teachers' perceptions of the test of English listening comprehension (TELC)* (Unpublished master's thesis). National Tsing Hua University, Hsinchu, Taiwan.
- Huang, S. C. (2004). *Washback effects of the basic competence English test on EFL teaching in junior high school in Taiwan* (Unpublished master's thesis). National Yunlin University of Science and Technology, Yunlin, Taiwan.
- Hughes, A. (2003). *Testing for language teachers*. Cambridge: Cambridge University Press.
- Khodabakhshzadeh, H., Zardkanloo, R. & Alipoor, I. (2017). The effect of mock tests on Iranian EFL learners' test scores. *International Journal of Education & Literacy Studies*, 5(3), 47-51.
- Lin, Y. Z. (2010). *A study on the washback effect of the basic competence English test* (Unpublished master's thesis). National Chengchi University, Taipei, Taiwan.
- Liu, M. Z. (2016). *Analysis and comparison of the BCT and CAP English test trends from 2012 to 2015* (Unpublished master's thesis). Asia University, Taichung, Taiwan.
- Madaus, G. F. (1988). The influence of testing on the curriculum. In L. N. Tanner (Ed.), *Critical Issues in Curriculum: Eighty-seventh Yearbook of the National Society for the Study of Education, Part I* (pp. 83-121). Chicago: University of Chicago Press.
- Marshall, C., & Rossman, G. B. (2016). *Designing qualitative research*. Sage Publications.
- Matsuda, A. (2012). Incorporating world Englisher in teaching English as an international language. *TESOL Quarterly*, 37(4), 719-729.
- Messick, S. (1996). Validity and washback in language testing. *Language Testing*, 13, 241-256.
- MOE. (2008). *General guidelines of grade 1-9 curriculum of elementary and junior high school education*. Taipei, Taiwan.
- Peng, J. S. (2017, June 28). *The new era of comprehensive assessment program: The way of surviving*

- for students with low and intermediate proficiency levels. UDN News Net. Retrieved from <https://udn.com/news/story/6887/2551742>
- Qi, L. (2005). Stakeholders' conflicting aims undermine the washback function of high-stakes test. *Language Testing*, 22(2), 142-173.
- 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*, 14, 51-74.
- Saif, S. (2006). Aiming for positive washback: A case study of international teaching assistants. *Language Testing*, 23(1), 1-34.
- Shohamy, E. (1992). Beyond proficiency testing: A diagnostic feedback testing model for assessing foreign language learning. *The Modern Language Journal*, 76, 513-521.
- Shohamy, E., & Donitsa-Schmidt, S., & Ferman, I. (1996). Test impact revisited: Washback effect over time. *Language Testing*, 13(3), 298-317.
- Shohamy, E. (2001). *The power of tests: A critical perspective on the uses of language tests*. New York: Pearson.
- Spolsky, B. (1995). The examination- classroom backwash cycle: Some historical cases. In D. Nunan, R. Berry & V. Berry (Eds.), *Bringing about Change in Language Education* (pp. 55-66). Hong Kong: University of Hong Kong.
- Turner, C. E. (2001). The need for impact studies of L2 performance testing and rating: identifying areas of potential consequences at all levels of the testing cycle. In C. Elder, Brown, A., Grove, E., Hill, K., Iwashita, N., Lumley, T., McNamara, T., & O'Loughlin, K. (Eds.), *Experimenting with Uncertainty: Essays in Honor of Alan Davies* (pp. 138-149). Cambridge University Press.
- Wall, D., & Alderson, J. C. (1993). Examining washback: The Sri Lankan impact study. *Language Testing*, 10, 41-69.
- Wall, D. (1997). Impact and washback in language testing. In C. Clapham, & D. Corson (Eds.), *Encyclopedia of Language in Education: Vol. 7. Language Testing and Assessment* (pp. 291-302). Dordrecht, the Netherlands: Kluwer Academic.
- Wang, S. M. (2015). *The washback effect of comprehensive assessment program on junior high school students* (Unpublished master's thesis). Ming Chuan University, Taoyuan, Taiwan.

Author biodata

Huei-Chun Teng is a professor in the Department of Applied Foreign Languages at National Taiwan University of Science and Technology, Taiwan. She received her Ph.D. degree from University of Minnesota, USA. Her research interests include listening research, communication study, language assessment, and learning strategy.

Chun-Wei Fu is an English teacher at Taipei Municipal Hong Dao Junior High School, Taiwan. She received her M.A. degree from National Taiwan University of Science and Technology.