

Taiwan University Students' Perceptions of Summative and Formative Classroom Assessment in English Courses

Yi-Ching Pan*

National Pingtung University

Abstract

Classroom assessment is an essential component of teaching and learning. Still, the literature on language testing often highlights teachers' perceptions of designing classroom assessments with little concern about students' perceived realities. This study thus explores Taiwan university students' perceptions of summative and formative classroom assessment tasks in their English courses. To address this issue, 107 first-year undergraduates at one Taiwan university were recruited to complete one summative written exam and two formative assessments, after which they filled in a questionnaire consisting of six subscales: congruence with planned learning, authenticity, student consultation, transparency, diversity, and washback effects. We discover four major findings: 1) students were in favor of the summative assessment due to its congruence with planned learning and transparency; most students admitted that they learn more from preparing for the summative assessment; 2) students viewed the cooperative group assessment positively, because of its diversity; 3) preparation for summative assessments elicited a greater degree of test-oriented learning for respective skills, whereas formative assessments enhanced students' motivation to learn English for productive skills; and 4) students believed that an appropriate combination of summative and formative assessment tasks benefits their learning. The findings provide further pedagogical implications.

Key words: consequences of test use, washback, classroom assessment, summative assessment, formative assessment

* yichingpan@yahoo.com.tw

Introduction

Since the literature has increasingly scrutinized the consequences of tests, many studies have investigated the washback on large-scale standardized language tests to determine the effects brought about by educational policies in different countries. For example, studies by Hayes & Read (2003), Ferman (2004), Gebril & Brown (2014), Cheng (2005), Qi (2005, 2007), Xie (2015), Zhan & Andrews (2014), and Smyth & Banks (2012) explored how high-stakes entrance examinations affect teaching and learning in New Zealand, Israel, Egypt, Hong Kong, China, and Ireland, respectively. Wall & Horak (2006, 2008, 2011) and Green (2007) analyzed the impacts of TOEFL (Test of English as a Foreign Language) and IELTS (The International English Language Testing System) on teaching in European ESL (English as a second language) classes Shih (2007), Tsai & Tsou (2009), Gan (2009), Pan, (2014, 2018), and Xie (2013) investigated the positive and negative effects brought about by English exit requirements using standardized language tests in Hong Kong and Taiwan higher-education institutes. These aforementioned traditional large-scale standardized tests, also known as assessment of learning (Earl, 2003), usually focus on learning outcomes, yet hardly address learning processes. Different types of assessments, such as assessment for learning and assessment as learning (Earl, 2005), have therefore been subjected to heightened levels of consideration. In addition to the raised awareness of multiple methods of assessments, many countries have proposed the utilization of standardized tests due to their fairness and reliability, following pressure to seek educational accountability and quality control (Black, 2000). The current trend is to adopt different types of classroom assessments as a fair and reliable means of monitoring students' learning processes and evaluating their classroom performance.

Assessments play an essential role in teaching and learning because they are crucial for informing teachers about their work, while at the same time, if the assessment is improved, then the resultant learning can be enhanced. A few studies exist that focus on the washback or impact of high-stakes testing of teaching and learning, especially from teachers' perceptions (Cheng, 2008, 2014; Spratt, 2005; Wall, 2000). However, very little research has explored how classroom assessments influence learning from students' viewpoints. To bridge this gap in the literature, we investigate Taiwan university students' perceptions of summative and formative classroom assessment tasks in their English classes and how these classroom assessments affect their learning.

Context of the study

Students in Taiwan are required to learn English from the time they are in elementary school so that by the time they attend a university, they will have learned English for at least eight years. Many students are not very motivated to learn English, probably because Taiwan's education system is test-oriented and teacher-centered. At the university level, all first-year students are required to take two-three hours of English classes every week. Moreover, the majority of these students have stated that they are taking English classes, because they want to earn credits for graduation, with around 30% of these students passing the CEFR A2 level and 50% of them not passing it. In other words, quite a few

students had insufficient English proficiency, even though they had learned English for more than eight years.

Literature review

This section reviews (1) the historical perspectives of washback, focusing on the contributing factors for various degrees of test effects and (2) the possible dimensions for understanding test-takers' perceptions of classroom assessments. The review on washback effects and assessment tasks serves as a theoretical basis for the purpose of this study to explore both students' perceptions of assessment tasks and the influence of the assessments on learning.

The historical perspectives of washback focusing on the contributing factors for various degrees of test effects

Hughes' (1993, 2003) trichotomy of the backwash model describes test effects in terms of "participants", "process(es)", and "product(s)". This model explains how participants interpreting and reacting toward a test affect how and what their responses to it will be, indicating that the quality of a test is essential to predict the degree of washback test effects.

Alderson and Wall (1993) propose 15 washback hypotheses and illustrate various possible effects brought about by tests primarily on teaching and learning, ranging in detail from the most basic to quite specific effects, as listed in Appendix 1. Alderson and Wall's hypotheses, like Hughes' model, address what washback effects might look like (i.e. the consequences) more than they focus on what factors other than a test (i.e. the mechanisms) lead to these effects. Nevertheless, several of the hypotheses imply that, in addition to a test, there might be some other factors that elicit more effects from some learners and teachers than they did for others. Not only does the quality of a test affect teaching and learning, but how a test is used under different situations also affects the strength of the washback, as in Hypothesis 13: "Tests that do not have important consequences will have no washback" (ibid.).

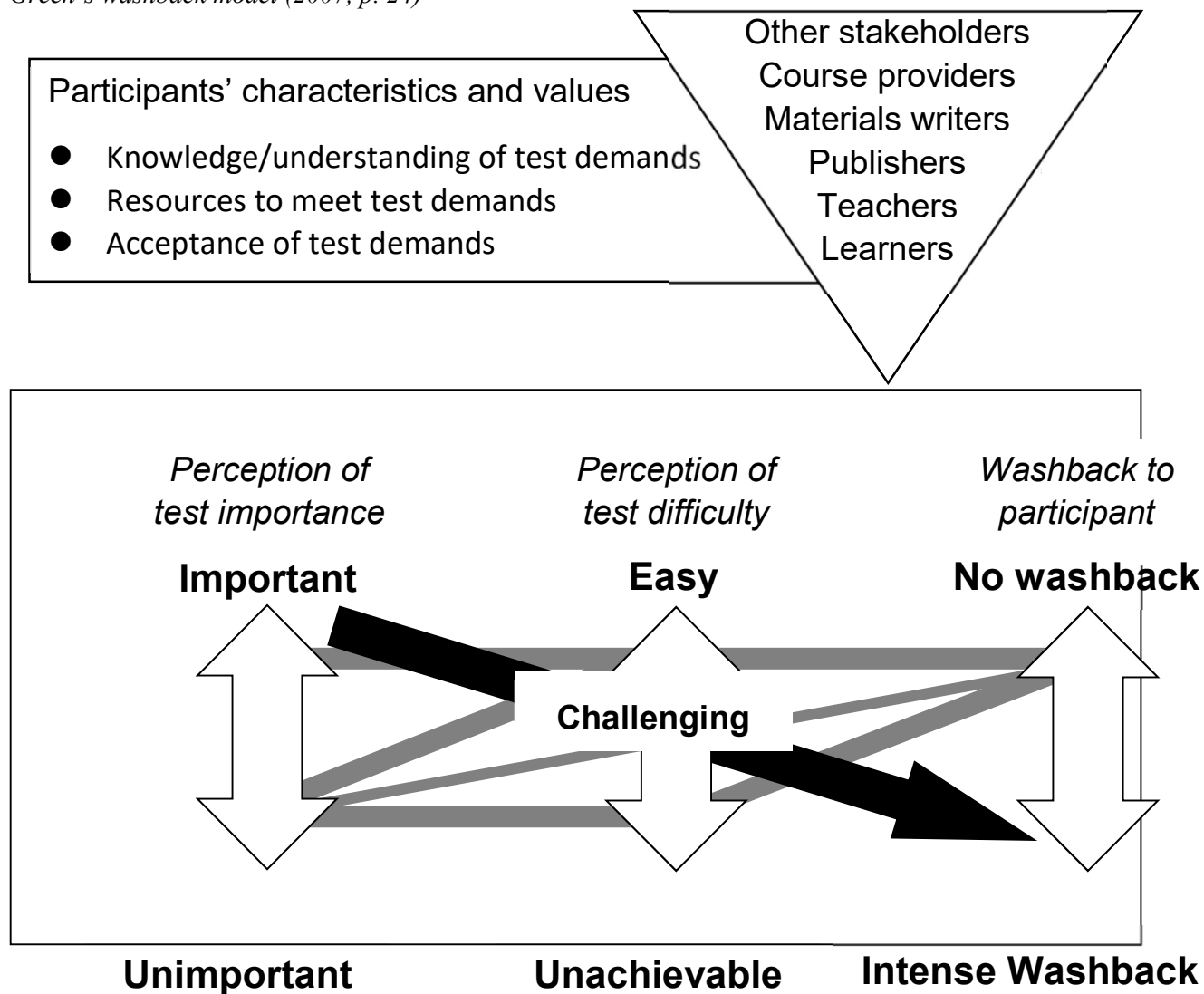
Green (2007) presents the concept of washback variability (see Figure 1), in which participants' perceptions of test stakes, test quality, and test difficulty tend to vary from person to person and therefore lead to differences in the washback effects experienced by individuals. In Green's opinion, students' perceptions, attitudes, and reactions toward the following seven factors may result in various degrees of washback intensity: (a) test demands (content, format, and complexity), (b) purpose of the test, (c) stakes of the test, (d) difficulty of the test, (e) test preparation resources, (f) teaching methods, and (g) learning content.

In review of the aforementioned studies, students' perceptions of classroom assessment can be explored from washback perspectives to understand how the assessment influences students in various ways. Using the above three washback models as the theoretical base for the questionnaire in

this study, Table 2 lists 11 items to investigate to what extent various assessment-related factors influence students' learning.

Figure 1

Green's washback model (2007, p. 24)



The possible dimensions for understanding test-takers' perceptions of classroom assessments

Dorman and Knightly (2006) develop an instrument called Perceptions of Assessment Tasks Inventory (PATI) to observe students' perceptions of assessment tasks under five categories to understand the qualities of assessments, which served as the theoretical basis for developing the 12 items in the student questionnaire (see Table 2). These five categories are described below:

These five categories are congruence with planned learning, authenticity, student consultation in the assessment process, transparency about the purposes and forms of the assessment, and diversity.

Congruence with planned learning denotes that students perceive the assessment tasks as being aligned with their learning objectives and activities. Authenticity means that the assessment tasks are related to students' daily life. Student consultation implies that students are involved and consulted in the assessment process. Transparency signifies that students are well informed about the purposes and forms of the assessment. Diversity is when students perceive that they can finish the assessment tasks at their speed.

The overall assessment can be categorized into two parts. One is summative assessment whose goal is to *evaluate student learning* at the end of the semester by comparing it against some standard or benchmark. Examples include traditional paper-and-pencil tests such as mid-term and final exams and proficiency tests.

Summative assessment usually lacks feedback or any suggestions to improve performance (Brown, 2004). The other is formative assessment, whose goal is to *monitor student knowledge and understanding in order* to provide ongoing and useful feedback that can be used both by instructors to improve their teaching and by students to improve their learning. Much of the classroom assessments such as oral discussion, group/pair work, and completing a portfolio can be formative because students form their knowledge by analyzing and internalizing teachers' comments (Brown, 2004). However, it seems that formative assessment has not always been the focus of attention in ESL/EFL studies, especially in a test-oriented educational environment. If the students are not very proficient in English, as mentioned in the context of study, and they have become accustomed to summative written assessments, then a question arises: Would they hold more favorable opinions toward summative assessments than formative assessments, or the other way around? We shall investigate this to help fill the gap in the field of language testing.

Research questions

Based on the literature review, this study thus explores two research questions.

1. What are the differences between students' perceptions of summative and formative assessments?
2. What are the differences between students' perceptions of summative and formative assessments based on their levels of proficiency?

Method

Subjects

One hundred and seven first-year students at one university were recruited for this study, of which 29 were male, and 78 were female. At the time of the study, they were taking a required 3-hour English class every week. Based on their mid-term exam scores, they were split into two groups: 54 students whose scores were in the top 50% were in the high-proficiency group, whereas the others (53) were in the low proficiency group.

Instruments

The instruments utilized include three assessments, one questionnaire, students' learning reflections, and semi-structured interviews. This study reports mainly quantitative data due to length limitations.

Three Classroom Assessments

The three assessments, including one summative and two formative assessments, were all designed

based on lecture material from the English reading class and were completed within a timeframe of one semester. The summative assessments consisted of two traditional multiple-choice listening and reading questions. As for the formative oral presentations, students were given a choice to select one question out of two and make a two-minute presentation. For the formative group audio PowerPoint presentation, students were asked to form groups of two to three people, create an audio PowerPoint, and write reflections based on what they had learned from this project. The students were also told in detail about the guidelines for the assessments, the purposes of completing these two formative assessments, when they should finish, and how they would be graded.

Questionnaire

The questionnaire consists of two parts. The first part adopts Dorman and Knightley's (2006) Perception Inventory and includes five sections: Planned Learning, Authenticity, Student Consultation, Transparency, and Diversity. The second part is based on a review of washback studies, such as Hughes' washback model, 15 Washback Hypotheses, and Green's concept of washback variability. Table 1 describes the 23 items listed in the student questionnaire. The Cronbach Alpha reliability for the 23 items on a 5-point Likert scale is 0.83. The 24th item asked which assessment did the students favor the most.

Table 1

Items on the Classroom Assessment Questionnaire

Part I	Items
Congruence with Planned Learning	1. How I am assessed and what I am assessed upon relate to what is done in the English class. 2. What is taught in the English class can help me to prepare for the assessments.
Authenticity	1. I have learned useful knowledge from the assessment. 2. I find the English assessment tasks to be relevant to the real world. 3. I find the English assessment tasks to be relevant to the future workplace.
Student Consultation	1. I am aware of how I will be assessed and what I will be assessed upon in the English class. 2. My teacher has explained to me the purpose of each assessment.

Transparency	<ol style="list-style-type: none"> 1. I understand how the English assessment tasks will be marked. 2. I am told in advance when I will be assessed. 3. I know what is needed to successfully accomplish the English assessment and get high marks.
Diversity	<ol style="list-style-type: none"> 1. I am given a choice of assessment tasks that suit my ability. 2. I am allowed to complete assessment tasks at my own speed.
Part II	Items
Washback (a) Test demands (content, format, and complexity) (b) Purpose of the test (c) Stakes of the test (d) Test difficulty (e) Test preparation resources (f) Teaching methods (g) Learning content	<ol style="list-style-type: none"> 1. The assessment is important to me. 2. The assessment is easy to me. 3. The assessment can measure my English ability. 4. Preparing for the assessment can enhance my motivation to learn English. 5. Preparing for the assessment has enhanced my four English skills (listening, reading, writing, and speaking). 6. Preparing for the assessment has enhanced my English proficiency levels. 7. I have experienced a variety of learning methods and activities while preparing for the assessment. 8. Preparing for the assessment is a meaningful and worthwhile experience. 9. I have spent a lot of time preparing for the assessment. 10. I am under pressure while preparing for/taking the assessment. 11. I am in favor of this assessment.
Part III	Items
Favorite classroom assessment	<ol style="list-style-type: none"> 1. Traditional pen-and-pencil test. 2. Individual presentation. 3. Group project. 4. A combination of the above three.

Data collection procedure

The study gave clear instructions over three assessments, including 1) two traditional multiple-choice listening and reading examinations (summative assessment in Tables 2 & 3), 2) a two-minute presentation (formative assessment 1 in Tables 2 & 3), and 3) group audio PowerPoint presentation (formative assessment 2 in Tables 2 & 3), as to their purpose, test tasks, test format, test time, marking criteria, and preparation. For the two formative assessments, the students were given oral feedback and written feedback during the process of their presentation.

After the students completed the three assessments, the classroom assessment questionnaires were distributed in class. Each student was rewarded with a free glass of fruit juice to thank them on their effort in answering the questionnaires.

Data analysis

IBM SPSS Statistics 22 was used to analyze the quantitative data. We utilized descriptive statistics to calculate the frequency distributions, means, and standard deviations of the questionnaire items and test scores. We also used inferential statistics (e.g. independent t-test, one-way ANOVA) to check for a statistical significance level of .05 ($p < 0.05$). The effect size is also reported to help readers

understand the magnitude of the significant differences found (Larson-Hall 2012; Larson-Hall & Plonsky, 2015), where .10, .30, and .50 are the benchmarks for the small, medium, and large effects in Cohen's (1988) effect size interpretation system.

Results & discussion

1. What are the differences between students' perceptions of summative and formative assessments?

According to one-way ANOVA analysis, we present the main findings that illustrate the statistically significant differences and effect sizes between students' perceptions of these three types of assessments, as shown in Tables 2 & 3.

In terms of consultation transparency, students are in favor of summative assessment more than the other types to a small degree. The students have a slightly greater understanding of the purpose and the marking scale of the summative assessment, possibly due to the fact that they have primarily taken summative written exams during the course of their studies. In terms of transparency, students have a greater understanding of how to prepare for the formative oral assessment than the other two assessments between a small and medium degree. This finding is likely, because there is only one oral question for the formative oral assessment, and students can devote their full attention to preparing for that single question. However, there is a lot more material covered in the summative and formative cooperative assessments, and so what students have prepared for may not even appear on the test, and accordingly, students have spent much more time completing these two kinds of assessments. In terms of diversity and authenticity, students favor formative assessments more than summative assessments to a medium degree. In their opinion, they can choose the assessment tasks that suit their ability at their speed. In addition, students contend that they can learn useful knowledge relevant to the real world from either the formative oral assessment or formative cooperative assessment.

In terms of washback effects, students favor summative assessments to a medium degree, because they are fair, can measure students' abilities, and help them to improve their vocabulary, listening, and reading skills. However, students favor formative assessments because they think

preparing for these two types of assessments are a meaningful and worthwhile learning experience. They stated that they had improved their writing and oral skills from formative assessments. Interestingly, although students stated that formative oral assessments increased the amount of pressure they experienced, they prefer this type of assessment than the summative assessment to a medium degree.

When asked which type of classroom assessments they preferred to take, 62% of the participants stated they liked the combination of the traditional pen-and-pencil test, individual presentation, and group project. Another 10-15% of the participants preferred each of the three classroom assessments respectively.

Based on the aforementioned findings, the preparation for summative assessments elicited a greater degree of test-oriented learning, whereas formative assessments enhanced students' motivation to learn English. Students have favorable opinions of different assessments, because of the positive effects they brought about on students' learning. Therefore, it is believed that an appropriate combination of summative and formative assessment tasks is beneficial for their learning.

Table 2
Students' perceptions of the three types of assessments (1)

Category	Questionnaire items	Assessment type (I)	Assessment type (J)	Mean (I-J)	Standard error	Sig.	r
Student consultation	My teacher has explained to me the purpose of each assessment.	Summative	Formative 2	.233	.083	.015	.23
		Formative 1	Formative 2	.289	.083	.002	.31
Transparency	1 I understand how the English assessment tasks will be marked.	Summative	Formative 1	.308	.114	.020	.22
		Summative	Formative 2	.414	.114	.001	.31
	2 I know what is needed to successfully accomplish the English assessment and get high marks.	Formative 1	Summative	.346	.106	.003	.28
		Formative 1	Formative 2	.368	.106	.002	.30
Diversity	1 I am given a choice of assessment tasks that suit my ability.	Formative 2	Summative	.506	.116	.000	.37
	2 I am allowed to complete assessment tasks at my own speed.	Formative 2	Summative	.263	.108	.041	.20
Authenticity	1 I have learned useful knowledge from the assessment.	Formative 2	Summative	.291	.106	.018	.23
		Formative 2	Formative 1	.300	.106	.014	.24
	2 I find the English assessment tasks to be relevant to the real world.	Formative 1	Summative	.346	.121	.013	.24

Table 3

Table 3: Students' perceptions of the three types of assessments (2)

Washback: Questionnaire items	Assessment type (I)	Assessment type (J)	Mean (I-J)	Standard error	Sig.	r
1 The assessment is fair.	Summative	Formative 1	.243	.095	.030	.20
	Summative	Formative 2	.602	.095	.000	.37
	Formative 1	Formative 2	.359	.095	.001	.31
2 The assessment can measure my English ability.	Summative	Formative 2	.595	.114	.000	.37
	Formative 1	Formative 2	.427	.114	.001	.31
3 I am nervous and feel pressure when completing the assessment tasks.	Summative	Formative 1	-.421	.145	.011	.24
	Summative	Formative 2	.724	.146	.000	.37
	Formative 1	Formative 2	-1.14	.146	.000	.37
4 Preparing for the assessment is a meaningful and worthwhile experience.	Summative	Formative 2	-.310	.095	.003	.28
5 I have studied a variety of learning methods and activities from the assessment.	Summative	Formative 1	-.458	.118	.000	.37
	Summative	Formative 2	-.891	.119	.000	.37
	Formative 1	Formative 2	-.433	.119	.001	.31
6 Preparing for the assessment has increased my knowledge of grammar and bank size of vocabulary.	Summative	Formative 1	.280	.089	.005	.27
	Formative 1	Formative 2	-.329	.089	.001	.31
7 Preparing for the assessment has enhanced my listening skills.	Summative	Formative 1	.467	.121	.001	.31
	Summative	Formative 2	.528	.131	.000	.31
8 Preparing for the assessment has enhanced my reading skills.	Summative	Formative 1	.776	.107	.000	.37
	Summative	Formative 2	.384	.107	.001	.31
	Formative 1	Formative 2	-.392	.107	.001	.31
9 Preparing for the assessment has enhanced my oral skills.	Summative	Formative 1	-1.290	.123	.000	.37
	Summative	Formative 2	-1.330	.123	.000	.37
10 Preparing for the assessment has enhanced my writing skills.	Summative	Formative 2	-.622	.151	.000	.37
	Formative 1	Formative 2	-.791	.151	.000	.37
11 I am in favor of the assessment.	Summative	Formative 1	-.561	.133	.000	.37
	Summative	Formative 2	-.503	.134	.001	.31

2. What are the differences between students' perceptions of summative and formative assessments based on their levels of proficiency?

According to an independent t-test, we present the main findings regarding the statistically significant differences and effect sizes between high- and low-proficiency students' perceptions of the three types of assessments, as shown in Tables 4-7.

High-proficiency students favor summative assessments more than low-proficiency students because they can increase the size of their vocabulary bank. Low-proficiency students favor summative assessments, only because these are important tests for them, and they might need to repeat the class if they cannot pass the summative assessments.

High-proficiency students favor formative assessments more than low-proficiency students because these tests can enhance their vocabulary, and they are more aware of the assessment tasks in terms of how to manage their time to prepare for them and when they will take place. Low-proficiency students favor formative assessments more than high-proficiency students because they can measure their English ability, enhance their listening and reading skills, and believe that these assessment tasks are relevant to real life. These findings appear to indicate that high-achieving students are in favor of the summative assessment due to its congruence with planned learning and transparency; most of these students admitted that they learned more from preparing for the summative assessment. Furthermore, low-achieving students viewed the cooperative group assessment positively because of its diversity.

The effect sizes for the significant differences between high-proficiency and low-proficiency students in terms of their preferences in summative and formative assessments are to a small degree. This finding indicates that the type of assessment may not be the major concern for students, but how students view the stakes, purpose, task demands, and difficulty of the assessment will influence their learning effects and learning attitude. Given these findings, low-proficiency students seem to benefit more from formative assessments, because they believe these assessments can measure their English proficiency, and preparing for formative assessments can improve their listening and reading skills. However, these formative assessments did not produce a significant improvement in their learning outcomes. Traditional paper-and-pencil written assessments may be beneficial for high-proficiency students to increase the size of their vocabulary bank. However, in order to motivate low-proficiency students, formative oral and cooperative assessments may be another solution for teachers in the English classroom.

As suggested by Peterson & Siadat (2009) and O'Neill (2102), this study echoes the above findings to propose the type of balanced assessment in which the teacher brings many and various strands of assessment together in a coherent way that addresses the desired goals and takes account of opportunities and constraints in the setting concerned. No matter whether a summative or formative assessment is practiced in class, it is essential to provide congruence with planned learning, authenticity, student consultation, transparency, diversity, and intended positive washback effects, such as meaningful learning experience, enhanced motivation, and proficiency levels.

High-achieving and low-achieving students, based on the findings, have different preferences toward the summative and formative assessments regarding their washback effects. In particular, low-achieving students focused on the stakes, the demand, and the purposes of the assessment. In contrast, high-achieving students paid attention to the quality of classroom

assessment, such as its student consultation, transparency, and diversity, in addition to washback effects. Given these findings, when implementing both formative and summative assessments, teachers should clearly inform their students how, what, and when they will be assessed in order for them to know what they should prepare or work on for the assessment with the final goal of helping teachers understand their students' learning processes and outcomes.

Another finding that should catch teachers' attention is that students have more favorable perceptions of formative assessment over summative assessment. This finding appears to promote the concept of assessment for learning (Klenowski, 2009), where teachers design assessment tasks that require students to do what they are interested in by using a variety of classroom activities, so that teachers can use the data collected from the students to help them improve students' learning. In Taiwan's education, university teachers usually instruct a large class size (around 45-60 for general English classes), and therefore summative assessments are usually given to assist teachers in understanding their students' learning outcomes, because it is easier to mark summative assessments. Therefore, time and labor constraints may hinder teachers' use of formative assessments. It is thus suggested that teachers be provided with more educational resources to encourage them to use formative assessments for understanding students' learning process and to help improve their learning as the end goal.

Table 4

High- and low-proficiency students' perceptions of the three types of assessments (1)

High-proficiency learners' perceptions of summative written assessment				Low-proficiency learners' perceptions of summative written assessment			
Ranking	Items	M	SD	Ranking	items	M	SD
Top 1	I am told in advance when I will be assessed.	4.63	.486	Top 1	I am told in advance when I will be assessed.	4.47	.573
Top 2	My teacher has explained to me the purpose of the assessment.	4.37	.561	Top 2	Passing the assessment is important to me.	4.47	.573
Top 3	Preparing for the assessment has increased my knowledge of grammar and amount of vocabulary.	4.35	.590	Top 3	Preparing for the assessment has enhanced my reading skills.	4.29	.658
Top 4	Preparing for the assessment has enhanced my reading skills.	4.35	.653	Top 4	My teacher has explained to me the purpose of the assessment.	4.20	.524
Top 5	I understand how the assessment is evaluated.	4.21	.776	Top 5	How the assessment is evaluated is fair.	4.20	.558
Btm 1	Preparing for the assessment has enhanced my written skills.	2.87	1.253	Btm 1	Preparing for the assessment has enhanced my oral skills.	2.95	1.161
Btm 2	Preparing for the assessment has enhanced my oral skills.	2.92	.837	Btm 2	Preparing for the assessment has enhanced my written skills.	3.05	1.096
Btm 3	I am nervous and feel pressure when taking the assessment.	3.25	1.064	Btm 3	I have spent a lot of time preparing for the assessment.	3.29	.956
Btm 4	I am in favor of the assessment.	3.25	1.046	Btm 4	I am nervous and feel pressure when taking the assessment.	3.29	1.100
Btm 5	I have experienced a variety of	3.44	1.037	Btm 5	I am in favor of the assessment.	3.33	1.055

learning methods and activities while preparing for the assessment.						
---	--	--	--	--	--	--

Note: summative= written assessment; formative 1 = individual oral presentation; formative 2 = group project.

Table 5

High- and low-proficiency students' perceptions of the three types of assessments (2)

High-proficiency learners' perceptions of formative assessment (1)				Low-proficiency learners' perceptions of Formative assessment (1)			
Ranking	Items	M	SD	Ranking	items	M	SD
Top 1	I am told in advance when I will be assessed.	4.52	.641	Top 1	Passing the assessment is important to me.	4.44	.536
Top 2	My teacher has explained to me the purpose of the assessment.	4.37	.595	Top 2	My teacher has explained to me the purpose of the assessment.	4.31	.540
Top 3	I am aware of how I will be assessed and what I will be assessed upon in the English class.	4.35	.556	Top 3	I am told in advance when I will be assessed.	4.31	.635
Top 4	Preparing for the assessment has enhanced my oral skills.	4.23	.703	Top 4	Preparing for the assessment has enhanced my oral skills.	4.29	.875
Top 5	Passing the assessment is important to me.	4.19	.817	Top 5	I am in favor of the assessment.	4.25	.700
Btm 1	Preparing for the assessment has enhanced my written skills.	2.63	1.138	Btm 1	Preparing for the assessment has enhanced my written skills.	2.95	1.113
Btm 2	I have spent a lot of time preparing for the assessment.	3.21	1.016	Btm 2	I have spent a lot of time preparing for the assessment.	3.53	.979
Btm 3	Preparing for the assessment has enhanced my listening skills.	3.33	1.024	Btm 3	Preparing for the assessment has enhanced my reading skills.	3.58	.875
Btm 4	Preparing for the assessment has enhanced my reading skills.	3.50	1.038	Btm 4	I am nervous and feel pressure when taking the assessment.	3.73	1.079
Btm 5	The assessment can measure my English ability.	3.54	.753	Btm 5	Preparing for the assessment has enhanced my listening skills.	3.75	.947

Table 6
High- and low-proficiency students' perceptions of the three types of assessments (3)

High-proficiency learners' perceptions of formative assessment (2)				Low-proficiency learners' perceptions of formative assessment (2)			
Ranking	Items	M	SD	Ranking	Items	M	SD
Top 1	I am told in advance when I will be assessed.	4.65	.480	Top 1	I have experienced a variety of learning methods and activities while preparing for the assessment.	4.44	.536
Top 2	Preparing for the assessment has increased my knowledge of grammar and amount of vocabulary.	4.33	.550	Top 2	I am told in advance when I will be assessed.	4.31	.540
Top 3	I have experienced a variety of learning methods and activities while preparing for the assessment.	4.29	.776	Top 3	Preparing for the assessment has enhanced my oral skills.	4.31	.635
Top 4	Preparing for the assessment has enhanced my oral skills.	4.29	.723	Top 4	Passing the assessment is important to me.	4.29	.875
Top 5	I am aware of how I will be assessed and what I will be assessed upon in the English class.	4.27	.717	Top 5	Preparing for the assessment is a meaningful and worthwhile experience.	4.25	.700

Note: summative= written assessment; formative 1 = individual oral presentation; formative 2 = group project.

Table 6
(Continued)

High-proficiency learners' perceptions of formative assessment (2)				Low-proficiency learners' perceptions of formative assessment (2)			
Ranking	Items	M	SD	Ranking	items	M	SD
Btm 1	I am nervous and feel pressure when taking the assessment.	2.62	1.105	Btm 1	I am nervous and feel pressure when taking the assessment.	2.95	1.113
Btm 2	The assessment can measure my English ability.	3.21	.977	Btm 2	The assessment can measure my English ability.	3.53	.979
Btm 3	Preparing for the assessment has enhanced my listening skills.	3.27	.931	Btm 3	How the assessment is evaluated is fair.	3.58	.875
Btm 4	I find the English assessment tasks to be relevant to the real world.	3.50	1.038	Btm 4	I have spent a lot of time preparing for the assessment.	3.73	1.079
Btm 5	Preparing for the assessment has enhanced my listening skills.	3.54	.979	Btm 5	Preparing for the assessment has enhanced my writing skills.	3.75	.947

Table 7

A summary of high- and low-proficiency students' perceptions of the three types of assessments (4)

A comparison of high-proficiency learners' and low-proficiency learners' perceptions of the three types of assessments

Assessment type	Items	<i>t</i>	<i>Sig.</i>	<i>r</i>
Summative	Passing the assessment is important to me.	1.989	.049	.19
	I am allowed to complete assessment tasks at my own speed.	2.347	.021	.22
	Preparing for the assessment has increased my knowledge of grammar and amount of vocabulary.	2.432	.017	.23
Formative 1	The assessment can measure my English ability.	2.690	.008	.25
	Preparing for the assessment has increased my knowledge of grammar and amount of vocabulary.	2.069	.041	.20
	Preparing for the assessment has enhanced my listening skills.	2.197	.030	.21
	I find the English assessment tasks to be relevant to the real world.	2.450	.016	.23
	I am in favor of the assessment.	2.017	.046	.19
Formative 2	Passing the assessment is important to me.	2.003	.048	.19
	I am aware of how I will be assessed and what I will be assessed upon in the English class.	2.198	.030	.21
	I am told in advance when I will be assessed.	2.668	.009	.25
	Preparing for the assessment has enhanced my listening skills.	2.122	.036	.20
	I find the English assessment tasks to be relevant to the real world.	2.489	.014	.24

Note: summative= written assessment; formative 1 = individual oral presentation; formative 2 = group project.

Conclusions

The purpose of this study has been to explore Taiwan university students' perceptions of summative and formative classroom assessment tasks in their English classes. To address this issue, a questionnaire was designed, based upon Dorman and Knightley's (2006) PATI and Green's (2007) model of washback. According to survey questionnaires given to 107 first-year

undergraduate students at one Taiwan university, this study discovers that: 1) students were in favor of the summative assessment due to its congruence with planned learning and transparency, with most students admitting that they learned more from preparing for the summative assessment; 2) students viewed the cooperative group assessment positively, because of its diversity; 3) preparation for summative assessments elicited a greater degree of test-oriented learning for respective skills, whereas formative assessments enhanced students' motivation to learn English for productive skills; and 4) students believed that an appropriate combination of summative and formative assessment tasks was beneficial for their learning. Based on the findings, a combination of summative and formative assessments should be given appropriately to better understand students' learning outcomes and learning processes. Since the findings were drawn from a small sample size, further research should recruit a larger size of student participants for better generalizing the research findings. Teachers, who are also the major stakeholder of the assessments, should also be queried in order to present additional evidence for understanding the quality and washback of classroom assessment.

References

- Alderson, C., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14, 115-129.
- Black, P. (2000). Research and the development of education assessment. *Oxford Review of Education*, 26(3&4), 407-419.
- Brown, H. D. (2004). *Language assessment: principles and classroom practices*. New York: Pearson Education/Noyce.
- Cheng, L. (2005). *Changing language teaching through language testing: washback study*. Cambridge: Cambridge University Press.
- Cheng, L. (2014). Consequences, impact, and washback. In A. J. Kunnan (Ed.), *The companion to language assessment* (pp. 1130–1146). John Wiley & Sons.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum
- Dorman, J. P., & Knightley, W. M. (2006). Development and Validation of an Instrument to Assess Secondary School Students' Perceptions of Assessment Tasks. *Educational Studies*, 32(1), 47-58.
- Earl, L. (2003). *Assessment as learning: Using classroom to maximize student learning*. Thousand Oaks, CA: Corwin Press.
- Earl, L., & Katz, S. (2005). *Rethinking Classroom Assessment With Purpose in Mind*. Winnipeg, Manitoba: Western Northern Canadian Protocol.
- Ferman, I. (2004). The washback of an EFL national oral matriculation test to teaching and learning. In L. Cheng, Y. Watanabe & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods* (pp. 191-210). Mahwah, New Jersey: Lawrence Erlbaum.
- Gan, Z. (2009). IELTS Preparation Course and Student IELTS Performance: A Case Study in Hong Kong. *RELC Journal*, 40(1) 23-41.
- Gebril, A., & Brown, G. (2014). The effect of high-stakes examination systems on teacher beliefs: Egyptian teachers' conceptions of assessment. *Assessment in Education: Principles, Policy & Practice*, 21(1), 16-33.
- Green, A. (2007). *IELTS washback in context: Preparation for academic writing in higher education*. Cambridge: Cambridge University Press.
- Hayes, B., & Read, J. (2003). IELTS Test preparation in New Zealand: Preparing students for the IELTS academic module. In R. Tolloh (Ed.), *IELTS Research Report 4* (pp. 153-206). Canberra: IELTS Australia Pty Limited.
- Hughes, A. (1993). *Backwash and TOEFL 2000*. Unpublished manuscript, University of Reading, England.
- Hughes, A. (2003). *Testing for language teachers* (2nd ed.). Cambridge: Cambridge University Press.
- Klenowski, V. (2009) Assessment for learning revisited: an Asia-Pacific perspective. *Assessment in Education: Principles, Policy and Practice*, 16(3), 263-268.

- Larson-Hall, J. (2012). Our statistical intuitions may be misleading us: Why we need robust statistics. *Language Teaching*, 45(4), 460-474.
- Larson-Hall, J., & Plonsky, L. (2015). Reporting and interpreting quantitative research findings: What gets reported and recommendations for the field. *Language Learning*, 65(S1), 127-159.
- O'Neill, G. (2012). UCD Assessment ReDesign Project: The Balance Between Assessment FOR and OF Learning, available at <https://www.ucd.ie/t4cms/UCDTLA0044.pdf>
- Peterson, E., & Siadat, M. V. (2009). Combination of Formative and Summative Assessment Instruments in Elementary Algebra Classes: A Prescription for Success. *Journal of Applied Research in the Community College*, 16(2), 92-102.
- Qi, L. (2005). Stakeholders' conflicting aims undermine the washback function of a 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(1), 51-74.
- Shih, C. (2007). A new washback model of students' learning. *The Canadian Modern Language Review*, 64(1), 135-162.
- Smyth, E., & Banks, J. (2012). High stakes testing and student perspectives on teaching and learning in the Republic of Ireland. *Education Assessment Evaluation and Accountability*, 24:283-306. DOI 10.1007/s11092-012-9154-6.
- Tsai, Y., & Tsou, C. (2009). A standardized English Language Proficiency test as the graduation benchmark: student perspectives on its application in higher education. *Assessment in Education: Principles, Policy & Practice*, 16(3), 319-330.
- Wall, D., & Horák, T. (2006). *The impact of changes in the TOEFL® examination on teaching: Phase 1, The baseline study* (TOEFL-MS-34). Princeton, NJ: ETS.
- Wall, D., & Horák, T. (2008). *The impact of changes in the TOEFL® examination on teaching and learning in Central and Eastern Europe: Phase 2, Measuring change* (TOEFL iBT-05). Princeton, NJ: ETS,
- Wall, D., & Horák, T. (2011). *The Impact of Changes in the TOEFL® Exam on Teaching in a Sample of Countries in Europe: Phase 3, The Role of the Coursebook. Phase 4, Describing Change* (TOEFL iBT-17). Princeton, NJ: ETS.
- Xie, Q. (2013). Does test preparation work? Implications for score validity. *Language Assessment Quarterly*, 10, 196-218.
- 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, 56-88.
- Zhan, Y., & Andrew, 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.

Appendix 1: Alderson and Wall's fifteen washback hypotheses
(1993, pp. 120-121)

- 1) A test will influence teaching.
- 2) A test will influence learning.
- 3) A test will influence what teachers teach; and
- 4) A test will influence how teachers teach; and therefore by extension from (2) above:
- 5) A test will influence what learners learn;
- 6) A test will influence how learners learn.
- 7) A test will influence the rate and sequence of teaching; and
- 8) A test will influence the rate and sequence of learning.
- 9) A test will influence the degree and depth of teaching.
- 10) A test will influence the degree and depth of learning.
- 11) A test will influence attitudes to the content, method, etc. of teaching and learning.
- 12) Tests that have important consequences will have washback; conversely,
- 13) Tests that do not have important consequences will have no washback.
- 14) Tests will have washback on all learners and teachers.
- 15) Tests will have washback effects for some learners and some teachers, but not for others.