Research Report

ETS RR–17-45

The Case of Taiwan: Perceptions of College Students About the Use of the *TOEIC®* Tests as a Condition of Graduation

Ching-Ni Hsieh

December 2017

ETS Research Report Series

Since its 1947 founding, ETS has conducted and disseminated scientific research to support its products and services, and to advance the measurement and education fields. In keeping with these goals, ETS is committed to making its research freely available to the professional community and to the general public. Published accounts of ETS research, including papers in the ETS Research Report series, undergo a formal peer-review process by ETS staff to ensure that they meet established scientific and professional standards. All such ETS-conducted peer reviews are in addition to any reviews that outside organizations may provide as part of their own publication processes. Peer review notwithstanding, the positions expressed in the ETS Research Report series and other published accounts of ETS research are those of the authors and not necessarily those of the Officers and Trustees of Educational Testing Service.

The Daniel Eignor Editorship is named in honor of Dr. Daniel R. Eignor, who from 2001 until 2011 served the Research and Development division as Editor for the ETS Research Report series. The Eignor Editorship has been created to recognize the pivotal leadership role that Dr. Eignor played in the research publication process at ETS.

The Case of Taiwan: Perceptions of College Students About the Use of the *TOEIC*® Tests as a Condition of Graduation

Ching-Ni Hsieh

Educational Testing Service, Princeton, NJ

This study uses online surveys and phone interviews with Taiwanese college students to investigate their perceptions of using the *TOEIC*® test scores to meet an English-language graduation requirement. Results indicate that students have positive views about the use of the TOEIC test scores for graduation and believe that preparing to take the test has a positive impact on their language proficiency and future employment prospects. The TOEIC test scores are also perceived favorably by the participants as having high levels of reliability and validity. The study provides empirical evidence to support the use of the TOEIC test as a high-stakes test for college exit requirement in Taiwan and has implications for similar test use in other countries

Research has shown that high-stakes tests can have a critical impact on test takers and other stakeholders (e.g., Alderson & Hamp-Lyons, 1996; Cheng, Andrews, & Yu, 2010; Hamp-Lyons, 1997, 1998; Murray, Riazi, & Cross, 2012; Qi, 2005, 2007; Wall, 2000). High-stakes tests are those that are used for making important educational and professional decisions about individuals, such as admissions, graduation, employment, or promotion. One of the high-stakes testing situations that has important consequences for students is the use of an exit test as a condition of graduation. Exit tests are generally considered to represent one of the mechanisms to ensure accountability in education (Berry & Lewkowicz, 2000). In the field of language testing and assessment thus far, relatively few studies have examined the use of standardized language proficiency tests as exit tests and their impact on teaching and learning (e.g., Berry & Lewkowicz, 2000; Nhan, 2013; Tasi & Tsou, 2009). The current study adds to this line of research with an examination of college students' perceptions of the English-language graduation requirement policy implemented by Taiwan's higher education (HE) institutions and the appropriateness of the *TOEIC*® tests as an exit test in this context. The use of the TOEIC as an exit test has a direct impact on students who take the test to meet the graduation requirement, and it has important educational and economic implications for students and society at large. Language learners or students whose education and language learning are directly influenced by the implementation of the language graduation requirement policy are perhaps the most important stakeholders in this testing situation. However, their views on the policy implementation are rarely consulted, if at all (Shih, 2010). Previous research has suggested that learners' attitudes toward and perception of a test and its use can affect their motivation and performance on the test (Bachman & Palmer, 1996). Given that the test takers' perceptions can have wide-ranging consequences, opinions about the use of the TOEIC tests as an exit test warrant further investigation to ensure the tests' validity.

Research Context

Since 2003, the Ministry of Education (MOE) in Taiwan has encouraged HE institutions to set English proficiency thresholds for undergraduates and to implement an English-language requirement policy for graduation. The objective is to raise students' English language proficiency (ELP) and to better prepare students to cope with global competition and meet the ELP requirements of the workplace (Shih, 2009, 2010, 2012). To this end, the MOE has advocated the use of standardized English proficiency tests as exit requirements. The TOEIC tests, along with other standardized tests such as the *TOEFL*® test, International English Language Testing System (IELTS), and the locally developed General English

Corresponding author: C.-N. Hsieh, E-mail: chsieh@ets.org

Proficiency Test (GEPT) have been recommended by the MOE to meet this objective. As a result, the majority of colleges and universities in Taiwan have set a requirement for students to achieve a satisfactory score on one of the recommended English proficiency tests before graduation.

Taiwan has two major types of 4-year HE institutions: general universities and technical colleges. General universities usually have more rigorous admissions standards with respect to test scores compared to technical colleges that focus on vocational education or training (although there are different tiers of general universities and technical colleges). Taiwan's MOE has advised HE institutions that adopt the exit requirement policy to set varied criteria according to students' language learning and communication needs and levels of proficiency. The MOE has recommended that general universities set an English-language graduation criterion at the B1 or above level on the Common European Framework of Reference (CEFR) as put forth by the Council of Europe (2001). Furthermore, the MOE has suggested that technical colleges adopt the A2 level as their benchmark for graduation. Apart from recommending the CEFR levels as the English-language graduation benchmark, the MOE does not prescribe what English language skills should be tested or what test scores on which tests should be required. University authorities could formulate their own policies and decide upon the specific ELP requirements for their students (Shih, 2010, 2012).

The TOEIC, one of the MOE's recommended proficiency tests for college graduation, is an ELP test for those whose native language is not English (Educational Testing Service [ETS], 2015). It measures the everyday English skills of people working in an international environment. Test scores indicate how well test takers can communicate with others in English in business, commerce, and industry. The test does not require specialized knowledge or vocabulary beyond that of a person who uses English in everyday work activities. The primary uses of the TOEIC tests are to allow test takers to verify their current level of English proficiency, qualify for a new position or promotion in a company, enhance their professional credentials, monitor their progress in English, set their own learning goals, and involve their employer in advancing their English ability. In the past decade, the TOEIC has gained increasingly wide recognition by test takers and score users in Taiwan, and the test scores are now being used extensively for recruitment and promotion by both domestic and multinational corporations and organizations (Pan & Roever, 2016).

Since the mid-2000s, many of Taiwan's HE institutions began to accept the TOEIC as one of the language proficiency tests suitable for exit purposes. The majority of those that accept the TOEIC tests require students to take only the TOEIC Listening and Reading (TOEIC LR) test and earn test scores that meet the minimum requirements for graduation. The TOEIC Listening section assesses how well test takers understand spoken English, and the TOEIC Reading section tests how well test takers understand written English. The TOEIC LR is designed to enable test takers to demonstrate their English listening and reading skills and thereby qualify for better employment opportunities and gain a competitive edge in the global workplace.

As of 2015, 133 (79%) of Taiwan's 169 HE institutions had complied with the MOE directive and had implemented the English-language graduation requirement policy; all of these accept TOEIC LR scores (Nichols, 2016). TOEIC Speaking and Writing test scores, in contrast, are less commonly required for college graduation, except for certain academic disciplines such as English or business, primarily because of the concerns about testing fees and the possible low passing rates.

Scores from the TOEIC LR test have been mapped onto the CEFR (Tannenbaum & Wylie, 2013). The minimum cut score for B1 level is a total TOEIC LR score of 550, and the minimum cut score for A2 is 225. The majority of the universities and colleges that have adopted the TOEIC tests have set a cut score for graduation between 450 and 550. Some top-tier institutions have set a higher bar; others have set different criteria for different majors (Shih, 2012). For example, English and business majors are usually required to obtain a higher passing score than their non-English or nonbusiness major counterparts (Pan & Roever, 2016).

A growing body of empirical studies has investigated the impact of the English language graduation requirement policy on language learning and teaching in Taiwan. Most studies have examined the viewpoints of students and teachers from technical colleges (e.g., Chu, 2009; Hsu, 2009; Pan, 2014; Pan & Newfields, 2011; Shih, 2009, 2010; Tasi & Tsou, 2009). Few studies have included perceptions of stakeholders from general universities (e.g., Chen, 2008; Vongpumivitch, 2006; Wu, 2012). These studies have found conflicting views regarding whether students are in support of or against the language requirement policy. In addition, students in technical colleges broadly reported that the policy causes pressure and anxiety. These feelings were especially strong for low-proficiency students and for those who had little interest or motivation to learn English, who felt that they were forced to "study for the test" (Hsu, 2009; Tasi & Tsou, 2009). Most of the studies (e.g.,

Pan, 2014; Shih, 2010) also showed that the policy has resulted in limited or no washback (positive or negative effects in the classroom). In contrast, Chen (2008) reported that students from a top-tier, general university thought that the exit requirement policy helped students improve their ELP, even though the policy did not play a significant role in their learning motivation.

Interestingly, but perhaps unsurprisingly, most of the publicly available empirical studies conducted in the Taiwanese context have examined the use of the GEPT, the locally developed test, which was once the most popular exit test (Roever & Pan, 2008; Shih, 2007). Little research exists that examines the use of the TOEIC as an exit test in the Taiwanese HE context and how the Taiwanese college students perceive the utility of the test. As Shohamy, Donitsa-Schmidt, and Ferman (1996) suggested, the perceived status of a test, such as whether a test is locally or internationally known, is directly linked to students' motivation, time devoted, and effort exerted to prepare for the test. It remains unclear whether the use of the TOEIC—a test with an international reputation and recognition for validity—is more likely to be perceived positively as a college exit test by test takers.

Within the context of the exit requirement policy in Taiwan, previous studies have suggested that students' views about the policy are often not considered when university authorities set the policy (Shih, 2010). Students are arguably the most important stakeholders in this testing situation, and their opinions about the implementation of the policy and its impact warrant further investigation. Previous research has also suggested that students of varying proficiency could have differing views about the policy and its impact on learning. Investigations into these issues could provide the MOE and policymakers with important insights into the appropriateness of the graduation benchmarks or cut scores. This study examined the opinions of students on the language requirement policy and the efficacy of the TOEIC as an exit test. The research questions guiding this study are as follows:

1. How do Taiwanese college students perceive the English-language graduation requirement policy? Are their perceptions related to their ELP?
2. How do Taiwanese college students who take the TOEIC for graduation perceive the efficacy of the TOEIC as an exit test? Do students of high and low proficiency differ in their perceptions?

Methodology

Instruments

The instruments used in this study included an online student survey and a set of semistructured interview questions. The survey questions asked about students' biographic, educational, and language learning backgrounds; TOEIC test-taking experiences; and perceptions about the English-language graduation requirement policy and the use of the TOEIC as an exit test. The survey was created using the online survey software SurveyMonkey (www.surveymonkey.com). The interview further explored factors that influenced students' perceptions.

Participants

Student Survey

The ETS representative in Taiwan served as the site coordinator for the project, sending out the link to the survey to more than 22,000 college students who had received valid TOEIC LR scores (i.e., TOEIC LR scores obtained within the past 2 years) at the time of data collection. A total of 1,527 surveys with valid responses were returned (response rate = 7.1%). Information about the respondents' latest TOEIC LR scores and the number of times test takers had taken the TOEIC were extracted from the official TOEIC test database. There were 361 males and 1,145 females; 21 respondents preferred not to respond to the gender question. The students' ages ranged between 19 and 24 ($M = 21.6$, $SD = 0.9$), and they had taken the TOEIC LR between one and 11 times ($M = 3.11$, $SD = 1.33$). The students' most recent TOEIC LR scores ranged between 175 and 980 ($M = 610$, $SD = 169.4$). The survey respondents represented 123 Taiwan HE institutions that used TOEIC as an exit test, with a total of 75 general universities and 48 technical colleges. The students came from a wide range of academic disciplines, broadly representing arts and humanities (24%), social sciences (6%), natural sciences and engineering (20%), education (2%), and business (34%), with 14% missing data.

All survey respondents were asked about their reasons for taking the TOEIC and their test-preparation activities. Both questions allowed students to include all relevant answers. The most commonly reported reason was "to meet the

graduation requirement" (71.1%), followed by "to improve English language proficiency" (59.7%), "to prepare for the job market" (51.4%), "to get an internationally recognized English proficiency test certificate" (37%), "to qualify for a better paying job" (31.6%), and "to communicate with English speakers" (22.7%). When asked how they prepared for the TOEIC, 69.0% reported using practice tests, 38.6% enrolled in language courses offered by their university or college, and 8.1% took courses at private language schools. Few students (12.9%) took TOEIC test-preparation courses at test-preparation training schools. The majority of the respondents (71.4%) also reported watching movies in English, reading magazines in English, listening to radio programs in English, and the like to prepare for the test. Very few students (2.8%) hired personal tutors or made foreign friends to help improve their proficiency in English.

Student Interviews

Twenty-six randomly selected survey respondents who expressed interest in participating in the follow-up interviews were interviewed by the researcher. The interviewees included 12 males and 14 females between the ages of 19 and 23 ($M = 21.38$, $SD = 1.06$). They had taken the TOEIC between two and five times ($M = 2.73$, $SD = 0.87$). Their most recent TOEIC LR scores ranged from 405 to 945, and the mean score was 681.15 ($SD = 157.18$). Although the interviewees' demographics largely mirrored those of the survey respondents in terms of age and gender, their TOEIC LR mean score was higher than the mean score of the total pool of the survey respondents, suggesting that the interviewees were overall more proficient language users. There were 21 students from general universities and five from technical colleges. The students represented 10 different institutions.

Procedures

Student Survey

The link to the online survey was sent to the site coordinator, who distributed it to TOEIC test takers who met the selection criteria (i.e., college students who had taken the TOEIC LR at least once, had valid TOEIC LR scores, and whose university or college used TOEIC LR as one of its English language exit tests). The survey stayed live for roughly 1 month to allow the students sufficient time to respond.

Student Interviews

The 26 students who agreed to participate in the follow-up interviews were contacted by the site coordinator to set up the interview sessions. The interviews were conducted one-on-one between the researcher and the students through phone calls. All interviews were conducted in Mandarin Chinese, the researcher's and the students' first language. Each interview session lasted between 15 and 20 minutes and was audio-recorded.

Data Analysis

Student Survey

The responses to the online survey were extracted from the SurveyMonkey website. Frequency counts, descriptive statistics, and cross-tab analyses were performed to answer the research questions.

Student Interviews

The audio recordings of the interviews were translated from Mandarin Chinese to English and then transcribed into English semantically by the researcher. The transcripts were used as data for analysis. First, the researcher read and reread the transcripts in their entirety several times to obtain a holistic impression of the responses and to determine a preliminary list of eight analytical categories that had the potential to provide answers to the research questions. For example, one of the categories was policy goals, intended to examine the students' viewpoints on the goals of the language exit requirement policy. Further analysis was conducted to identify different variants under each analytical category. For example, under the analytical category, policy's impact on ELP, three variants emerged from the data: *positive*, *negative*, and *no*

impact. Each variant was defined and described in the coding scheme following the procedures for analyzing semistructured interviews proposed by Schmidt (2004). The coding scheme was subsequently revised to incorporate new analytical categories or variants and to remove those that were redundant or irrelevant. Once the coding scheme was finalized, six of the transcripts were double-coded by the researcher and a research assistant to establish coder reliability. Interrater agreement was reached at 95%. Discrepant cases were resolved through discussion. The researcher then coded the entire dataset using the qualitative data analysis software NVivo 11 (www.qsrinternational.com/nvivo-product/nvivo11-for-windows).

Results and Discussion

This section presents the results by research question and discusses their implications.

Research Question 1: How do Taiwanese college students perceive the English-language graduation requirement policy? Are their perceptions related to their ELP?

Student Survey Results

The students were asked to respond to statements about their viewpoints regarding the necessity of the ELP requirement policy and the policy's impact on their language learning (see Table 1). The majority (82.3%) supported the policy and thought that HE institutions should require students to achieve a certain level of English proficiency before graduation. Interestingly, more than 90% of the students indicated that they would be motivated to study English even if the policy were not in place, implying that the policy is not a primary motivator for many students.

Although previous case studies found that Taiwanese college students felt that they were forced to take standardized language proficiency tests due to the exit requirement (e.g., Chu, 2009; Hsu, 2009), 81.9% of the respondents indicated that they would still be motivated to prepare and take standardized language proficiency tests, regardless of the presence of the policy. Contrary to previous findings, this result suggests that in the test-driven educational environment in Taiwan, students would take standardized ELP tests for reasons other than the graduation requirement. It can be argued that because there was a relatively high percentage of respondents who took the TOEIC as a graduation requirement, the use of the TOEIC as an exit test could have impacted how respondents perceived the necessity and implementation of the policy. These students may perceive added value in obtaining an internationally recognized ELP certificate such as the TOEIC, insofar as it is beneficial to have a policy that encourages them to obtain a satisfactory score on the TOEIC in preparation for their future employment. The students' future job prospects and impact on credentials needed for job applications may serve as a strong impetus for them to take standardized ELP tests.

To explore whether students' opinions about the exit requirement policy are related to their language proficiency, the survey respondents were first split into two proficiency groups, *high* and *low* (i.e., falling above and below respondents' median TOEIC LR score of 615). It should be noted that these were relative proficiency groups only for the convenience of the analysis and should not be interpreted as representing different proficiency levels based on the TOEIC LR scores of a more representative sample. The mean difference in the total TOEIC LR scores between the high- and low-proficiency groups was 278.3 total score points, and the difference was statistically significant, $t(1,483) = 56.24$, $p < .001$, $d = 2.87$.

Cross-tabulations were performed to determine if language proficiency could predict students' viewpoints for or against the language graduation requirement policy (see Table 2). The analysis was conducted within each type of HE institution.

Table 1 Students' Opinions About the Exit Requirement Policy

Statement	Agree (%)	Disagree (%)
It is necessary for universities to require their students to obtain a certain score on an English proficiency test before graduation.	83.2	16.7
I would still be motivated to study English in college even if there were no English graduation requirement.	91.7	8.3
I would still be motivated to prepare to take standardized English language proficiency tests even if there were no English graduation requirement.	81.9	18.1

Note. Total sample $N = 1,527$.

Table 2 Opinions About the Exit Requirement Policy by Proficiency

Type of higher education institution	TOEIC scores		Opinion about exit requirement policy		N	χ^2
	Mean	*SD*	Necessary	Not necessary	Total	Sig. (2-sided)
General university						.001
High	754.57	86.47	507	89	596	
Low	506.46	80.49	312	93	405	
Technical college						.006
High	733.87	91.39	152	13	165	
Low	433.15	114.95	300	61	361	

Table 3 Student Interviewees' Perceptions About the English-Language Graduation Requirement Policy

Main category	Subcategory	Number of students	Number of comments
Policy goals	To ensure students have minimum ELP for communication	18	23
	To help improve students' ELP	18	28
	To motivate students to study English	12	17
	To prepare students for further study	5	6
	To prepare students for future employment	12	17
Policy impact on ELP	Positive	22	39
	Negative	0	0
	No impact	3	4
Policy impact on motivation	Positive	18	30
	Negative	3	4
	No impact	9	12
	Short-term impact	2	2

Note. ELP = English language proficiency.

Significant differences were found between the high- and low-proficiency groups in their opinions about the necessity of the policy, both for students of general universities, $\chi^2(1) = 10.45$, $p < .01$, and for students of technical colleges, $\chi^2(1) = 7.61$, $p < .01$. Based on the odds ratio effect size, the high-proficiency students of general universities were 1.69 times more likely to support the policy than their low-proficiency counterparts; similarly, the high-proficiency students of technical colleges were 2.38 times more likely to support the policy than their low-proficiency counterparts. The results suggest that more proficient students tend to have a more positive attitude toward the policy, perhaps in part because the policy is less likely to pose a threat for their graduation. It is also possible that they believe that the policy could serve as an incentive for them to obtain an English language certificate that could be beneficial for their future.

Student Interview Results

The interview data extended the survey results and helped illuminate potential factors that influence students' perceptions of the English-language graduation requirement policy. The students' responses centered on three major themes: policy goals, policy's impact on ELP, and policy impact on motivation. Table 3 indicates the number of students who commented on each of the subcategories within each main category and the number of comments made.

The students made numerous comments on what they thought the language policy was designed for. Collectively, the data showed that the students had a fair understanding of the policy goals and were generally in support of its implementation. Eighteen students considered that the policy was meant to ensure that college graduates have the minimum level of English proficiency for communication and that the policy could help students improve their ELP. Twelve students believed that the policy was put in place to motivate students and to prepare them for future employment; five students thought that the policy could help prepare students for their future studies.

When asked about the policy's impact on learning, the interviewees largely commented positively about its influence on their ELP. "I think it's a great policy, because when you go on the job market, English is always required. So it's a good motivator," commented Student 26 (high proficiency, technical college, sophomore). In contrast, several students felt that

the policy had limited or no impact on their ELP or motivation. A few pointed out that the policy's impact was short term in nature: "I'm a senior student now, so the policy pushed me to study English for a while. After I passed the test, I lost my motivation to study English," said Student 13 (low proficiency, general university, senior). Another factor that influenced the students' views about the policy pertained to their perception that an appropriate cut score for graduation was lacking. Student 1 (high proficiency, general university, senior) reported:

> I don't think the policy is meaningful or has any impact on students' learning or motivation because the requirement is so low. In my school, the requirement is 450 on the TOEIC LR. Even if you don't prepare for the test and you just guess randomly, you might be able to get 200 or 300 points. The requirement set by my university was just the English level required for middle-school students. If the requirement were higher, I think it would help students improve their proficiency better.

Similar concerns about the inadequate cut scores were raised in the responses of seven students who generally supported the policy but were skeptical about the positive washback the policy could induce. For instance, Student 2 (low proficiency, technical college, junior) stated:

> Well, I've discussed this issue [low cut score] with many of my friends, and we all agree that a higher graduation requirement would really motivate us to study English. Since the requirement in my school is so low, most students think that it's a piece of cake, and they don't need to work hard and they'll be able to meet the requirement anyways. Although I think it's a good policy, it doesn't really make a dent in motivating students.

Interestingly, Student 2, whose most recent TOEIC LR score was 455 and was one of the least proficient students in the entire interviewee pool, seemed confident that students at his college, which was a second-tier technical college, would not have much difficulty meeting the requirement. Several universities in Taiwan had in fact set a cut score of 450 or lower on the TOEIC LR to avoid having a small passing rate, which could impact the ratings they receive from routine program evaluations conducted by the MOE and the funding they obtain from the government (Shih, 2012). The students' comments above, however, reveal that an inadequate cut score could undermine the goals of the policy and adversely cause students to lose the incentive to work hard, knowing that the threat of failing was minimal.

Research Question 2: How do Taiwanese college students who take the TOEIC for graduation perceive the efficacy of the TOEIC as an exit test? Do students of high and low proficiency differ in their perceptions?

Student Survey Results

Given that college students have multiple standardized tests to choose from for meeting the exit requirement, it is logical to assume that students who choose to take the TOEIC tests would perceive the use of the TOEIC as an exit test in a more positive light. To determine whether this assumption holds, the following analysis included only responses of students ($N = 1,086$) who indicated that they took the TOEIC LR test in order to pass the exit requirement, among other reasons.

Table 4 shows the frequency and percentage of students who agreed or disagreed with each statement related to their perceptions about the use of the TOEIC test. The majority of the students (86.4%) reported that preparing for the TOEIC had a positive impact on their ELP. Students' views about the validity of the test scores, however, were somewhat mixed. A slight majority of respondents (54.7%) indicated that the test scores accurately reflected their level of ELP, and the other half disagreed. The respondents' differing views on score validity may have been influenced by the exit requirement policy that required only TOEIC LR test scores. Students who did not show confidence in the validity of the scores may feel that, as a measure of receptive (listening and reading) and not productive (speaking and writing) skills, the TOEIC LR test could not fully reveal their language competency. The result could also reflect an issue of how the statement was phrased. Had the students been asked about whether the scores demonstrated their English listening and reading abilities in the workplace, the results could have been different.

Students' perceptions regarding the requirement of TOEIC speaking and writing tests for graduation was divided, with slightly less than half (48.3%) admitting that productive skills were important and should be required. It was interesting to note that the students' mean TOEIC LR scores differed significantly between the two groups of students who were in favor

Table 4 Survey Respondents' Perceptions About the Use of the TOEIC as an Exit Test

Statement	Agree (%)	Disagree (%)
Preparing for the TOEIC helps me improve my ELP.	86.4	13.6
TOEIC LR scores accurately reflect my ELP.	54.7	45.3
TOEIC speaking and writing tests should be required for graduation.	48.3	51.7
If there were no English exit requirement, I would still be motivated to prepare for and take the TOEIC.	74.6	25.4

Note. Total sample $N = 1,086$.

Table 5 Survey Respondents' Perceptions About the Efficacy of TOEIC by Proficiency

	Proficiency group		N	χ^2
Survey question	High	Low	Total	Sig. (2-sided)
TOEIC helps improve my proficiency				
Agree	392	546	938	.342
Disagree	68	80	148	
TOEIC scores are valid				
Agree	248	346	594	.657
Disagree	212	280	492	
TOEIC speaking and writing should be required				
Agree	270	255	525	.000
Disagree	190	371	561	
I would take TOEIC without requirement				
Agree	373	437	810	.000
Disagree	87	189	276	

Note. Total sample $N = 1,086$.

or not in favor of the requirement of productive skills, $t(1,084) = 6.53$, $p < .001$, $d = 0.39$ for listening and $t(1,084) = 5.62$, $p < .001$, $d = 0.34$ for reading. The results indicated that students with better listening and reading skills as measured by the TOEIC test had a significantly stronger preference for the requirement of an adequate level of speaking and writing proficiency before graduation. Students with lower receptive skills are less likely to pass the graduation requirement, and thus they might have a fear of adding another obstacle if productive skills are also required.

When asked about their test-taking motivation, almost 75% of the students responded that they would be motivated to prepare to take the TOEIC even if they were not required to pass an English proficiency test before graduation. This high level of test-taking motivation once again demonstrated that the students believe in the utility of the TOEIC test scores for purposes other than passing the exit requirement.

The second part of Research Question 2 addressed whether students of high and low ELP view the efficacy of the TOEIC as an exit test differently. Four separate cross-tab analyses were performed to determine if the perception differences exist (see Table 5). The analysis showed that the high- and low-proficiency students did not differ in their views on the use of the TOEIC to improve their ELP, $\chi^2(1) = .904$, $p = .342$. In general, the majority were positive about its impact on language learning. A nonsignificant difference was also found for students' perceptions on the validity of the TOEIC scores, $\chi^2(1) = .197$, $p = .657$.

With regard to the requirement of the TOEIC speaking and writing test scores for graduation, results of the cross-tab analysis yielded a significant difference between proficiency groups, $\chi^2(1) = 34.25$, $p < .001$. The odds ratio effect size indicated that the high-proficiency group was 2.06 times more supportive of requiring speaking and writing tests compared to the low-proficiency one. A significant difference between proficiency groups was also found for test-taking motivation, $\chi^2(1) = 17.97$, $p < .001$. The odds ratio effect size indicated that high-proficiency students were 1.85 times more likely to prepare to take the TOEIC regardless of the presence of the policy, compared to the low-proficiency ones.

Student Interview Results

The interviewees were asked to discuss their experiences and opinions about preparing for and taking the TOEIC tests and the use of the TOEIC as an exit test. To help answer the research question, the discussion focused on students' opinions

Table 6 Student Interviewees' Perceptions About the Use of TOEIC as an Exit Test

Main category	Subcategory	Number of students	Number of comments
Test design	Content is authentic	9	14
	Measure workplace English	8	9
	Include various native accents	4	6
	A well-designed test	4	6
	Reading section too long	5	5
Test purpose	Future employment	15	9
	Further study	2	3
	Scholarship application	1	1
	Measure student's ELP	11	20
Score validity	Scores reflect actual ELP	18	22
	Scores do not reflect actual ELP	2	3

Note. ELP = English language proficiency.

about the use of the TOEIC as an exit test. Students' perceptions involved three major themes: test design, test purpose, and score validity. Table 6 summarizes the students' responses.

When responding to the question about the utility of the TOEIC as an exit test, the student interviewees generally commented positively, specifying that the test has good design features. Specifically, several students believed that the test content is authentic and the TOEIC LR is a good measure of workplace English. Others stated that the listening test reflects real-world English language use because it includes various native accents. A few students reported that the TOEIC is a much better and preferred test for graduation requirements. Five students (19%) felt that the reading test is too long and that they had difficulty finishing the test in time.

The second major theme pertains to the students' purposes for taking the test. The interviewees reported that obtaining a satisfactory score on the TOEIC before graduation could simultaneously serve multiple purposes, such as qualifying for certain jobs, applying for graduate schools and scholarships, or simply assessing one's ELP. One sample quote follows (Student 3, high proficiency, general university, senior):

The TOEIC test is very popular among companies and organizations and it has a good reputation for its good discrimination, so I decided to give it a try for my graduation requirement. Also since I'm about to graduate, the test certificate is very helpful when I apply for jobs.

Eleven students added that the TOEIC test assesses not only workplace English but also language use in real-world contexts. These students were sensitive to what is assessed in the test and considered that the test could also measure their general ELP. Student 5 explained in detail (low proficiency, technical college, freshman):

I think the TOEIC test scores are very helpful and have practical utility because the test focuses on workplace English.. .. After I prepared for the TOEIC, I also realized that the TOEIC test content involves a lot of real-life situations and you have to pay special attention to all the "wh" questions, like why, how, what, etc. The test content is a good reflection of the language skills you need in daily life and it can measure your general English language proficiency as well.

The third theme that emerged from the data was related to students' perceptions about TOEIC score validity. Eighteen students made comments about their perceptions of score validity and were positive about the accuracy of the TOEIC LR scores they obtained, suggesting that the scores were perceived as very reliable and reflected their actual English language abilities. "I took the test two times, and my scores were very close both times. So I think the scores were very consistent and reflected my English abilities," said Student 19 (high-proficiency, general university, senior). When asked why he thought that the test score was an accurate reflection of his English skills, Student 3 commented:

I think the scores were quite accurate because when I was responding to the questions, I knew which items I answered correctly and which ones I did not know the answers. I mean, I know what my level is, and the scores reflected that.

Student 5 made a comment about score interpretations that helped explain the differing views on score validity seen from the study survey. She stated:

> I think speaking and writing abilities are different from reading and listening abilities. So, yes, I think the TOEIC listening and reading scores are good measures of my listening and reading abilities. But they can't reflect my speaking and writing abilities.. .. The university should not use listening and reading tests only to judge my overall English proficiency.

Two students did not regard the TOEIC LR scores as valid and felt that it is easy to achieve a high score on the TOEIC by intensive test-preparation training. The following quote exemplifies the key opinion expressed (Student 10, high proficiency, general university, senior):

> I think it's possible to get very high scores on the TOEIC if you just prepare for the test in a relatively short period of time. You know, preparing for the TOEIC test is very easy. All the items are multiple-choice items. The only thing is that there are many items on the test. If you can read faster and respond faster to the reading items, you can obtain high scores. So the scores cannot really reflect your English ability.

Although this comment about the possible effect of test preparation on score gains points to a concern of students with the use of the TOEIC as an exit test, the fact that the majority of the students believed that the TOEIC test scores have good reliability and validity suggests that the impact of test preparation might not be a serious problem within the context of this study. To better understand the role of test preparation and its impact, more empirical studies should be conducted to examine the nature of test-preparation activities and how these activities might, on the one hand, assist students in validly responding to test questions or, on the other hand, artificially inflate their test scores. Results of such investigations can help strengthen the validity argument for the TOEIC test, justify its use, and bring about more positive perceptions in the local context.

Conclusion and Implications

This study examined TOEIC test takers' perceptions about the English-language graduation requirement policy and the use of the TOEIC as an exit test within Taiwan's HE institutions. The survey results and the interview data collectively showed that the majority of the students were in favor of the policy and were positive about the use of the TOEIC as an exit test. Findings of the study suggest that students' levels of ELP are related to their perceptions about the policy and the appropriateness of the test use. The results also revealed that the cut scores set by some institutions might be too low to bring about positive washback, corroborating the findings of Shih (2007). In light of the results, we recommend that Taiwan's MOE policy makers and university decision makers periodically review the cut scores to ensure that the requirements are appropriate for their students and the intended purpose of the exam.

The survey respondents generally believed that preparing to take the TOEIC was helpful for improving their ELP and acknowledged that there could be added value in requiring students to take the TOEIC Speaking and Writing tests. To promote positive washback, future research should explore the social, educational, and economic impact of requiring the TOEIC Speaking and Writing tests and issues that could arise as a result of such a requirement. It is likely that additional teaching and learning resources will need to be provided to help students increase their speaking and writing proficiency in order to succeed.

Taken together, students' perceptions of the purposes of the TOEIC are in line with the intended uses of the TOEIC for preparing test takers to gain a competitive edge in the job market. The results of the study have provided an important piece of empirical evidence in support of the use of the TOEIC exam in Taiwan's HE context, and they have implications for the use of the TOEIC as a college exit test in other Asian contexts, such as Korea (Choi, 2008) and Vietnam (Nhan, 2013) and for exit tests in general (e.g., Berry & Lewkowicz, 2000; Spolsky, 1997). With the growing trend of using the TOEIC as a college gate-keeping test, more research that investigates this high-stakes test use across different contexts is called for. In light of the study results, future research will benefit from exploring individual learner factors (e.g., motivation, test anxiety, language proficiency) and educational backgrounds (e.g., academic major, year in school) and how these variables interact with students' perceptions about the use of the test scores for graduation within and across the testing contexts.

Findings of this study have implications for the creation and implementation of language testing policies. The study results suggest that test users, such as university staff or MOE policy makers who interpret and use language test scores, may not necessarily have a full understanding of what proficiency test cut scores mean and how to best use the information provided by language tests to make decisions. Given the high-stakes nature of their work, it is critical that these score users exercise their roles in an informed and ethical manner in the interest of valid test interpretation and use. To this end, the TOEIC program and the local partner could produce educational materials to help build the assessment literacy of test users. Specifically, the score users and language educators need to be much better informed about the TOEIC test, the test processes, and the principles and concepts that guide the test practices. An appropriate level of assessment literacy among score users can help mitigate the risk of misuse of test scores in making decisions for students (O'Loughlin, 2013).

In high-stakes testing situations such as college exit tests, there is also a need to inform students about the test design, intended use, and score interpretations. When students are well informed of the test practices, they can better establish a link between their learning goals and the assessment tasks, and positive washback can be promoted. Policymakers should keep in mind that students' perceptions about test use can play an important mediating role in policy implementation. As the study results revealed, students' beliefs may not always be congruent with the intentions underlying the language policy required of them. Future research should continue to examine test-taker perceptions of the impact of exit tests on language learning, focusing on individual students, their learning goals, and their understanding of test scores. Reasons for possible inconsistencies between the intended policy objectives and student perceptions should be identified, discussed, and adequately addressed to successfully achieve the intended purposes of a language testing policy.

References

Alderson, C., & Hamp-Lyons, L. (1996). TOEFL preparation courses: A study of washback. *Language Testing, 13*(3), 280–297.

Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Oxford, England: Oxford University Press.

Berry, V., & Lewkowicz, J. (2000). Exit-tests: Is there an alternative? *Hong Kong Journal of Applied Linguistics, 5*(1), 19–49.

Chen, T.-H. (2008). *Impacts of compulsory standardized exams on college students' L2 learning motivation* (Unpublished master's thesis). National Chiao Tung University, Hsinchu, Taiwan.

Cheng, L., Andrews, S., & Yu, Y. (2010). Impact and consequences of school-based assessment (SBA): Students' and parents' views of SBA in Hong Kong. *Language Testing, 28*(2), 221–249.

Choi, I.-C. (2008). The impact of EFL testing on EFL education in Korea. *Language Testing*, *25*(1), 39–62.

Chu, H.-Y. (2009). *Stakes, needs and washback: An investigation of the English benchmark policy for graduation and EFL education at two universities of technology in Taiwan* (Unpublished doctoral dissertation). National Taiwan Normal Unversity, Taipei, Taiwan.

Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge, England: Cambridge University Press.

Educational Testing Service. (2015). *TOEIC examinee handbook: Listening and reading*. Princeton, NJ: Author.

Hamp-Lyons, L. (1997). Washback, impact and validity: Ethical concerns. *Language Testing, 14*(3), 295–303.

Hamp-Lyons, L. (1998). Ethical test preparation practice: The case of the TOEFL. *TESOL Quarterly, 32*(2), 295–303.

Hsu, H.-F. (2009). *The impact of implementing English proficiency tests as a graduation requirement at Taiwanese universities of technology* (Unpublished doctoral dissertation). University of York, York, Canada.

Murray, J. C., Riazi, A. M., & Cross, J. L. (2012). Test candidates' attitudes and their relationship to demographic and experiential variables: The case of overseas trained teachers in NSW, Australia. *Language Testing, 29*(4), 577–595.

Nhan, T. (2013). The TOEIC® test as an exit requirement in universities and colleges in Danang City, Vietnam: Challenges and impacts. *International Journal of Innovative Interdisciplinary Research, 2*(1), 33–50.

Nichols, J. (2016). Do high-stakes English proficiency tests motivate Taiwanese university students to learn English? *American Journal of Educational Research, 4*(13), 927–930.

O'Loughlin, K. (2013). Developing the assessment literacy of university proficiency test users. *Language Testing*, *30*(3), 363–380.

Pan, Y.-C. (2014). Learner washback variability in standardized exit tests. *TESL-EJ, 18*(2), 1–30.

Pan, Y.-C., & Newfields, T. (2011). Teacher and student washback on test preparation evidenced from Taiwan's English certification exit requirements. *International Journal of Pedagogies and Learning, 6*(3), 260–272.

Pan, Y.-C., & Roever, C. (2016). Consequences of test use: A case study of employers' voice on the social impact of English certification exit requirements in Taiwan. *Language Testing in Asia, 6*(6), 1–22.

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.

Roever, C., & Pan, Y.-C. (2008). Test review: GEPT: General English Proficiency Test. *Language Testing, 25*(3), 403–408.

Schmidt, C. (2004). The analysis of semi-structured interviews. In U. Flick, E. v. Kardorff, & I. Steinke (Eds.), *A companion to qualitative research* (pp. 253–258). London, England: Sage.

Shih, C-M. (2007). A new washback model of students' learning. *The Canadian Modern Language Review, 64*(1): 135–162.

Shih, C.-M. (2009). How tests change teaching: A model for reference. *English Teaching: Practice and Critique, 8*(2), 188–206.

Shih, C.-M. (2010). The washback of the General English Proficiency Test on university policies: A Taiwan case study. *Language Assessment Quarterly, 7*, 234–254.

Shih, C.-M. (2012). Policy analysis of the English graduation benchmark in Taiwan. *Perspectives in Education, 30*(3), 60–68.

Shohamy, E., Donitsa-Schmidt, S., & Ferman, I. (1996). Test impact revised: Washback effect over time. *Language Testing, 13*(3), 298–317.

Spolsky, B. (1997). The ethics of gatekeeping tests: What have we learned in a hundred years? *Language Testing*, *14*(3), 242–247.

Tannenbaum, R. J., & Wylie, E. C. (2013). Mapping *TOEIC*® and *TOEIC Bridge*™ test scores to the Common European Framework of Reference. In D. E. Powers (Ed.), *The research foundation for the TOEIC tests: A compendium of studies: Volume II* (pp. 6.1–6.10). Princeton, NJ: Educational Testing Service.

Tasi, Y., & Tsou, C.-H. (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.

Vongpumivitch, V. (2006). *An impact study of Taiwan's General English Proficiency Test (GEPT)*. Paper presented at the annual meeting of the Language Testing Research Colloquium, Melbourne, Australia.

Wall, D. (2000). The impact of high-stakes testing on teaching and learning: Can this be predicted or controlled? *System, 28*, 499–509.

Wu, J. (2012). GEPT and English language teaching and testing in Taiwan. *Language Assessment Quarterly, 9*(1): 11–25.

Suggested Citation: