

Compendium Study

The Relationship Among *TOEIC*[®] Listening, Reading, Speaking, and Writing Skills

Jinghua Liu and Kate Costanzo

September 2013

The *TOEIC*® test is an English-language proficiency test for people whose native language is not English. It measures the everyday English skills of people working in an international environment. The scores indicate how well people can communicate in English in business, commerce, and industry. The test does not require specialized knowledge (i.e., it is not a subject test) or vocabulary beyond that of a person who uses English in everyday work activities.

The TOEIC test serves multiple purposes. TOEIC scores can be used for job recruitment, placement, and/or promotion within a corporation where everyday English at the workplace is a required job skill. In addition, TOEIC scores can be used as a measurement of everyday English proficiency levels of students in schools and as a measurement of an individual’s progress in English proficiency levels over time.

Test Configuration

The TOEIC test focuses on four essential English-language skills used in real life in the workplace: listening, reading, speaking, and writing. The *TOEIC*® Listening and Reading test is a paper-and-pencil-based test divided into listening comprehension and reading comprehension. The *TOEIC*® Speaking and Writing tests are computer-based tests administered separately from the TOEIC Listening and Reading test. Test takers can choose to take both the TOEIC Speaking and Writing tests or can take one without taking the other. More specific test configurations are presented in Table 1.

Table 1
TOEIC Test Configurations

	Listening	Reading	Speaking	Writing
Format	Paper-and-pencil test	Paper-and-pencil test	Computer-based test	Computer-based test
Questions	100 multiple-choice questions	100 multiple-choice questions	11 spoken and/or written prompts that require test takers to respond	8 written prompts that require test takers to respond
Administration	Administered in secure test centers or through institutions	Administered in secure test centers or through institutions	Administered via computer in secure test centers or through institutions	Administered via computer in secure test centers or through institutions
Scores	Score ranges from 5–495	Score ranges from 5–495	Score ranges from 0–200; score falls into one of the eight proficiency levels	Score ranges from 0–200; score falls into one of the nine proficiency levels

Listening

The listening section on the TOEIC Listening and Reading test measures how well a test taker understands spoken English. Test takers are asked to answer questions based on a variety of statements, questions, conversations, and talks recorded in English. The listening section consists of 100 multiple-choice questions. The testing time is about 45 minutes. Each test taker receives a score for listening on a scale from 5 to 495 points with increments of 5 points.

Reading

The reading section on the TOEIC Listening and Reading test measures how well a test taker understands written English. Test takers read a variety of materials and respond at their own pace. The reading section consists of 100 multiple-choice questions and takes approximately 75 minutes. Similar to the listening section, for the reading section each test taker receives a score on a scale from 5 to 495, with increments of 5 points. In addition, a total score of listening and reading is reported as well.

Speaking

The *TOEIC*® Speaking test is designed to measure a person's ability to communicate in spoken English in the context of daily life and of global workplace. The test is composed of 11 tasks and takes approximately 20 minutes to complete. The first two tasks require test takers to read a text aloud in an attempt to evaluate test takers' pronunciation, intonation, and stress. The third task shows a picture to test takers and requires them to describe the picture. Grammar, vocabulary, and cohesion are measured in addition to pronunciation, intonation, and stress. The next six tasks are integrated reading/listening/speaking tasks, which require test takers to read a short text, listen to a spoken text that pertains to the reading text, and then respond to questions. The remaining two tasks are listening/speaking tasks, requiring test takers to listen to a short spoken text and then propose a solution or express opinions.

Writing

The *TOEIC*® Writing test is designed to measure a person's ability to use written English to perform communication tasks that are typical of daily life and the global workplace. The test is composed of eight tasks and takes about one hour to complete. The first five tasks ask test takers to write a sentence based on a picture. The evaluation criteria include grammar and relevance of the sentences to the pictures. The next two tasks require test takers to read an e-mail message and respond to a written request. The quality and variety of the sentences, vocabulary, and organization are evaluated. For the last task, test takers write an essay in response to a question that asks them to state, explain, and support their opinions on an issue.

TOEIC Speaking scores and TOEIC Writing scores are reported separately. Each is reported on a scale of 0 to 200 in increments of 10. In addition, there are eight proficiency levels for the TOEIC Speaking test, and nine proficiency levels for the TOEIC Writing test. The proficiency levels correspond to a scaled score range and describe the types of general skills and proficiencies in spoken or written English that are common for most people who have similar scores.

Overall, the TOEIC test is designed to measure test takers' strengths and weaknesses in all four language skills and provide evidence of English proficiency that can be used to make informed decisions. The purpose of the current study was to examine the relationship among listening, reading, speaking, and writing skills measured by the TOEIC tests using empirical data and to explore the potential role of receptive skills (listening and reading) in the improvement of productive skills (speaking and writing).

In the next section of this paper, we review previous studies examining the relationship of different skills measured by different language tests. The methodology section describes empirical analysis procedures including correlation analyses and the improvement of speaking and writing based on different proficiency levels of listening and reading. The final section discusses the results.

The Relationship Among Different Language Skills: Previous Studies

Research Using *TOEFL*® Data

In a study examining the factor structure of the *TOEFL iBT*® test, Sawaki, Stricker, and Oranje (2008) reviewed research on whether language ability is unitary or divisible into components. They stated that “the current consensus in the field of language testing is that second language ability is multicomponential, with a general factor as well as smaller group factors” (p. 3). They also reviewed numerous studies that examine the structure of TOEFL (e.g., Hale et al., 1988; Hale, Rock, & Jirele, 1989; Manning, 1987; Stricker, Rock, & Lee, 2005; Swinton & Powers, 1980) and concluded that the multicomponential nature of the TOEFL test has been supported by these studies.

In their study, Sawaki et al. (2008) conducted an item-level confirmatory factor analysis (CFA) using a sample from a TOEFL iBT field trial. First, the study identified a single higher-order general factor: English as a second language (ESL)/English as foreign language (EFL) ability. Second, the study identified four first-order factors: reading, listening, speaking, and writing. The structure of the TOEFL iBT was best represented by a hierarchical factor structure that included the single higher-order general factor (ESL/EFL ability) and four first-order factors. The factor loading patterns of the integrated speaking and writing tasks indicate that these tasks well define the target constructs (speaking and writing) and are minimally involved in the reading and listening constructs. In general, the finding from the Sawaki et al. study is consistent with the consensus in the language assessment literature that language ability is multicomponential.

Research Using IELTS Data

While Sawaki et al. (2008) identified reading, listening, speaking, and writing as the four first-order factors of TOEFL iBT, Bozorgian (2012) examined the relationship among these four skills measured by the International English-Language Testing System (IELTS).

Data were collected from approximately 700 IELTS test takers. Correlation analysis revealed that the scores from the four skills were moderately correlated. The highest correlation was between TOEIC Listening and Reading scores ($r = .729$), and the lowest correlation was between TOEIC Reading/Speaking scores and TOEIC Writing/Speaking scores ($r = .498$). Further, it was found that TOEIC Listening scores had the highest correlation ($r = .893$) with the overall language proficiency (measured by the average of the four scores), followed by TOEIC Reading/Writing scores ($r = .792$). The Speaking scores had relatively the lowest correlation with the overall language proficiency ($r = .756$).

The author concluded that each of the four language skills contributes to second or foreign language learning. It is crucial to take all four skills into account in second language learning and evaluation.

Research Using TOEIC Test Data

To collect evidence to validate the meaning of the TOEIC scores, Powers and his colleagues designed and administered self-assessment surveys to TOEIC test takers in Japan and Korea that gathered perceptions of their abilities to perform a variety of everyday English-language tasks (Powers, Kim, & Weng, 2008; Powers, Kim, Yu, Weng, & VanWinkle, 2009).

Powers et al. (2008) assembled a survey that included a variety of can-do statements related to listening and reading tasks, ranging from easy tasks such as "I can understand the days of the week and the months of the year" to more complicated tasks such as "I can understand a complex presentation or demonstration in an academic or work-related setting." Test takers were asked to respond to each statement using a 5-point scale, with 1 (*not at all*) and 5 (*easily*). Test takers' self-assessments of their abilities to perform the can-do tasks were defined by the sum of responses to (a) all listening can-do tasks and (b) all reading can-do tasks. Correlations between TOEIC Listening and Reading scores and test takers' self-assessments of their abilities were analyzed. Later, another survey focusing on speaking and writing can-do statements was assembled and administered to test takers. Correlations were computed between TOEIC Speaking and Writing scores and test takers' self-assessments of their ability to perform speaking can-do tasks and writing can-do tasks (Powers et al., 2009).

In addition to the main finding that test takers at higher TOEIC score levels were more likely to report that they could successfully accomplish each of the everyday language tasks in English, Powers and his colleagues also found (a) modest discriminant validity of the listening and reading components of the TOEIC test, suggesting that each section contributes to the measurement of English-language skills (Powers et al., 2008) and (b) modest discriminant validity of the TOEIC Speaking and Writing

measures, suggesting that each measure contributes uniquely to the assessment of English-language proficiency (Powers et al., 2009). The results were consistent with those in the Sawaki et al. (2008) factor analysis study, suggesting that listening and reading as well as speaking and writing are related but distinct factors.

One limitation of the Powers et al. (2008, 2009) studies is that the data were collected separately for listening/reading and speaking/writing. Since the TOEIC Speaking and Writing tests were introduced in December 2006, the number of test takers who took these tests by summer 2007, when the first survey was administered, was limited. Hence, the first survey included only listening- and reading-related can-do statements. The second survey included speaking and writing can-do statements and was administered in fall 2008. Therefore, the correlation was computed between listening and reading and between speaking and writing, but not across listening/reading and speaking/writing.¹

The current study examined the relationship among all four skills measured by the TOEIC tests: listening, reading, speaking, and writing. Further, the relationship of listening and reading proficiency to score improvement of speaking and writing is explored.

Methodology

Data

The dataset included 4,935 test takers from Korea who took (a) the TOEIC Speaking and/or Writing tests multiple times from December 2006 to December 2008 and (b) the TOEIC Listening and Reading test from December 2006 to November 2008. In this dataset, each test taker had one record for a TOEIC Listening test scaled score and one record for a TOEIC Reading test scaled score but multiple records for TOEIC Speaking and/or TOEIC Writing tests scaled scores. The TOEIC Listening and Reading test scaled scores were the most recent ones prior to test takers' first TOEIC Speaking and/or TOEIC Writing scores collected for the current study. In addition, each TOEIC Speaking and TOEIC Writing test score is associated with the corresponding proficiency levels (1–8 for the TOEIC Speaking test, and 1–9 for the TOEIC Writing test).

¹While the current study was being completed, a four-skill survey that included listening, reading, speaking, and writing can-do statements was being conducted. The results are reported elsewhere (Powers, Yu, & Yan, 2013).

Table 2 presents the descriptive statistics for the four sets of scores. As can be seen in Table 2, this group of test takers had mean scores of 421.2 and 385.9 on the listening and reading sections, respectively, on the TOEIC Listening and Reading test. The mean TOEIC Speaking score was 138.6, whereas the mean score of the TOEIC Writing test was 154. Note that the summary statistics for TOEIC Speaking and Writing tests are based on the first-time scores.

Table 2

Descriptive Statistics of the Dataset

Statistic	TOEIC Listening	TOEIC Reading	TOEIC Speaking		TOEIC Writing	
			Score	Proficiency	Score	Proficiency
Mean	421.2	385.9	138.6	5.91	154.0	7.09
SD	65.3	69.5	27.2	1.04	28.5	1.05
Min	10	10	10	1	0	1
Max	495	495	200	8	200	9

Note. Summary statistics for the TOEIC Speaking and Writing tests are based on first-time scores. *N* = 4,935.

Statistical Analyses

Correlations among TOEIC Listening, Reading, Speaking, and Writing scores. The correlation between each pair of scores is examined using Pearson correlation coefficients. Further, since the total TOEIC Listening and Reading score is often used to evaluate test takers’ English-language proficiency, we also looked into the relationship between the TOEIC Speaking score and the total TOEIC Listening and Reading score, and the TOEIC Writing score and the total TOEIC Listening and Reading score.

Two sets of correlations are reported—observed score correlations and true score correlations. Observed score correlations are the correlations between the scores obtained on the different tests by individual test takers. True score correlations, or disattenuated correlations, are estimates of the correlations between scores that would have been obtained if the tests were perfectly reliable. True score correlations are estimated statistically from the observed score correlations using estimates of test reliability.

Criterion to evaluate correlations. Once we obtained the correlations between two test scores, we needed to evaluate the magnitude of the correlations in order to determine whether or not the two tests measure the same or similar construct. Dorans (2000, 2004) defined an index called reduction in uncertainty (RiU) to measure the statistical certainty that two variables differ. Let r represent the correlation coefficient between the two variables, then RiU is defined as

$$RiU = 1 - \sqrt{1 - r^2}. \quad (1)$$

When $r = 0$, there is 0% reduction; when $r = 1$, there is 100% reduction. Where should the threshold be for a predictor to serve as a valid surrogate for the variable being predicted? Dorans suggested that for test score linkage in high-stakes settings, a correlation coefficient of at least .866 between the predictor and the score to be predicted is needed to reduce the uncertainty by at least 50%. If a predictor cannot reduce uncertainty by at least 50%, it is unlikely that the predictor can serve as a valid surrogate for the score being predicted. We adopt the notion of RiU for the current study and use a correlation coefficient of .866 as the threshold to determine whether the correlation between two test score vectors is high enough.

The relationship of TOEIC Listening and Reading test proficiency to TOEIC Speaking and Writing test score improvement. We also examined TOEIC Speaking and Writing score change based on different TOEIC Listening and Reading score levels. Table 3 is a hypothetical example for illustration purposes only.

Table 3

Illustration of TOEIC Listening and Reading (L&R) Proficiency Impact on TOEIC Speaking (S) Score Improvement

Test taker	1st S score	2nd S score	Score difference	L&R	Average score change
1	100	90	-10	450	
2	100	110	10	450	
3	100	120	20	450	6.7
4	100	100	0	710	
5	100	130	30	710	
6	100	120	20	710	16.7

Note. Score changes are averaged over Test Takers 1–3 and Test Takers 4–6.

The procedure is as follows. First of all, tests takers are divided into 20 groups based on their first-time TOEIC Speaking scores (10, 20, 30, . . . , 180, 190, 200). In this example, these six test takers are in the same group because their first-time TOEIC Speaking score is 100. Second, we calculate the TOEIC Speaking score difference for each test taker after the second time they take the test, where

$score\ difference = 2nd\ score - 1st\ score$. For example, Test Taker 1 has a score difference of -10; Test Taker 2 has a score difference of 10, while Test Taker 3 has a score difference of 20. Third, for each test taker, the score difference is paired with his/her total TOEIC Listening and Reading score. Within each TOEIC Speaking score level (100 in this example), all the repeaters at the same total score of the TOEIC Listening and Reading test are grouped together, and an average of their TOEIC Speaking score change is calculated. For example, Test Takers 1, 2, and 3 all have a TOEIC Listening and Reading test total score of 450. So the average TOEIC Speaking score change at the TOEIC Listening and Reading test total score level of 450 is calculated across these three test takers. The average TOEIC Speaking score change is 6.7. In other words, for those test takers with their first-time TOEIC Speaking score of 100 and TOEIC Listening and Reading test total score of 450, the average TOEIC Speaking score change is 6.7 when they take the TOEIC Speaking test the second time. Similarly, for test takers with first-time TOEIC Speaking score of 100 and TOEIC Listening and Reading test total score of 710, the average score change is 16.7 when they take the TOEIC Speaking test the second time.

Results

Correlation and Reduction in Uncertainty (RiU)

Tables 4 and 5 contain the Pearson correlation coefficients for the six pairs of test scores. The four sets of test scores are all moderately correlated. For observed score correlations, the highest correlation is between TOEIC Listening and Reading scores, .726, with RiU = 31%, which reduces the uncertainty way below 50%; the lowest correlation is between TOEIC Listening and Writing scores, and between TOEIC Reading and Speaking scores, around .535, which only reduces the uncertainty approximately 16%. TOEIC Listening scores exhibit a higher correlation with TOEIC Speaking scores ($r = .634$, RiU = 23%) than with TOEIC Writing scores, while TOEIC Reading scores show slightly higher correlation with TOEIC Writing scores ($r = .564$, RiU = 17%) than with TOEIC Speaking scores. These correlations are very similar to what was found on other language tests such as the TOEFL iBT® test (TOEFL iBT Statistical Analysis Team, 2011) and the IELTS test (Bozorgian, 2012).

Table 4

Correlation Among TOEIC Listening (L), Reading (R), Speaking (S), and Writing (W) Scores

Correlation	L	R	S	W
L	1.000	0.789	0.739	0.612
R	0.726	1.000	0.625	0.645
S	0.634	0.537	1.000	0.727
W	0.535	0.564	0.592	1.000

Table 5***Reduction in Uncertainty (RiU) Among TOEIC Listening (L), Reading (R), Speaking (S), and Writing (W) Scores***

	<i>r</i>	RiU
L-R	0.726	31.3%
L-S	0.634	22.7%
L-W	0.535	15.5%
R-S	0.537	15.6%
R-W	0.564	17.4%
S-W	0.592	19.4%

The true score correlations are shown above the diagonal in Table 4. The estimated internal consistency reliabilities for both the listening section and the reading section on the TOEIC Listening and Reading test are .92; for the TOEIC Speaking test and the TOEIC Writing test, the estimated reliabilities are .80 and .83, respectively (Liao, Qu, & Morgan, 2010). The pattern is quite consistent with that found for observed scores. The reliabilities are comparable to those from the TOEFL iBT test where the reliabilities for listening and reading are around .90 and the reliabilities for speaking and writing range approximately from .80 to .85 (TOEFL iBT Statistical Analysis Team, 2011).

Overall, none of the correlations, either observed score correlations or true score correlations after adjusting for measurement errors, approach the threshold of .866. The highest degree of reduction in uncertainty is only 31%, much less than the 50%. In other words, although the four test scores were correlated, none of the correlations is high enough to suggest that any test can serve as a proxy for another. Each skill measures a distinct English-language proficiency component, and none of them can be a valid predictor for the other.

The observed score (and true score) correlations between the total TOEIC Listening and Reading score with the TOEIC Speaking and Writing scores are .628 (RiU = 22%) and .592 (RiU = 19%), respectively. Again, these moderate correlations and the small magnitude of RiU indicate that the TOEIC Listening and Reading test and the TOEIC Speaking and Writing tests measure related but distinct constructs.

Impact of Listening and Reading Proficiency on TOEIC Speaking and/or Writing Score Improvement

As described above, the analysis of relations of listening and reading proficiency to TOEIC Speaking and/or Writing score improvement was conducted, conditioned on each total TOEIC Listening and Reading score level. Since sample sizes were quite small at some score levels, we present only the results for total sample sizes larger than or equal to 100 at each TOEIC Speaking or Writing score level. Hence, for the TOEIC Speaking test, we included only the groups whose first-time TOEIC Speaking scores ranged from 70 to 180; for the TOEIC Writing test, we included the groups whose first-time TOEIC Writing scores ranged from 110 to 180. In addition, in order to present the results in a more concise way, we grouped the test takers into three categories based on their TOEIC Listening and Reading test total scores: low (10 ~ 400), medium (405 ~ 700), and high (705 ~ 990). Note that the categorization is for the convenience of presenting the results only.

Table 6 summarizes the average TOEIC Speaking score improvement for each of three listening and reading proficiency categories. To be concise, we present only every other score level. As can be seen from Table 6, for test takers whose first-time TOEIC Speaking score was 70 and whose TOEIC Listening and Reading test total scores were between 5 ~ 400, the average TOEIC Speaking score improvement was 12.9. If TOEIC Listening and Reading test total scores ranged between 405 ~ 700, average TOEIC Speaking score improvement increased by 22.5 points. The average TOEIC Speaking score improvement reached the highest level, 38.1, for those test takers whose TOEIC Listening and Reading test total scores were between 705 ~ 990.

Table 6

Average Speaking Score Improvement Conditioned on TOEIC Listening and Reading (L&R) Proficiencies

1 st TOEIC Speaking score	L&R		
	10 ~ 400	405 ~ 700	705 ~ 990
70	12.9	22.5	38.1
90	-2.1	13.1	24.1
110	-4.0	4.3	15.8
130	N/A ^a	-5.5	5.0
150	N/A ^a	-15.6	-1.6
170	N/A ^a	-31.3	-9.6

^aThe sample size was smaller than 5.

Figure 1 depicts the average TOEIC Speaking score change for this group of test takers along the entire TOEIC Listening and Reading test total score range. The X axis is the total score on the TOEIC Listening and Reading test, ranging from 0 to 990. The Y axis is the score difference between the second-time score and the first-time score. Each dot represents the average score change, conditioned on TOEIC Listening and Reading test total scores.

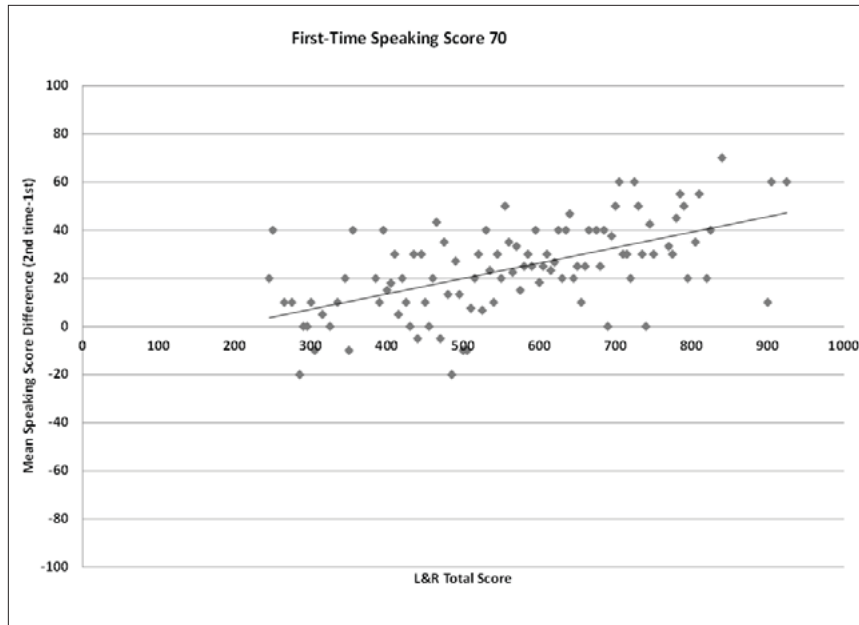


Figure 1. Mean score difference conditioned on the total score on the TOEIC Listening and Reading test (L&R): first TOEIC Speaking score = 70.

For this group of test takers, if their TOEIC Listening and Reading test total scores were below 500, most of their second-time TOEIC Speaking scores increased, whereas a few had score decreases. If scores were above 500, the second-time TOEIC Speaking score was more likely to increase than for those whose TOEIC Listening and Reading test total scores were below 500. The magnitude of the score increase in general was larger as TOEIC Listening and Reading scores increased. Although there were variations, the trend was clear: the higher the test takers' TOEIC Listening and Reading scores, the more likely test takers were to have larger score gains upon retesting.

The same trend can be observed at other TOEIC Speaking score levels. For the group of test takers whose first-time TOEIC Speaking score was 90, for instance, TOEIC Speaking score improvement was -2.1, 13.1 and 24.1, respectively, at the low, medium, and high listening and reading proficiency levels. At a given TOEIC Speaking score level, the higher the listening and reading proficiency was, the bigger the TOEIC Speaking score improvements were. Within the same listening and reading proficiency level, the higher the initial TOEIC Speaking scores were, the smaller the improvement was. The latter is probably an artifact due to regression to the mean.

It is interesting that even at the higher initial TOEIC Speaking score levels where there were more decreases than increases, the trend is for the increase to be greater and the decrease to be smaller. For example, at TOEIC Speaking score levels of 150 and 170, the second scores decreased for each listening and reading proficiency group. At score level 170, the increase of score change from the medium to high was 21.7 (from -31.3 to -9.6), whereas the decrease of score change from 150 to 170 for the high group was only 8 (from -1.6 to -9.6). See Table 6.

The association of listening and reading proficiency with TOEIC Writing score improvement is presented in Table 7. The results exhibit a similar pattern: at the same initial TOEIC Writing score level, the higher the listening and reading proficiency level, the higher the score improvement; at the same listening and reading proficiency level, the higher the initial TOEIC Writing score, the lower the score improvement.

Table 7
Average Writing Score Improvement Conditioned on Listening and Reading (L&R) Proficiencies

1 st TOEIC Writing score	L&R		
	10 ~ 400 ^a	405 ~ 700	705 ~ 990
110	N/A	14.7	27.0
130	N/A	7.2	17.0
150	N/A	-2.3	9.0
170	N/A	-20	0.25

^a The sample size was smaller than 5.

Figure 2 presents the results for test takers whose first-time TOEIC Writing score was 110 along the entire TOEIC Listening and Reading score range. Figure 2 exhibits a similar pattern to that shown in Figure 1: The more proficient the test takers are on the TOEIC Listening and Reading tests, the more likely they are to obtain a larger score increase the second time they take the TOEIC Writing test.

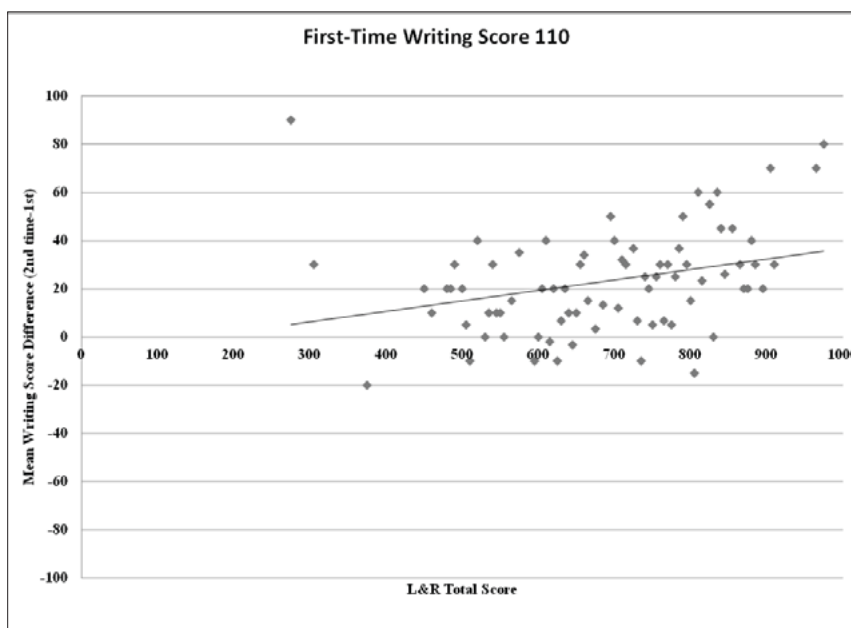


Figure 2. Mean score difference conditioned on the total score on the TOEIC Listening and Reading test (L&R): first TOEIC Writing (W) score = 110.

A complete set of results are plotted in the appendix. Even though there were variations, the overall trend is consistent across both the TOEIC Speaking and Writing tests: At each first-time TOEIC Speaking or Writing score level, the more proficient the test takers are on listening and reading tasks, the more likely that their second-time TOEIC Speaking and/or Writing scores are to increase.

Discussion

The TOEIC tests measure the everyday workplace English skills of people who work in an international environment. They focus on four essential English-language skills used in real life in the workplace: listening, reading, speaking, and writing. The purpose of this study was to examine the relationship among the four skills using empirical data and to examine the potential impact of listening and reading proficiencies on the improvement of speaking and writing skills.

The results suggest that the four skills are different. They are moderately correlated, but the correlation is not sufficient for one skill to serve as a valid surrogate for another skill. The results are consistent with research on other English-language tests such as the TOEFL and IELTS tests. Listening, reading, speaking, and writing each measure distinct aspects of English-language proficiency and each of them makes a unique contribution to the measurement of English-language abilities. Performing well on one test does not necessarily guarantee performing similarly well on the other TOEIC tests. Test takers are encouraged, therefore, to take all four skill tests to gain a more comprehensive understanding of their English-language abilities.


Further, it was found that scores for the TOEIC Listening and Reading test have the highest correlation of any pair of scores, followed by TOEIC Listening and Speaking scores. TOEIC Speaking and Writing scores have the third-highest correlation. TOEIC Listening and Writing scores and TOEIC Reading and Speaking scores, on the other hand, have weaker correlations. This may suggest that in the acquisition of a foreign language, listening is fundamental and is integrated with all other components such as speaking and reading. Some previous studies using TOEFL data all found a distinct listening comprehension factor, despite the differences on other factors (Hale et al., 1988; Manning, 1987; Sawaki et al., 2008; Swinton & Powers, 1980).

We also found that the receptive skills such as listening and reading might be fundamental to the improvement of productive skills (e.g., speaking and writing). Conversely, it is interesting that, using IELTS data, Bozorgian (2012) found that the development of productive skills may be required to foster the growth of receptive skills from the onset of learning. Unfortunately, although our dataset includes multiple TOEIC Speaking and/or TOEIC Writing scores, it contains only one TOEIC Listening/Reading score. Future studies can also explore the impact of speaking and writing proficiency on the improvement of TOEIC Listening and Reading scores.

In conclusion, this study provides evidence that the listening, reading, speaking, and writing components of the TOEIC tests measure related but distinct English-language proficiencies. Evaluating all four skills may therefore provide a reasonably complete picture of English communication competency, including strengths and weaknesses. Furthermore, receptive skills may be instrumental in helping to improve productive skills through learning and training.

References

- Bozorgian, H. (2012). Listening skill requires a further look into second/foreign language learning. *ISRN Education, Vol. 2012*, Article ID 810129doi:10.5402/2012/810129
- Dorans, N. J. (2000). *Distinctions among classes of linkages* (College Board Research Note RN-11). New York, NY: College Board.
- Dorans, N. J. (2004). Equating, concordance and expectation. *Applied Psychological Measurement*, 28(4), 227–246.
- Hale, G. A., Rock, D. A., & Jirele, T. (1989). *Confirmatory factor analysis of the Test of English as a Foreign Language* (TOEFL Research Report No. 32). Princeton, NJ: Educational Testing Service.
- Hale, G. A., Stansfield, C. W., Rock, D. A., Hicks, M. M., Butler, F. A., & Oller, J. W., Jr. (1988). *Multiple-choice Cloze items and the Test of English as a Foreign Language* (TOEFL Research Report No. 26). Princeton, NJ: Educational Testing Service.
- Liao, C., Qu, Y., & Morgan, R. (2010). *The relationship of test scores measured by the TOEIC Listening and Reading and TOEIC Speaking and Writing tests*. Unpublished manuscript.
- Manning, W. H. (1987). *Development of Cloze-elide tests of English as a second language* (TOEFL Research Report No. RR-87-18). Princeton, NJ: Educational Testing Service.
- Powers, D. E., Kim, H., & Weng, V. Z. (2008). *The redesigned TOEIC® (listening and reading) test: Relations to test-taker perceptions of proficiency in English* (Research Report No. RR-08-56). Princeton, NJ: Educational Testing Service.
- Powers, D. E., Kim, H., Yu, F., Weng, V. Z., & VanWinkle W. (2009). *The TOEIC® Speaking and Writing tests: Relations to test-taker perceptions of proficiency in English* (Research Report No. RR-09-18). Princeton, NJ: Educational Testing Service.
- Powers, D. E., Yu, F., & Yan, F. (2013). The TOEIC® Listening, Reading, Speaking, and Writing tests: Evaluating their unique contribution to assessing English-language proficiency. In D. Powers (Ed.), *The research foundation for the TOEIC tests: A compendium of studies* (Vol. II, 3.1-3.14). Princeton, NJ: Educational Testing Service.
- Sawaki, Y., Stricker, L., & Oranje, A. (2008). *Factor structure of the TOEFL Internet-based test (iBT): Exploration in a field trial sample* (Research Report No. RR-08-09). Princeton, NJ: Educational Testing Service.
- Stricker, L. J., Rock, D. A., & Lee, Y. W. (2005). *Factor structure of the LanguEdge test across language groups* (TOEFL Monograph Series No. MS-32). Princeton, NJ: Educational Testing Service.



Swinton, S. S., & Powers, D. E. (1980). *Factor analysis of the Test of English as a Foreign Language for several language groups* (TOEFL Research Report No. 6). Princeton, NJ: Educational Testing Service.

TOEFL iBT Statistical Analysis Team. (2011). *Test of English as a foreign language Internet-based test TOEFL iBT: Test analysis* (Statistical Report No. SR-11-030). Princeton, NJ: Educational Testing Service.

Appendix

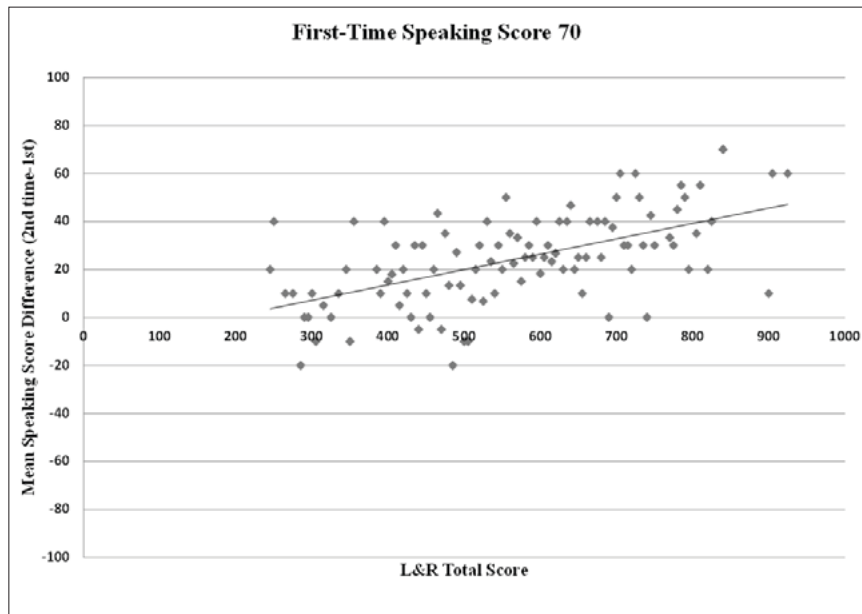


Figure A1. Mean score difference conditioned on the total score of the TOEIC Listening and Reading test (L&R): first TOEIC Speaking (SPK) score = 70.

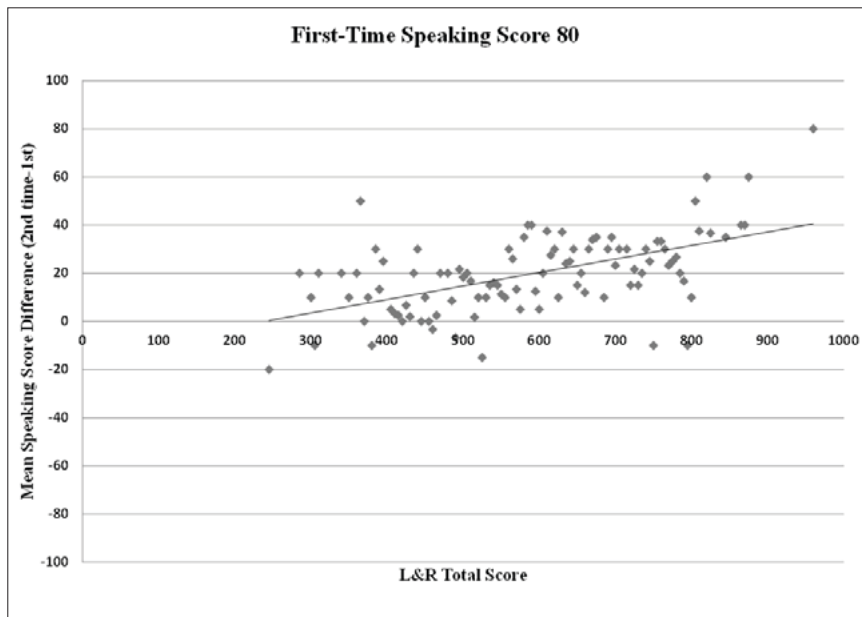


Figure A2. Mean score difference conditioned on the total score of the TOEIC Listening and Reading test (L&R): first TOEIC Speaking (SPK) score = 80.

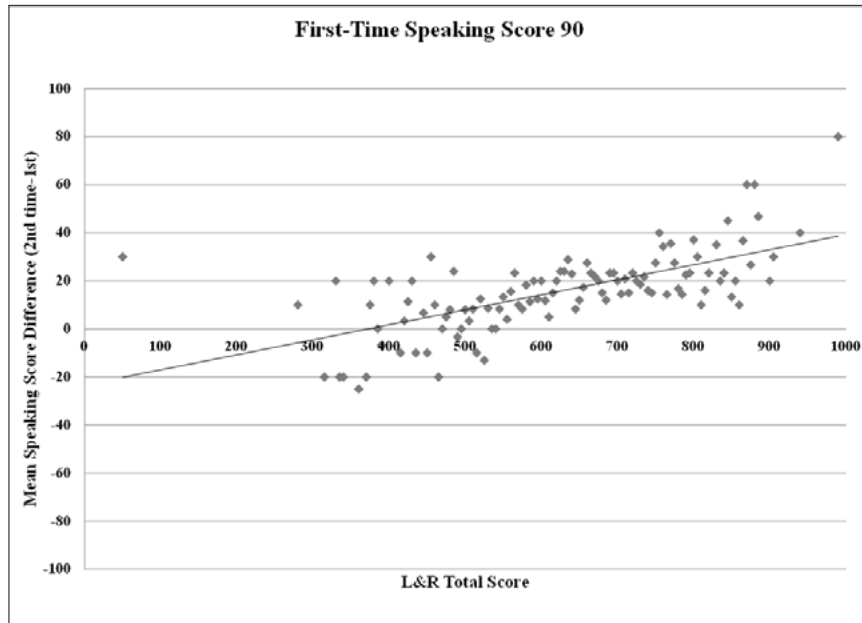


Figure A3. Mean score difference conditioned on the total score of the TOEIC Listening and Reading test (L&R): first the TOEIC Speaking (SPK) score = 90.

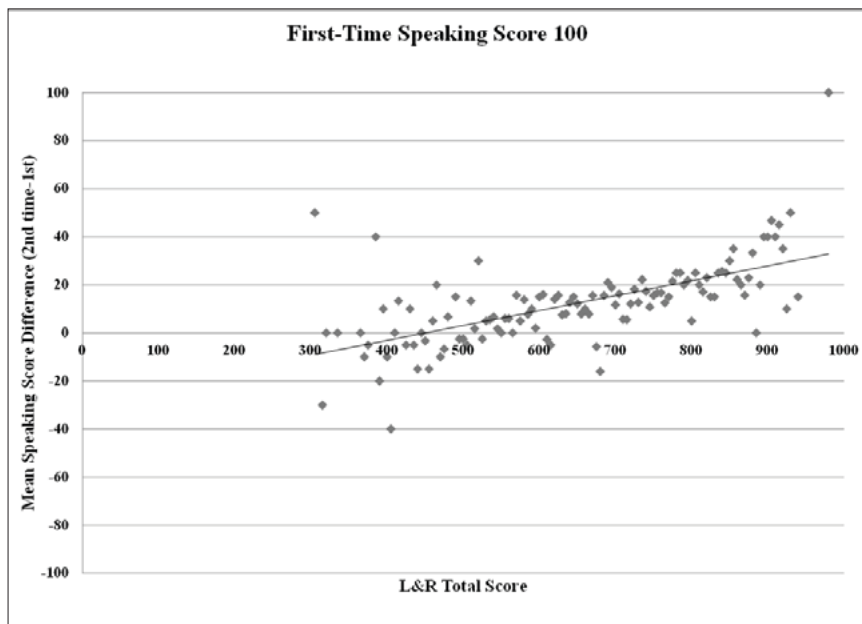


Figure A4. Mean score difference conditioned on the total score of the TOEIC Listening and Reading test (L&R): first the TOEIC Speaking (SPK) score = 100.

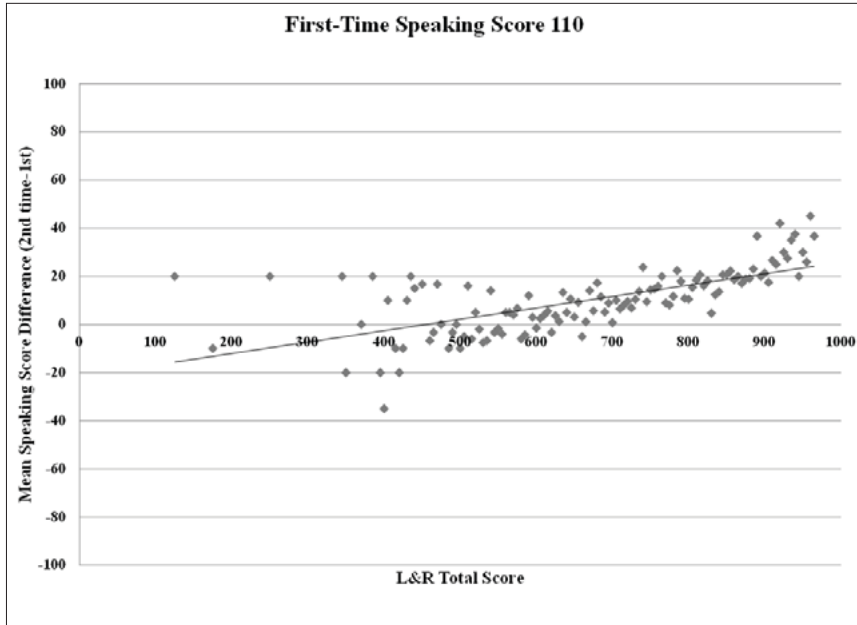


Figure A5. Mean score difference conditioned on the total score of the TOEIC Listening and Reading test (L&R): first TOEIC Speaking (SPK) score = 110.

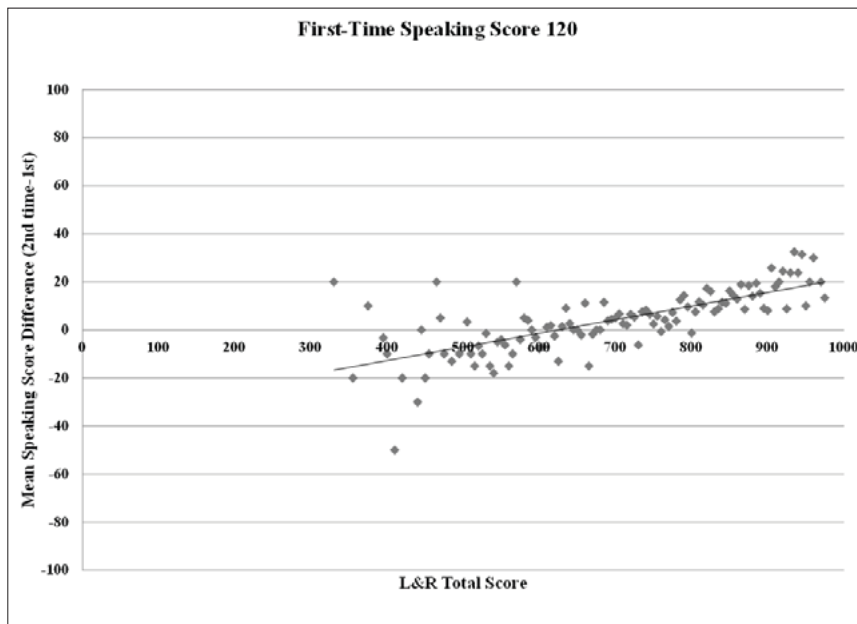


Figure A6. Mean score difference conditioned on the total score of the TOEIC Listening and Reading test (L&R): first TOEIC Speaking (SPK) score = 120.

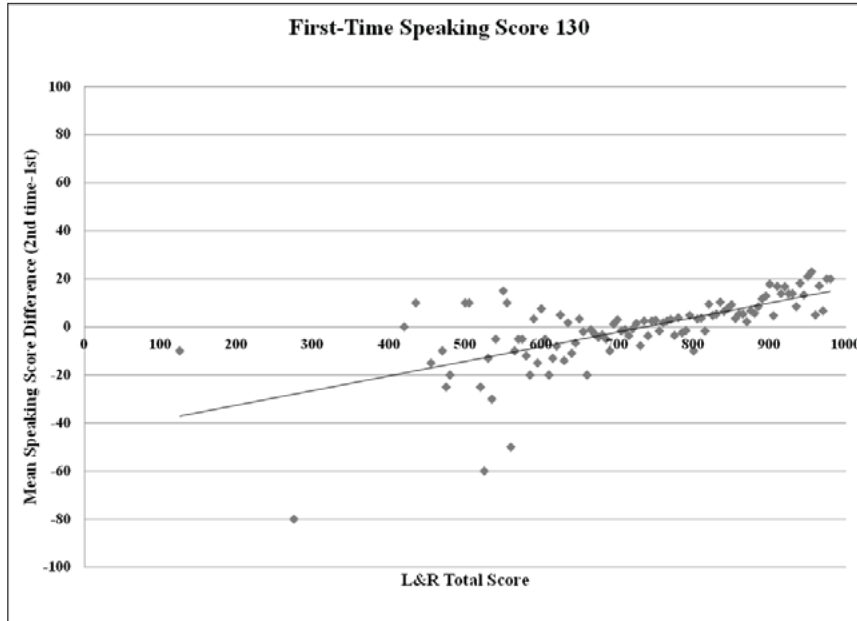


Figure A7. Mean score difference conditioned on the total score of the TOEIC Listening and Reading test (L&R): first TOEIC Speaking (SPK) score = 130.

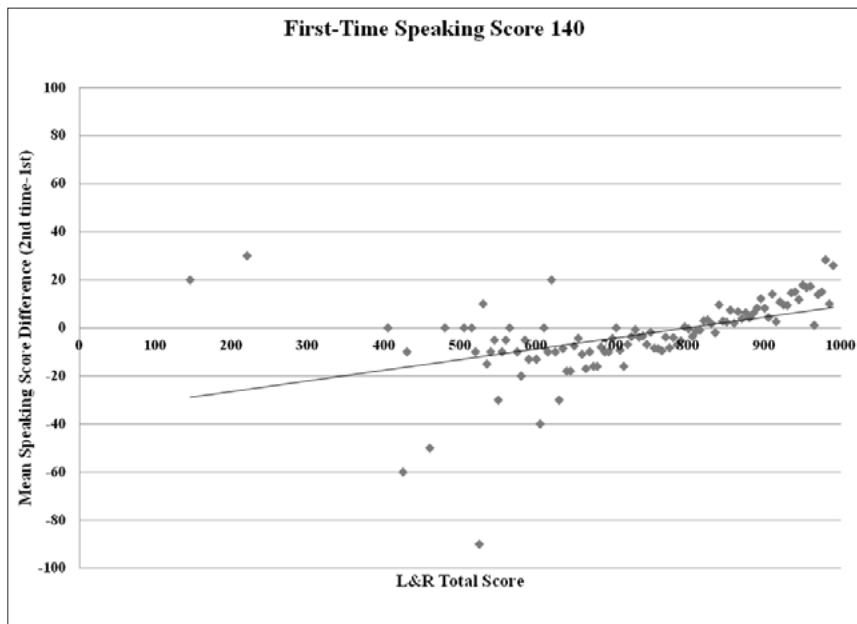


Figure A8. Mean score difference conditioned on the total score of the TOEIC Listening and Reading test (L&R): first TOEIC Speaking (SPK) score = 140.

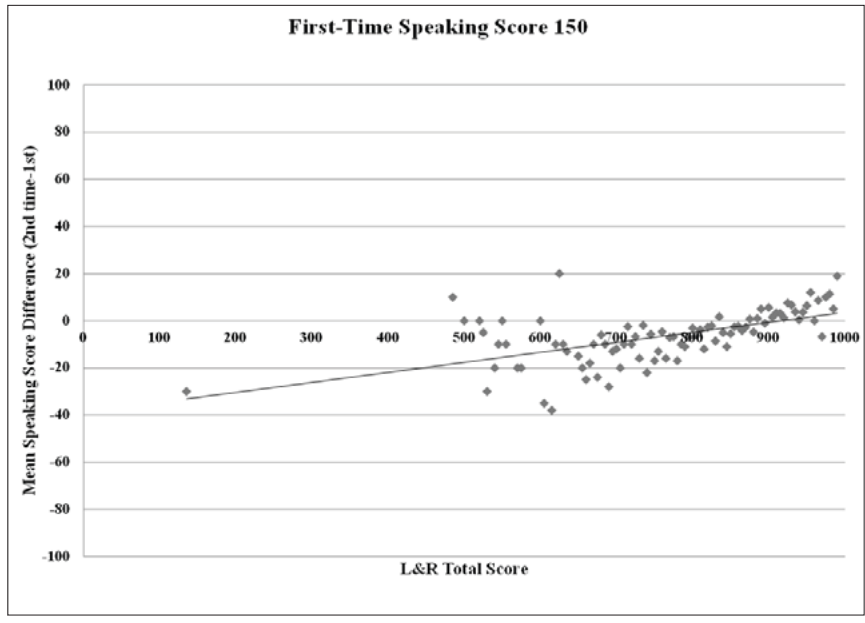


Figure A9. Mean score difference conditioned on the total score of the TOEIC Listening and Reading test (L&R): first TOEIC Speaking (SPK) score = 150.

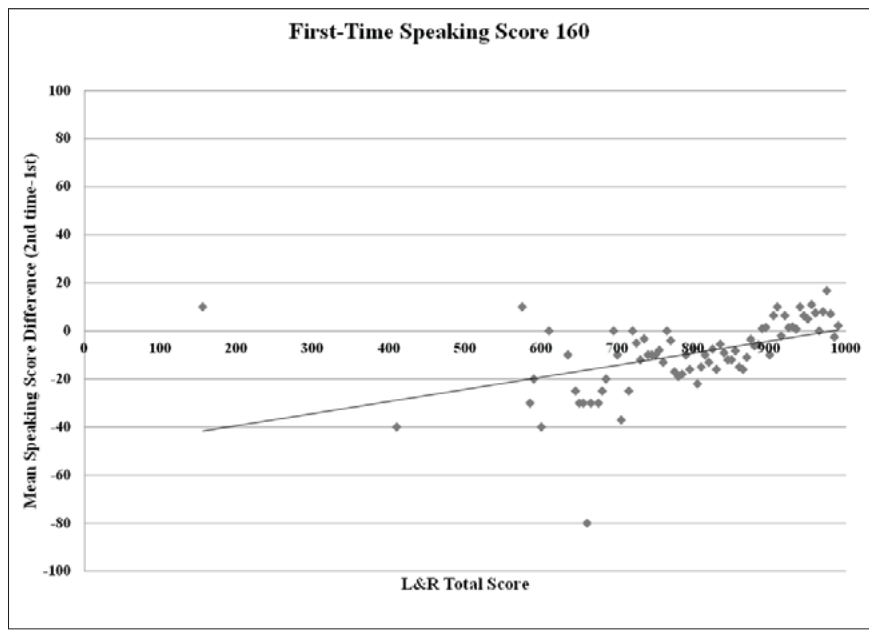


Figure A10. Mean score difference conditioned on the total score of the TOEIC Listening and Reading test (L&R): first TOEIC Speaking (SPK) score = 160.

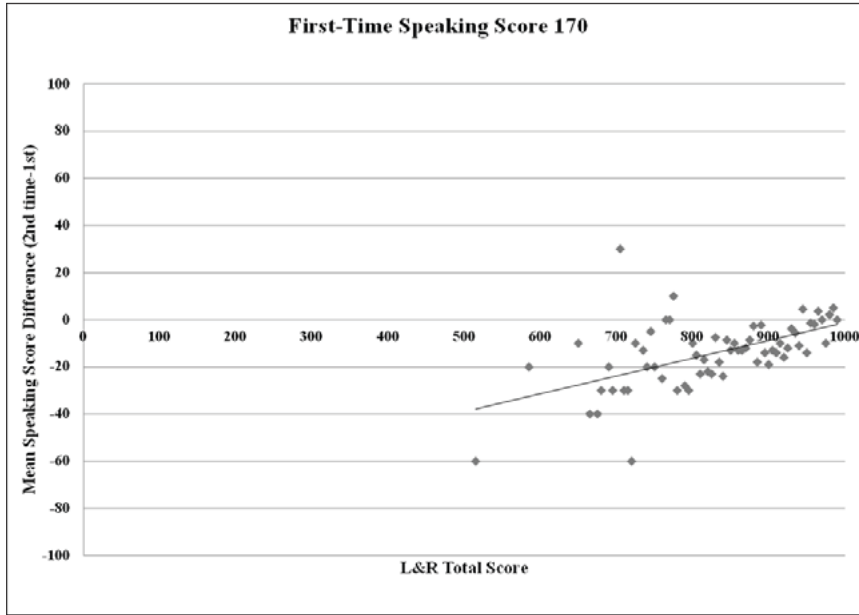


Figure A11. Mean score difference conditioned on the total score of the TOEIC Listening and Reading test (L&R): first TOEIC Speaking (SPK) score = 170.

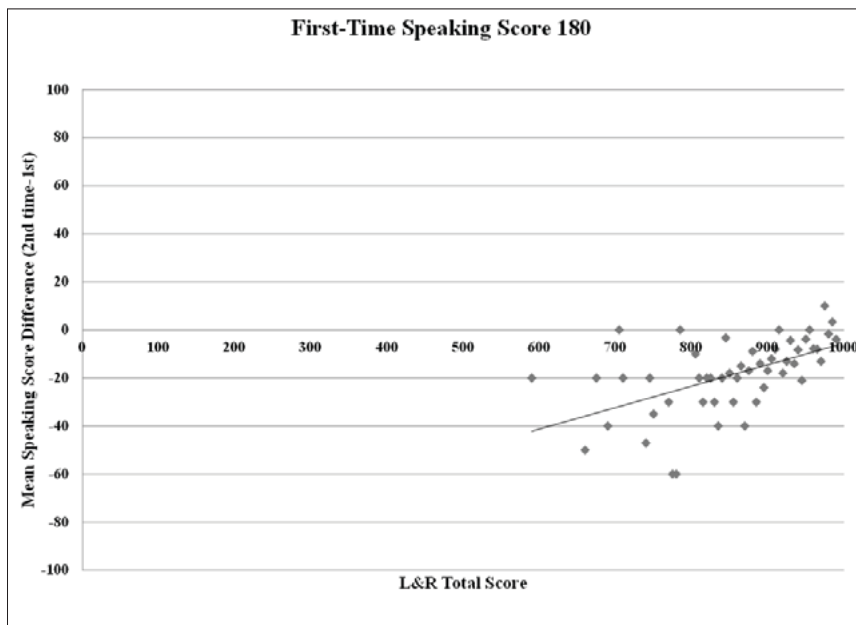


Figure A12. Mean score difference conditioned on the total score of the TOEIC Listening and Reading test (L&R): first TOEIC Speaking (SPK) score = 180.

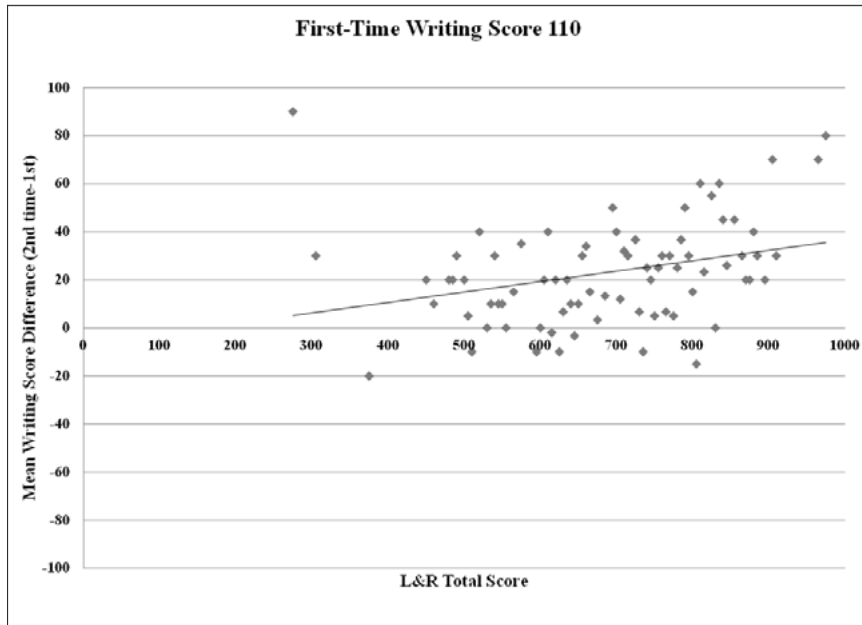


Figure A13. Mean score difference conditioned on the total score of the TOEIC Listening and Reading test (L&R): first TOEIC Writing (W) score = 110.

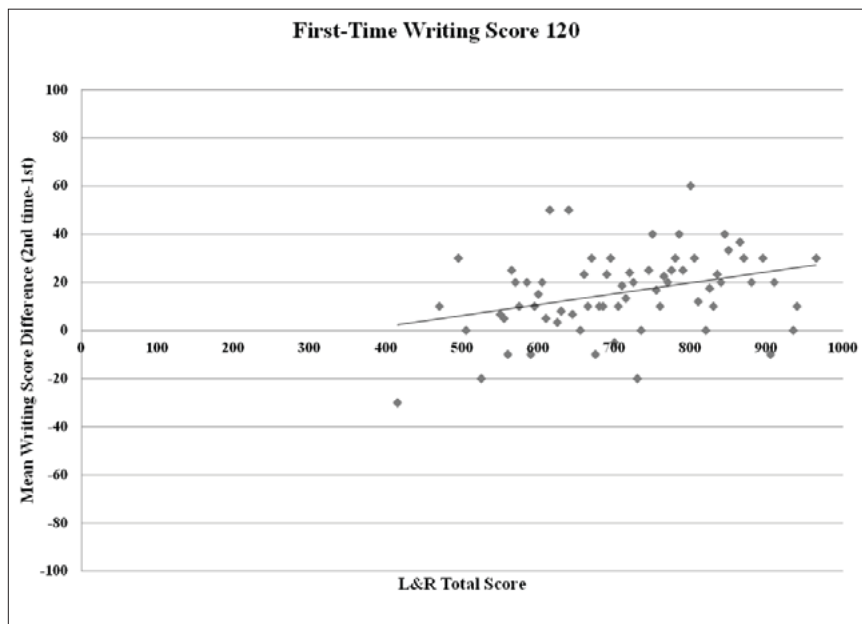


Figure A14. Mean score difference conditioned on the total score of the TOEIC Listening and Reading test (L&R): first TOEIC Writing (W) score = 120.

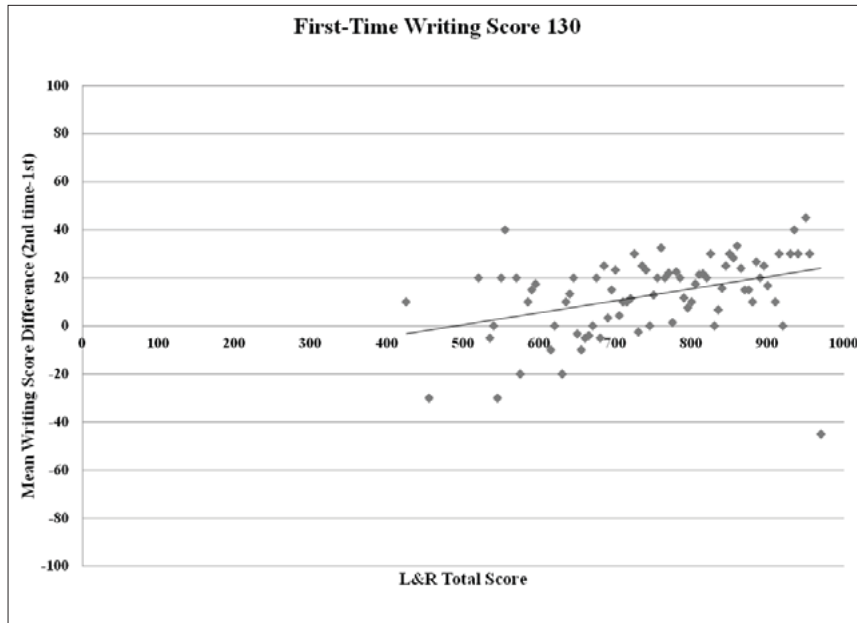


Figure A15. Mean score difference conditioned on the total score of the TOEIC Listening and Reading test (L&R): first TOEIC Writing (W) score = 130.

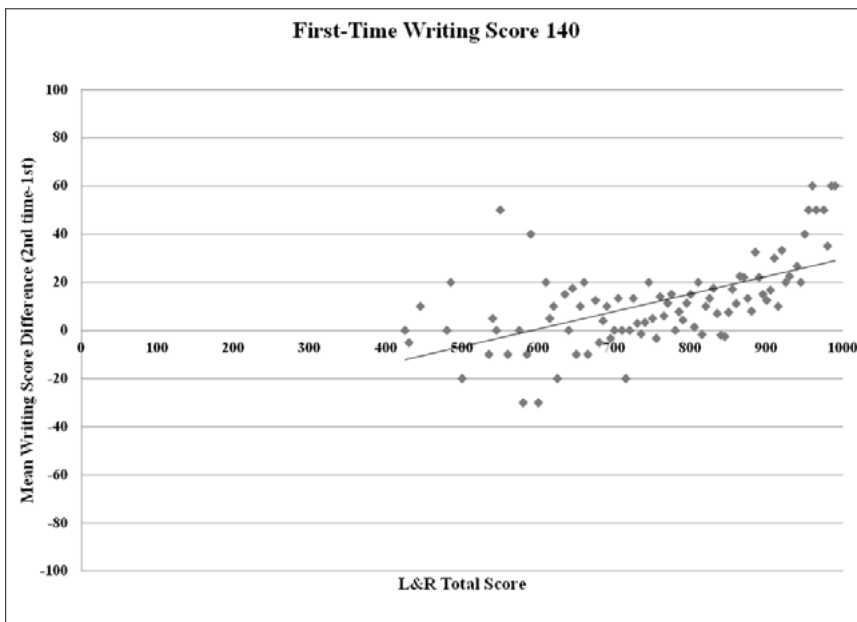


Figure A16. Mean score difference conditioned on the total score of the TOEIC Listening and Reading test (L&R): first TOEIC Writing (W) score = 140.

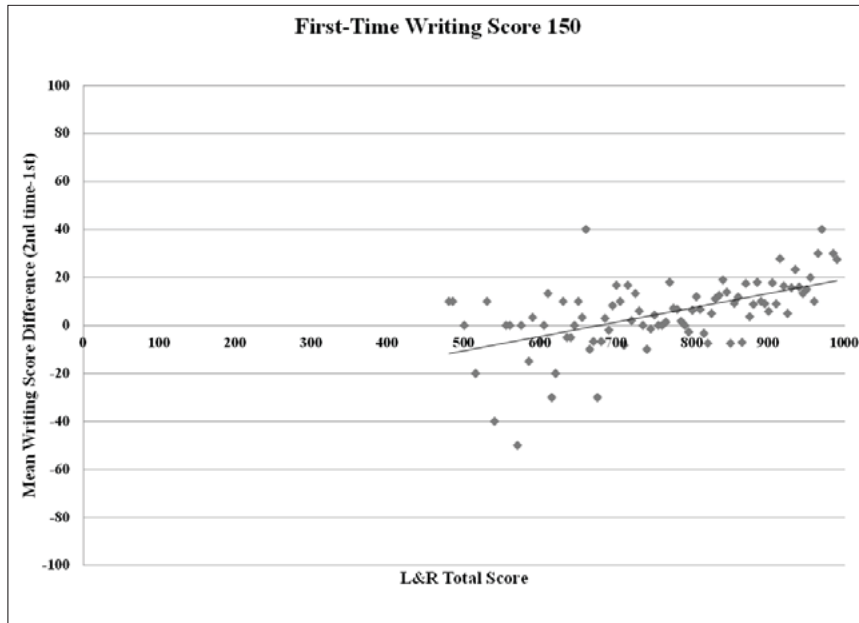


Figure A17. Mean score difference conditioned on the total score of the TOEIC Listening and Reading test (L&R): first TOEIC Writing (W) score = 150.

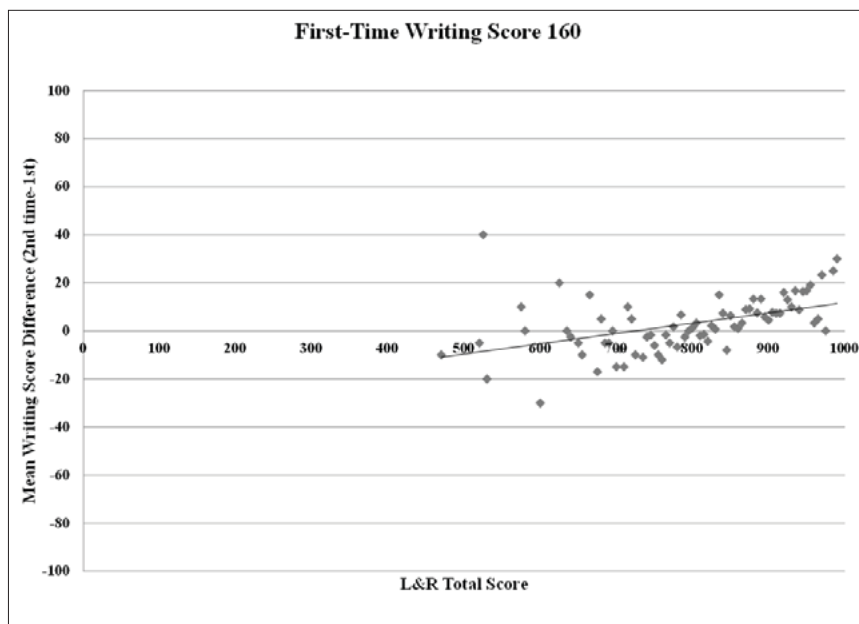


Figure A18. Mean score difference conditioned on the total score of the TOEIC Listening and Reading test (L&R): first TOEIC Writing (W) score = 160.

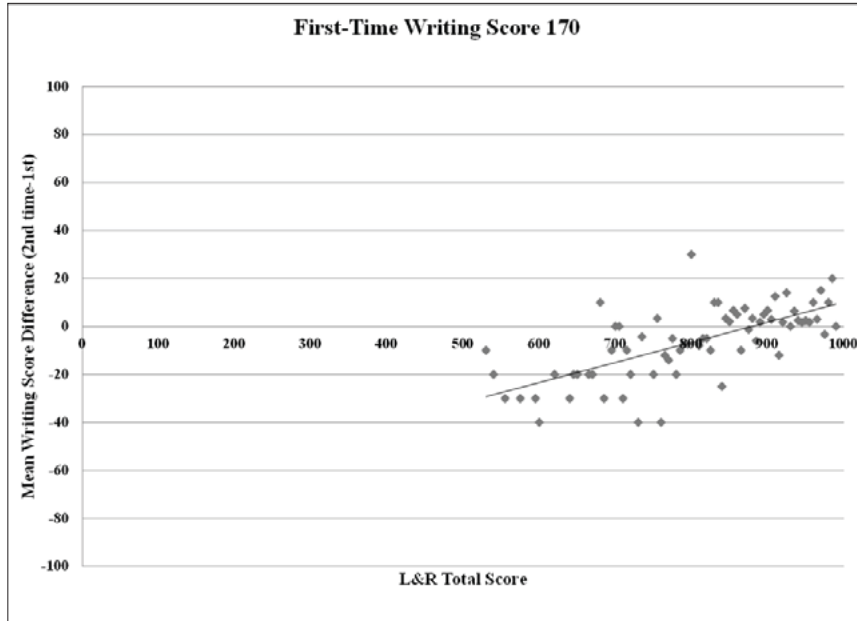


Figure A19. Mean score difference conditioned on the total score of the TOEIC Listening and Reading test (L&R): first TOEIC Writing (W) score = 170.

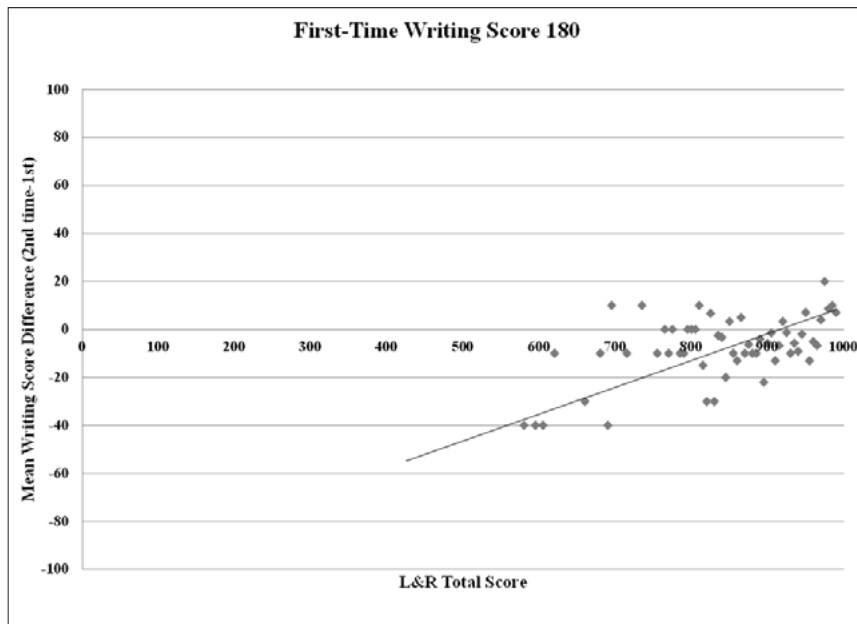


Figure A20. Mean score difference conditioned on the total score of the TOEIC Listening and Reading test (L&R): first TOEIC Writing (W) score = 180.