*Compendium Study*

# *TOEIC Bridge™* Scores: Validity Evidence From Korea and Japan

*Donald E. Powers and Fred Yan*

*September 2013*

It is generally accepted that no test can automatically be considered valid in all situations or for all purposes (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999). This statement is especially applicable to international assessments like the *TOEIC Bridge*™ test, a standardized multiple-choice test designed to measure the ability to listen to and read everyday English. Intended for beginning and (low) intermediate learners of English, the test has been found to be most appropriate for learners at the (basic) A1 and A2 levels and the (intermediate) B1 level of the Common European Framework (CEF; Tannenbaum & Wylie, 2008). Empirical studies have also deemed the test appropriate for students in South American countries by virtue of its relationship with (a) student self-assessments, (b) other locally developed tests, and (c) teacher ratings of student proficiency (Powers, Bravo et al., 2008; Powers, Mercadante, & Yan, 2012; Sinharay et al., 2009).

Because the standards for educational and psychological testing mandate that multiple sources of validity evidence are always desirable (and the more evidence, the better), the objective of this effort was to examine another source of validity evidence for *TOEIC Bridge* scores. Specifically, we sought evidence (for Japanese and Korean test takers) of the relationship between *TOEIC Bridge* scores and students' self-assessments of their own English-language skills, in particular their ability to perform a variety of everyday language tasks in English.

## Method

From June to September 2012, a total of approximately 2,600 *TOEIC Bridge* test takers in Japan and Korea completed a 50-item, self-assessment inventory that was previously used by Powers, Bravo, et al. (2008) in South America. (The questions were translated into Japanese and Korean.) This inventory consisted of everyday English-language listening and reading tasks that were adapted from an earlier, longer inventory devised for a validity study of the *TOEIC*® test (Powers, Kim, et al., 2008). Because the *TOEIC Bridge* test is intended to measure emerging English-language competencies, the tasks included in the inventory for this study were chosen so as to be significantly easier than those used for the earlier TOEIC test study. That is, we eliminated tasks that were rated in the earlier study as being the most difficult to perform. Response choices were on a 5-point scale as follows: 1 (*cannot do at all*), 2 (*can do with great difficulty*), 3 (*can do with some difficulty*), 4 (*can do with little difficulty*), and 5 (*can do easily*).

Analyses of data entailed computing correlations (for reading and for listening) between *TOEIC Bridge* scores and self-assessment ratings for composites based on all 25 listening tasks and all 25 reading tasks. For each task, cross-tabulations were also computed to display the percentages of test takers at each *TOEIC Bridge* score level who said that they could perform the task either easily or with little difficulty.

# Results

A total of 2,315 test takers (1,739 from Japan and 576 from Korea) completed all items on the 50-item inventory. For each country sample, the estimated reliability (coefficient alpha) was .96 to .97 for the listening and reading portions of the inventory.

The two country samples differed with respect to gender composition: 36% of the Japanese sample was female compared with 51% of the Korean sample. In addition, the test takers in the Korean sample were generally younger than those in the Japanese sample. The median age of the Korean sample test takers was approximately 13 years, while the median age of the Japanese sample test takers was approximately 18 years. About two thirds of the Korean participants were between the ages of 11 and 15; about two thirds of the Japanese participants were between 16 and 42 years old. The between-country age difference for the study samples is consistent with that found in the total population of *TOEIC Bridge* test takers for the 2012 testing year: Japanese test takers had a median age of approximately 29 years, while Korean test takers had a median age of approximately 13 years (Matthew Shotts, personal communication, April 29, 2013). The difference in the ages of the participants from the two countries may be due, at least in part, to the fact that English instruction is generally introduced at the third grade level in Korea and at the fifth grade level in Japan (Xiaoming Xi, personal communication, February 1, 2013).

Table 1 provides a summary of findings. The mean *TOEIC Bridge* scores of both the Japanese and the Korean samples were relatively high, indicating that, in general, study participants were relatively able in terms of English-language proficiency. The two samples were comparable in terms of listening proficiency as measured by the *TOEIC Bridge* test, while the Japanese sample was somewhat more able with respect to reading proficiency. With regard to their self-assessment of abilities, the two samples appear to be quite different. Self-assessment composite scores (based on all 25 listening tasks and on all 25 reading tasks) were higher for Korean test takers by about 1½ standard deviations on both the listening and reading portions of the inventory (mid to high 90s for Korean participants versus low 70s for Japanese participants). That is, on the 5-point response scale, the mean rating on all listening tasks was 3.92 for Korean test takers and 2.87 for Japanese test takers. The mean ratings for reading tasks were 3.78 and 2.82, respectively. Thus, the Korean test takers in our study gave ratings that were, on average, about a full point higher on the 5-point scale used to evaluate task difficulty. Despite the difference in mean ratings, both samples rank-ordered the tasks very similarly in terms of difficulty. In light of the significant mean difference in ages of the sample members and the substantial differences in self-assessment ratings (when compared with the more modest difference in test scores), we have elected to present results separately for the two country samples.

**Table 1**

*Correlations Among Test Taker Self-Assessments and TOEIC Bridge Scores*

| No. | Measure | M(SD) | 1 | 2 | 3 |
|-----|---------|-------|---|---|---|
| | | | Measure no. | | |
| | | Korean sample (*n* = 576) | | | |
| 1 | *TOEIC Bridge* Listening | 66.7 (9.9) | | | |
| 2 | *TOEIC Bridge* Reading | 58.1 (12.8) | 0.73 | | |
| 3 | Self-assess listening | 98.0 (15.8) | 0.35 | 0.25 | |
| 4 | Self-assess reading | 94.6 (18.3) | 0.31 | 0.22 | 0.90 |
| | | Japanese sample (*n* = 1,739) | | | |
| 1 | *TOEIC Bridge* Listening | 66.1 (8.6) | | | |
| 2 | *TOEIC Bridge* Reading | 66.6 (10.3) | 0.67 | | |
| 3 | Self-assess listening | 71.7 (17.7) | 0.39 | 0.26 | |
| 4 | Self-assess reading | 70.5 (18.3) | 0.40 | 0.35 | 0.85 |

*Note.* All correlations are significant at the *p* < .01 level or beyond.

As can be seen from Table 1, *TOEIC Bridge* Reading and Listening scores were, as expected, relatively highly related (*r* = .73 and .67) in both samples. Self-assessments (composites) of reading skills and listening skills were also strongly related in both samples (*r* = .90 and .85), suggesting that (a) the two rating scales generally reflect the same trait or (b) test takers were unable to differentiate their ability to perform real-world tasks in the listening and reading domains. Of greater interest, the correlations between the composite self-assessment ratings and *TOEIC Bridge* scores were modest (*r* = .22 to .35 in the Korean sample, and *r* = .26 to .40 in the Japanese sample).

Unfortunately, product-moment correlation coefficients do not always adequately capture the nature and strength of relationships between variables. Therefore, we have supplemented the correlational results in Table 1 by computing, for each listening and each reading task, the percentage of test takers at various *TOEIC Bridge* score levels who said that they could perform tasks either easily or with little difficulty. These results are shown for both the Korean sample (Tables 2 and 3) and for the Japanese sample (Tables 4 and 5). The mean (and standard deviation) of test-taker ratings for each of the 50 individual tasks is given also.

We note that six of the tasks in the listening task inventory and six in the reading task inventory were developed to correspond to the six levels of the Common European Framework of Reference for Language (CEF)—from basic user (Level A1) to independent user (Levels B1 and B2) to proficient user (Levels C1 and C2). Within each set of 25 tasks (listening and reading) for both the Korean and the Japanese samples, the CEF-derived tasks were rank-ordered exactly according to the level they were intended to represent. For instance, Listening Task 3 ("Recognize familiar words and simple phrases . . ."), which represents the A1 level, received the highest mean rating by each sample, indicating that it was perceived as the least difficult of the CEF-related listening tasks. On the other hand, Listening Task 25 ("Understand any kind of spoken language, even when people speak quickly"), which represents the highest CEF level (C2), was given the lowest mean rating of any of the

six CEF tasks by each sample, indicating that it was perceived as the most difficult listening task. The CEF-related reading tasks were also ordered completely consistently with the CEF levels. Test takers' mean ratings of the other, non-CEF-related tasks also corresponded quite well with our own expectations of the difficulty of the tasks, giving us some confidence in the trustworthiness of the ratings.

When compared to perceptions by the total sample, the percentages at various score levels (Tables 2 and 3 for the Korean sample) are somewhat less consistent with our expectations. In prior studies, we found that, almost without exception, the percentages of test takers who said they could perform a task increased consistently with each higher *TOEIC Bridge* score level. In the current study, this trend is somewhat less consistent for the Korean sample. For listening tasks, the percentages rise consistently with each higher test score level for nearly all of the least difficult tasks. For example, for Listening Task 11 in Table 2 ("Understand a sales person when he or she tells me the price of various items"), the increases are consistent and sharp, in this case from 51% at the lowest test score level to 93% at the highest. However, for the more difficult tasks (starting with Listening Task 15 on the list, "Understand the main ideas in news broadcasts on radio or TV"), there is a slight decrease in percentages from the lowest score level to the next lowest level, in this case, from 45% to 43%. Beyond these two lowest score levels, however, the expected pattern of consistently rising percentages begins to appear.

For the reading tasks shown in Table 3 (Korean sample), the same decreasing (or inconsistent) pattern of percentages is apparent for the most difficult of the 25 reading tasks at the three lowest *TOEIC Bridge* score levels. For example, for Listening Task 11 ("Find information I need in a telephone directory"), the percentages decrease from 70 at the lowest level to 68 at the next higher level then to 64 at the next before they begin to increase systematically starting at 81 at the next higher level. Only from the third lowest reading score level and higher do percentages exhibit the expected consistent increase. For the most difficult tasks, this expected pattern of increasing percentages begins to appear slightly later at the next higher score level.

Tables 4 and 5 display results for listening and reading tasks for the Japanese sample. As can be seen in Table 4 (the listening results for the Japanese sample), the expected pattern of increasing percentages appears for each of the 25 listening tasks: The percentage of test takers who say that they can perform a task increases consistently with each higher TOEIC Listening score level.

When results are examined for reading tasks for the Japanese sample (Table 5), increasing percentages at each higher score level are entirely consistent with expectations, that is, increasing percentages of test takers indicating ability to perform each at each higher reading score level. Thus, the results in the Japanese sample are much more consistent with expectations than are the results in the Korean sample, in part likely due to the differences in the sizes of the two country samples.

## Table 2

*Percentages of TOEIC Bridge Test Takers by Listening Score Level Who Indicated That They Could Perform Various English-Language Listening Tasks Either Easily or With Little Difficulty for Korean Sample*

| Task | TOEIC Bridge Listening score level | | | | | |
|---|---|---|---|---|---|---|
| | 42–50 | 52–60 | 62–70 | 72–80 | 82–90 | M |
| 1. Understand the days of the week and months of the year. | 85 | 96 | 98 | 98 | 100 | 4.76 |
| 2. Understand simple questions in social situations such as "How are you?" and "Where do you live?" | 95 | 95 | 99 | 99 | 100 | 4.86 |
| **3. Recognize familiar words and simple phrases when people speak slowly and clearly (CEF Level A1–basic user).** | 74 | 83 | 95 | 96 | 98 | 4.53 |
| 4. Understand a person's name when she or he gives it to me over the telephone. | 100 | 92 | 97 | 97 | 98 | 4.74 |
| 5. Understand some memorized words and phrases. | 79 | 83 | 92 | 93 | 98 | 4.50 |
| **6. Understand the main point of simple messages and short, clear announcements (CEF Level A2–basic user).** | 70 | 69 | 83 | 88 | 100 | 4.25 |
| 7. Understand someone who is speaking slowly and deliberately about his or her hobbies and interests. | 64 | 79 | 87 | 93 | 98 | 4.40 |
| 8. Understand directions about what time to come to a class and where it will be held. | 70 | 72 | 86 | 92 | 100 | 4.38 |
| 9. Understand the main points of standard speech on familiar matters when people speak slowly and clearly (CEF Level B1–independent user). | 62 | 72 | 81 | 88 | 95 | 4.22 |
| 10. Understand someone speaking slowly and deliberately, who is giving me directions on how to walk to a nearby location. | 74 | 70 | 84 | 92 | 100 | 4.38 |
| 11. Understand a salesperson when she or he tells me prices of various items. | 51 | 59 | 77 | 90 | 93 | 4.18 |
| 12. Understand a person in social situations talking about his/her background, family, or interests. | 55 | 65 | 75 | 86 | 88 | 4.09 |
| 13. Understand a classmate discussing a simple problem that arose in class. | 55 | 62 | 82 | 87 | 93 | 4.23 |
| 14. Understand explanations about how to perform a routine task related to my class. | 55 | 59 | 68 | 80 | 93 | 4.01 |
| 15. Understand the main ideas in news reports broadcast on the radio or TV. | 45 | 43 | 54 | 61 | 70 | 3.67 |
| 16. Understand public announcements that are broadcast. | 49 | 49 | 61 | 72 | 90 | 3.83 |
| 17. Understand headline news broadcasts on the radio. | 49 | 42 | 53 | 63 | 75 | 3.64 |
| 18. Understand play-by-play descriptions on the radio of sports events that I like (e.g., soccer). | 47 | 46 | 53 | 58 | 78 | 3.61 |
| 19. Understand an explanation given over the radio of why a road has been temporarily closed. | 36 | 24 | 47 | 53 | 73 | 3.38 |
| 20. Understand a complex presentation or demonstration given in a classroom. | 36 | 27 | 39 | 49 | 65 | 3.27 |
| **21. Understand extended speech and lectures, and follow complex arguments on familiar topics (CEF Level B2–independent user).** | 34 | 22 | 30 | 37 | 55 | 3.09 |
| 22. Understand a discussion of current events taking place among a group of persons speaking English. | 34 | 20 | 33 | 45 | 75 | 3.22 |
| **23. Understand extended speech even when it is not clearly structured (CEF Level C1–proficient user).** | 30 | 28 | 37 | 48 | 70 | 3.21 |
| 24. Understand an extended debate on a complex topic such as public health. | 19 | 17 | 27 | 36 | 58 | 2.92 |
| **25. Understand any kind of spoken language, even when people speak quickly (CEF Level C2–proficient user).** | 15 | 9 | 21 | 29 | 48 | 2.70 |
| *N* for score interval | 47 | 98 | 237 | 153 | 40 | |

## Table 3

*Percentages of TOEIC Bridge Test Takers by Reading Score Level Who Indicated That They Could Perform Various English-Language Reading Tasks Either Easily or With Little Difficulty for Korean Sample*

| Task | TOEIC Bridge Reading score level | | | | | | |
|---|---|---|---|---|---|---|---|
| | 32–40 | 42–50 | 52–60 | 62–70 | 72–80 | 82–90 | M |
| **1. Understand familiar words and very simple sentences (CEF Level A1– basic user).** | 83 | 91 | 95 | 96 | 98 | 95 | 4.71 |
| 2. Recognize memorized words and phrases (for example, "exit," "entrance," and "stop"). | 75 | 88 | 88 | 88 | 95 | 100 | 4.50 |
| 3. Read, on storefronts, the type of store or services provided (e.g., "dry cleaning," "book store"). | 65 | 76 | 86 | 91 | 96 | 90 | 4.41 |
| **4. Understand short, simple texts (e.g., personal letters; CEF Level A2– basic user).** | 78 | 81 | 91 | 92 | 95 | 95 | 4.48 |
| 5. Read and understand traffic signs. | 73 | 83 | 86 | 91 | 95 | 100 | 4.46 |
| 6. Read and understand a simple email from a friend. | 75 | 72 | 81 | 89 | 93 | 95 | 4.32 |
| 7. Read and understand a train or bus schedule. | 70 | 74 | 78 | 88 | 93 | 95 | 4.28 |
| 8. Understand texts that consist mainly of everyday words or vocabulary (CEF Level B1–independent user). | 63 | 75 | 77 | 89 | 95 | 100 | 4.24 |
| 9. Read and understand a restaurant menu. | 70 | 73 | 74 | 85 | 96 | 95 | 4.24 |
| 10. Read and understand simple, step-by-step instructions (for example, how to operate an electronic machine, like a CD player). | 55 | 61 | 59 | 71 | 81 | 75 | 3.87 |
| 11. Find information that I need in a telephone directory. | 70 | 68 | 64 | 81 | 87 | 80 | 4.10 |
| 12. Read and understand a letter of thanks from a friend or a relative. | 60 | 71 | 70 | 79 | 89 | 95 | 4.18 |
| 13. Read information about products (for example, advertisements) | 55 | 58 | 56 | 74 | 83 | 90 | 3.90 |
| 14. Read and understand directions and explanations presented in manuals written for beginning users. | 55 | 63 | 59 | 72 | 73 | 90 | 3.89 |
| 15. Read entertainment-related information (for example, tourist guides). | 45 | 46 | 49 | 58 | 69 | 75 | 3.62 |
| 16. Read and understand main points of an article on a familiar academic topic. | 53 | 48 | 46 | 63 | 67 | 95 | 3.68 |
| 17. Read English to translate text into my own language (for example, letters and technical documents). | 35 | 35 | 35 | 46 | 53 | 75 | 3.25 |
| 18. Read and understand a popular novel. | 40 | 32 | 36 | 42 | 49 | 80 | 3.22 |
| 19. Read and understand magazine articles without using a dictionary. | 40 | 34 | 30 | 33 | 47 | 40 | 3.09 |
| 20. Read highly technical material related to my classes with little use of a dictionary. | 40 | 35 | 33 | 33 | 41 | 40 | 3.11 |
| **21. Understand the viewpoints expressed in articles and reports about contemporary issues or problems (CEF Level B2–independent user).** | 50 | 41 | 41 | 42 | 54 | 65 | 3.30 |
| 22. Read a newspaper editorial and understand its meaning as well as the writer's intent. | 45 | 33 | 35 | 40 | 55 | 45 | 3.17 |
| 23. Identify inconsistencies or differences in points of view in two newspaper interviews with politicians of opposing parties. | 43 | 32 | 35 | 33 | 46 | 50 | 3.12 |
| **24. Understand long, complex texts, even technical ones that do not relate to my classes (CEF Level C1–proficient user).** | 33 | 27 | 27 | 28 | 31 | 50 | 2.81 |
| **25. Understand virtually all forms of written language, including abstract and linguistically complex text (CEF Level C2–proficient user).** | 35 | 22 | 25 | 26 | 35 | 45 | 2.71 |
| *N* for score interval | 40 | 149 | 138 | 144 | 83 | 20 | |

*Note.* Number of responses for each statement is 576. Responses were on a 5-point scale from 5 (*can do easily*) to 1 (*cannot do at all*). CEF= Common European Framework. CEF-related statements are shown in bold. SDs for responses ranged from 0.60 to 1.20.

## Table 4

*Percentages of TOEIC Bridge Test Takers by Listening Score Level Who Indicated That They Could Perform Various English-Language Listening Tasks Either Easily or With Little Difficulty for Japanese Sample*

| Task | *TOEIC Bridge* Listening score level | | | | | |
|---|---|---|---|---|---|---|
| | 42–50 | 52–60 | 62–70 | 72–80 | 82–90 | M |
| 1. Understand the days of the week and months of the year. | 69 | 82 | 84 | 89 | 98 | 4.34 |
| 2. Understand simple questions in social situations such as "How are you?" and "Where do you live?" | 41 | 66 | 77 | 87 | 98 | 4.14 |
| **3. Recognize familiar words and simple phrases when people speak slowly and clearly (CEF Level A1–basic user).** | 32 | 57 | 67 | 77 | 93 | 3.86 |
| 4. Understand a person's name when she or he gives it to me over the telephone. | 28 | 45 | 56 | 63 | 86 | 3.64 |
| 5. Understand some memorized words and phrases. | 20 | 39 | 55 | 65 | 86 | 3.62 |
| **6. Understand the main point of simple messages and short, clear announcements (CEF Level A2–basic user).** | 8 | 32 | 43 | 59 | 86 | 3.41 |
| 7. Understand someone who is speaking slowly and deliberately about his or her hobbies and interests. | 12 | 28 | 39 | 54 | 82 | 3.31 |
| 8. Understand directions about what time to come to a class and where it will be held. | 7 | 22 | 38 | 50 | 81 | 3.29 |
| **9. Understand the main points of standard speech on familiar matters when people speak slowly and clearly (CEF Level B1–independent user).** | 9 | 24 | 38 | 51 | 86 | 3.26 |
| 10. Understand someone speaking slowly and deliberately, who is giving me directions on how to walk to a nearby location. | 9 | 24 | 33 | 46 | 74 | 3.21 |
| 11. Understand a salesperson when she or he tells me prices of various items. | 8 | 21 | 32 | 43 | 75 | 3.16 |
| 12. Understand a person in social situations talking about his/her background, family, or interests. | 0 | 9 | 17 | 27 | 61 | 2.75 |
| 13. Understand a classmate discussing a simple problem that arose in class. | 3 | 8 | 14 | 25 | 56 | 2.63 |
| 14. Understand explanations about how to perform a routine task related to my class. | 0 | 7 | 14 | 23 | 61 | 2.64 |
| 15. Understand the main ideas in news reports broadcast on the radio or TV. | 3 | 7 | 11 | 18 | 42 | 2.54 |
| 16. Understand public announcements that are broadcast. | 0 | 5 | 11 | 17 | 46 | 2.49 |
| 17. Understand headline news broadcasts on the radio. | 0 | 6 | 10 | 15 | 39 | 2.44 |
| 18. Understand play-by-play descriptions on the radio of sports events that I like (e.g., soccer). | 4 | 9 | 12 | 16 | 35 | 2.42 |
| 19. Understand an explanation given over the radio of why a road has been temporarily closed. | 1 | 4 | 9 | 13 | 40 | 2.31 |
| 20. Understand a complex presentation or demonstration given in a classroom. | 0 | 3 | 6 | 8 | 33 | 2.14 |
| **21. Understand extended speech and lectures, and follow complex arguments on familiar topics (CEF Level B2–independent user).** | 0 | 2 | 6 | 9 | 35 | 2.15 |
| 22. Understand a discussion of current events taking place among a group of persons speaking English. | 0 | 3 | 6 | 7 | 32 | 2.09 |
| **23. Understand extended speech even when it is not clearly structured (CEF Level C1–proficient user).** | 0 | 1 | 5 | 7 | 30 | 2.04 |
| 24. Understand an extended debate on a complex topic such as public health. | 0 | 1 | 4 | 5 | 26 | 1.88 |
| **25. Understand any kind of spoken language, even when people speak quickly (CEF Level C2–proficient user).** | 1 | 1 | 4 | 7 | 32 | 1.91 |
| *N* for score interval | 75 | 327 | 890 | 385 | 57 | |

*Note.* Number of responses for each statement is 1,739. Responses were on a 5-point scale from 5 (*can do easily*) to 1 (*cannot do at all*). SDs for responses ranged from 0.79 to 1.01. CEF = Common European Framework. CEF-related statements are shown in bold.

**Table 5**

*Percentages of TOEIC Bridge Test Takers by Reading Score Level Who Indicated That They Could Perform Various English-Language Reading Tasks Either Easily or With Little Difficulty for Japanese Sample*

| Task | TOEIC Bridge Reading score level | | | | | | |
|---|---|---|---|---|---|---|---|
| | 32–40 | 42–50 | 52–60 | 62–70 | 72–80 | 82–90 | M |
| **1. Understand familiar words and very simple sentences (CEF Level A1–basic user).** | 30 | 60 | 61 | 71 | 81 | 86 | 4.01 |
| 2. Recognize memorized words and phrases (for example, "exit," "entrance," and "stop"). | 30 | 56 | 59 | 68 | 82 | 87 | 3.99 |
| 3. Read, on storefronts, the type of store or services provided (e.g., "dry cleaning," "book store"). | 30 | 51 | 53 | 63 | 74 | 85 | 3.84 |
| **4. Understand short, simple texts (e.g., personal letters; CEF Level A2–basic user).** | 10 | 34 | 42 | 55 | 66 | 80 | 3.64 |
| 5. Read and understand traffic signs. | 10 | 37 | 38 | 49 | 63 | 75 | 3.56 |
| 6. Read and understand a simple email from a friend. | 10 | 25 | 36 | 48 | 61 | 80 | 3.49 |
| 7. Read and understand a train or bus schedule. | 10 | 24 | 33 | 46 | 58 | 71 | 3.45 |
| **8. Understand texts that consist mainly of everyday words or vocabulary (CEF Level B1–independent user).** | 0 | 18 | 24 | 33 | 49 | 68 | 3.26 |
| 9. Read and understand a restaurant menu. | 10 | 18 | 27 | 33 | 45 | 58 | 3.22 |
| 10. Read and understand simple, step-by-step instructions (for example, how to operate an electronic machine, like a CD player). | 0 | 15 | 22 | 25 | 40 | 59 | 3.05 |
| 11. Find information that I need in a telephone directory. | 10 | 11 | 18 | 23 | 32 | 49 | 2.91 |
| 12. Read and understand a letter of thanks from a friend or a relative. | 0 | 9 | 12 | 22 | 35 | 57 | 2.89 |
| 13. Read information about products (for example, advertisements). | 0 | 8 | 15 | 23 | 35 | 51 | 2.93 |
| 14. Read and understand directions and explanations presented in manuals written for beginning users. | 0 | 6 | 15 | 19 | 31 | 51 | 2.84 |
| 15. Read entertainment-related information (for example, tourist guides). | 0 | 8 | 12 | 18 | 27 | 47 | 2.78 |
| 16. Read and understand main points of an article on a familiar academic topic. | 0 | 5 | 11 | 13 | 22 | 39 | 2.59 |
| 17. Read English to translate text into my own language (for example, letters and technical documents). | 0 | 1 | 6 | 8 | 13 | 22 | 2.27 |
| 18. Read and understand a popular novel. | 0 | 2 | 5 | 5 | 9 | 20 | 2.10 |
| 19. Read and understand magazine articles without using a dictionary. | 0 | 1 | 4 | 6 | 10 | 20 | 2.13 |
| 20. Read highly technical material related to my classes with little use of a dictionary. | 0 | 1 | 6 | 5 | 8 | 20 | 2.07 |
| **21. Understand the viewpoints expressed in articles and reports about contemporary issues or problems (CEF Level B2–independent user).** | 0 | 2 | 2 | 5 | 5 | 19 | 2.01 |
| 22. Read a newspaper editorial and understand its meaning as well as the writer's intent. | 0 | 1 | 4 | 4 | 6 | 20 | 2.01 |
| 23. Identify inconsistencies or differences in points of view in two newspaper interviews with politicians of opposing parties. | 0 | 1 | 3 | 4 | 4 | 14 | 1.85 |
| **24. Understand long, complex texts, even technical ones that do not relate to my classes (CEF Level C1–proficient user).** | 0 | 0 | 3 | 3 | 5 | 14 | 1.82 |
| **25. Understand virtually all forms of written language, including abstract and linguistically complex text (CEF Level C2–proficient user).** | 0 | 0 | 3 | 4 | 4 | 14 | 1.80 |
| *N* for score interval | 10 | 131 | 332 | 704 | 423 | 138 | |

*Note.* Number of responses for each statement is 1,739. Responses were on a 5-point scale from 5 (*can do easily*) to 1 (*cannot do at all*). SDs for responses ranged from 0.67 to 1.00. CEF = Common European Framework. CEF-related statements are shown in bold.

# Discussion

The results of the study presented here suggest that *TOEIC Bridge* scores relate significantly to test-taker assessments of their own English-language skills. The size of the correlations computed here between self-assessments and *TOEIC Bridge* scores can, according to conventional standards for effect sizes in the social sciences, be characterized as being in the moderate range (Cohen, 1988). The relationships noted in the current study are somewhat weaker than those reported in previous similar studies. In particular, in prior studies we noted strong, consistent patterns in terms of steadily increasing percentages of examinees at each higher test score level who indicated that they could perform a variety of everyday language tasks in English. Here, we noted inconsistencies in one of our two country samples (Korea), especially for reading tasks at the lowest score levels (where *TOEIC Bridge* scores are somewhat less discriminating): The percentage of test takers who reported that they could perform a task was, contrary to expectations, sometimes slightly higher at a lower score level than at the next higher level. This was true especially for the most difficult tasks.

In the larger Japanese sample, the results were entirely consistent with expectations: The percentages of test takers who reported that they could perform a task rose consistently with each higher score level for every reading and every listening task. We strongly suspect that the results differed according to sample for several reasons. The most apparent difference between samples was country of residence, thus raising the prospect of cross-cultural differences in self-reporting of proficiencies or in test-taking ability. A possibly more germane factor, however, may be age of test takers: The test takers in the Korean sample were, on average, significantly younger than those in the Japanese sample. This Korean sample had mean test scores that were lower for reading than those of the Japanese sample; mean scores on the listening portion of the test were, on the other hand, more nearly comparable. However, perceptions of the ability to perform each of the 50 language tasks were substantially more positive for the Korean sample than for the Japanese sample. While cross-cultural factors cannot be ruled out as a cause of this discrepancy, it seems more likely that, because we did not detect differences of this sort in prior studies, the difference in the age of the test takers in the samples is responsible. Younger test takers may have different standards than older test takers for what it means to be able to successfully perform English-language tasks. In addition, the Japanese sample was approximately three times as large as the Korean sample, thus producing more stable results across test score levels, especially at levels containing only small numbers of test takers. A major lesson learned here is that caution should be exercised when pooling possibly diverse samples for analysis, as this pooling may mask true relationships in more homogeneous samples.

In conclusion, within the limits of the study, the results provide some modest additional evidence of the validity of the *TOEIC Bridge* test as an indicator of English-language proficiency. These results are both statistically and practically significant with regard to making decisions about students' emerging proficiency in English. The data support this conclusion despite certain study limitations, which may have hindered finding even stronger relationships.

# References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Powers, D. E., Bravo, G. M., Sinharay, S., Valdivia, L. E., Simpson, A. G. & Weng, V. Z. (2008). *Relating scores on the TOEIC Bridge to student perceptions of proficiency in English* (Research Memorandum no. RM-08-02). Princeton, NJ: Educational Testing Service.

Powers, D. E., Kim, H.-J., & Weng, V. (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, Mercadante, R., & Yan, F. (2012). *Validating TOEIC Bridge scores against teacher ratings for vocational students in China* (Research Memorandum No. RM-22-12). Princeton, NJ: Educational Testing Service.

Sinharay, S., Powers, D. E., Feng, Y., Saldivia, L., Guinta, A., Simpson, A., & Weng, V. (2009). Appropriateness of the *TOEIC Bridge* test for students in three countries of South America. *Language Testing*, *26*, 589–619.

Tannenbaum, R. J., & Wylie, E. C. (2008). *Linking English-language test scores onto the Common European Framework of Reference: An application of standard-setting methodology* (TOEFL iBT Research Report No. 06). Princeton, NJ: Educational Testing Service.