

Ms Jenny Collis
Australian Health Practitioners Regulatory Authority
GPO Box 9958
Melbourne
VIC 3001
20 December 2013

Educational Testing Service
P.O. Box 6151
Princeton, NJ 08541-6151, USA
Phone: (877) 863-3546
(609) 771-7100
Fax: (609) 771-7500
e-mail: toefl@ets.org

Dear Jenny

Thank you for the opportunity to respond to the Australian Health Practitioners Regulatory Authority (AHPRA) **Review of English language skills registration standard**

Educational Testing Service (ETS) appreciates the importance of setting appropriate English language proficiency standards for the health professions. We understand that English proficiency plays a critical role among health practitioners as errors caused by insufficient language ability can be life-threatening, a far higher level of consequence than making a poor admission decision at an institution.

ETS wishes to submit the following information relating to the questions posed regarding the review of the standards for your consideration.

Additional English language tests to meet the English proficiency standards

The internet based TOEFL® (TOEFL iBT®) test launched in 2005 provides the National Boards with an additional global English proficiency test with which professional bodies can measure the English proficiency of those candidates whom Boards determine must undertake an English proficiency test. As noted in the Summary of Research Findings, *IELTS and TOEFL are the main tests accepted worldwide with TOEFL dominate in selected part of Asia and North America*

Although the TOEFL test is less well known in Australia than the IELTS tests, the TOEFL test is the most long standing English proficiency test, and will celebrate its 50th anniversary in 2014. The TOEFL iBT test is the most widely available English proficiency test with around 4000 test centres worldwide. It is also the most widely accepted English proficiency test. TOEFL test scores are accepted by over 9000 institutions, government agencies and professional bodies globally to make important decisions.

As noted in the Summary of Research, TOEFL test results are accepted in Australia for student visas. Thus many students entering Australian institutions come with TOEFL test results. However currently, most medical and allied professions require IELTS only, thus restricting the pool of applicants for these courses.

Educational Testing Service
P.O. Box 6151
Princeton, NJ 08541-6151, USA
Phone: (877) 863-3546
(609) 771-7100
Fax: (609) 771-7500
e-mail: toefl@ets.org

AHPRA will be interested to learn that the Department of Immigration and Border Protection (DIBP) has completed a Review of the Implementation of Alternative Tests (to IELTS) commenced in 2011 and found that: “ Following analysis of submissions, quantitative data, interviews and survey results, the Department is satisfied that the Test of English as a Foreign Language (TOEFL) internet-based test (iBT) ..., met the objectives of the Student visa programme during the review period.”

In terms of fulfilling service delivery deed of agreement obligations, the Review found the TOEFL test “..., demonstrated that they updated their test content regularly, employed trained markers and invigilators and implemented standardised testing conditions in all test centres. Their storage of test material and availability for audit purposes by the Department complied with the requirements for the Student visa programme. Security measures for internet based testing, minimizing risk of identity substitution and document fraud were also of a high standard.”

“The introduction of alternative English language tests has delivered more choice to Student visa applicants who are required to provide evidence of English language proficiency with their visa application.”

DIBP is now working through the processes for approval for extending the English proficiency requirements to other visa categories such as skilled migration.

It would therefore be timely for the national standards review to consider additional English language proficiency tests for the National Standards.

Accepting test scores from multiple sittings

The ETS TOEFL position regarding resitting of tests is similar to that advised by IELTS for its test. The four component modules are not offered as separate tests to be available at different times.

We understand the frustration of test takers who receive variable scores from sitting an English proficiency test. We would draw the Board's attention to Repeater Research on TOEFL iBT results.

<http://www.ets.org/Media/Research/pdf/RM-08-05.pdf> This research shows a very small mean score difference for over 12000 test takers retesting shortly after taking a prior TOEFL iBT test.

We would submit that policy should not be driven by the fact that candidates for some English proficiency tests regularly achieve inconsistent results from retest scores. It might be more pertinent to look at the test instruments, their scoring policies and the period of retesting/retraining for markers rather than altering the Standards.

ETS invests heavily on ensuring objectivity and consistency of test item marking. For example, scoring processes ensure that multiple markers score Speaking and Writing items of each candidate. Markers are not only trained and closely supervised in real time during scoring but have to pass their own calibration test each time they begin marking a question. Markers never meet candidates or are aware of any aspects of their identity so cannot be influenced by personal issues. Markers score results from around the world so that globally consistent scores are maintained. TOEFL test taker score data are not maintained at the test centre and score results are not published from the centre itself.

This can be compared to marking for the IELTS test, where after an initial period of training, markers are examined once every two years. IELTS tests are marked in the same centre where the test was administered, test taker data is retained at the test centre and score results are prepared at the test centre. IELTS markers receive responses only from local test takers and do not receive an anonymous, randomised worldwide set of responses.

Online Verification of TOEFL iBT results

Access to a free, easy-to-use service of Online Verification is available to professional bodies and institutions. The service is available 24/7. ETS will support new users with additional information and webinars if required.

Access to standard setting for professional bodies

In response to a DIBP requirement, ETS conducted score comparison research between the TOEFL test and the IELTS academic test. For the research, ETS

Educational Testing Service
P.O. Box 6151
Princeton, NJ 08541-6151, USA
Phone: (877) 863-3546
(609) 771-7100
Fax: (609) 771-7500
e-mail: toefl@ets.org

obtained a sample of 1,153 students who had both IELTS and TOEFL scores. Under this approach, TOEFL iBT score ranges are identified for test takers meeting a particular score level on the equivalent IELTS score band. The research report is attached for your information.

However, because the TOEFL program has long experience with the health professions, and understands the seriousness of determining appropriate score requirements, ETS strongly encourages the health bodies to determine minimum score requirements for the TOEFL iBT test via standard setting. Attached is an outline of a standard setting approach for the health professions.

ETS offers a service in Australia to institutions and professional bodies to assist them to identify the most suitable minimum scores for practitioners in specific disciplines. This practice is widely used in North America to provide a valid, professional approach to this important issue. A copy of a standard setting report for the Nursing profession in the United States is attached.

Additional Option for Chinese Medicine

The current standard includes a specific test result for Chinese medicine. The score of 237 comes from the computer based TOEFL test (TOEFL CBT). This test was discontinued quite a few years ago and no valid test scores exist. The computer based TOEFL test was replaced by the internet based TOEFL (TOEFL iBT). We would recommend that TOEFL iBT scores be determined for this option to be useful to the Chinese Medicine profession.

Thank you again for the opportunity to respond to the Review. Please do not hesitate to contact me if you require more information.

Kind regards

Helen Cook
Associate Director, Client Relations
ETS TOEFL
[Redacted]

Enc *Linking TOEFL®iBT to IELTS scores – a Research Report*

Standard Setting Fact Sheet

*A Minimum English Proficiency Standard for the Test of
English as a Foreign Language Internet-Based Test – NCSBN
USA*

Educational Testing Service
P.O. Box 6151
Princeton, NJ 08541-6151, USA
Phone: (877) 863-3546
(609) 771-7100
Fax: (609) 771-7500
e-mail: toefl@ets.org

Linking *TOEFL iBT*™ Scores to
IELTS® Scores –

A Research Report

Abstract

TOEFL® test scores have been widely accepted as evidence of nonnative English speakers' proficiency in the use of English. The TOEFL test measures test takers' ability to communicate in English-medium colleges and universities.

The TOEFL iBT™ test measures a candidate's ability to combine listening, reading, speaking and writing skills with authentic and innovative integrated tasks, making it the most comprehensive English-language proficiency test. Accepted by over 7,500 score users in more than 130 countries, the TOEFL test is the most widely recognized English-language test in the world.

Recently, ETS conducted score comparison research between the TOEFL test and an alternative test, IELTS (International English Language Testing System). For the research, ETS obtained a sample of 1,153 students who had both IELTS and TOEFL scores. Using equipercentile linking, the corresponding TOEFL score that would pass the same percentage of test takers for each IELTS scores was obtained. The score comparison results for each section (Listening, Speaking, Reading, and Writing) and the total test showed that most of the students scored in the middle to mid-high score ranges on both tests. Limitations of the research and suggestions for future research are also discussed.

Linking TOEFL iBT™ Scores to IELTS Scores

Introduction

TOEFL® test scores have been widely accepted by educational and other institutions throughout the world as evidence of nonnative English speakers' proficiency in English language use, and the TOEFL test has received very positive feedback from score users and instructors in the English as a Second Language (ESL) community.

There are many occasions when scores from different tests or even from different versions of the same test need to be linked for decision making or other purposes. For example, the scores on the SAT® I and ACT® tests, two well known college admissions tests in the United States, have been linked to each other to help university admissions offices use the scores to evaluate applicants. A research study was carried out by both ETS and ACT research staff to develop a score comparison table for this purpose (Pommerich, Hanson, Harris, & Scoring, 2000.) Another example is that, when the TOEFL iBT test, was developed to be taken via the Internet and to replace the computer-adaptive test (cBT) version, a study was conducted to link TOEFL iBT scores to TOEFL cBT scores (Wang, Eignor, & Enright, 2007.)

Recently, ETS conducted score comparison research between the TOEFL test and an alternative test, IELTS (International English Language Testing System). As a best practice required by the Guidelines for Practice by the International Language Testing Association (ILTA, 2007), the Standards for Educational and Psychological Testing (AERA, APA, NCME, 1999), and the ETS standards (ETS, 2002, p. 45,), appropriate psychometric procedures should always be used to link scores from two different assessments if the scores are to be compared. To comply with this standard, ETS psychometric staff conducted a research study to explore and establish empirical relationships between the scores on the two tests. This research was designed and carried out to answer the following two research questions:

1. What TOEFL iBT section scores are comparable to IELTS section scores?
2. What TOEFL iBT total scores are comparable to IELTS total scores?

Linking *TOEFL iBT*™ Scores to IELTS® Scores – A Research Report

Finding the corresponding TOEFL iBT scores for each of the IELTS scores was accomplished by what is known as score linking or score comparison (called score comparison hereafter in this report) in the field of educational measurement.

In this report, the samples and the analyses will be described in the Method section. The section for Results will present score comparisons for the four skill sections and the total scores. This is followed by the Discussion section that considers the interpretation of the results and their implications in terms of the generalization of the findings from this study, evaluates the limitations of the research, and makes suggestions for future research.

Method

Data Collection

In order to acquire empirical evidence for linking the scores on the two tests, scores were obtained from students who had taken both tests (Pommerich, 2007). The data collection effort for this research started in 2008 and continued into 2009. The first batch of data, obtained in 2008, contained 217 students who submitted to ETS copies of their TOEFL and IELTS score reports. In order to increase the sample size and the stability of the results, ETS then contacted a large number of test takers via e-mail and encouraged the participants to submit their latest scores on both tests to ETS. Over 1,000 students responded and reported their scores on the two tests to ETS. The two batches of data were then combined for the study.

Before any analyses started, the first task was to clean the data, which involved removing unusable records. Records that had missing scores on one or more sections or that had invalid scores were removed from subsequent analyses. For example, one reported IELTS score of 4.7 was identified as an invalid score and was eliminated from the analysis data set¹. After this data cleaning procedure, 1,153 records were retained for the analyses described in the next section. Close to 1,000 records included the participants' country of testing information, including about 70 countries from around the world. The largest group was from China including Hong Kong (41%), followed by USA (6%), and Japan (5%). The others were all below 4%.

Data Analysis

To understand the statistical characteristics of the scores on the two tests from the final sample of 1,153 students and to establish the relationships between the scores on the two tests, the following types of analysis were conducted. The findings from the analyses are presented in the Results section.

1. Descriptive statistics were computed to evaluate such information as the average scores on each test at both the section and the total score level, and how much variation (standard deviation) was observed in the scores. The correlation between the scores on the two tests was also computed to see to what extent the scores on the two tests were related to each other.

¹ IELTS scores have 9 bands in one-half increments (e.g., 3, 3.5, 4, 4.5...)

2. Linking scores is basically a transformation from a score on one scale to a score on the other scale (Holland, 2007) and can be accomplished in a variety of ways depending on the nature of the data and the purpose of the linking (Kolen & Brennan, 1995; Holland & Dorans, 2006). To find the correspondence between the TOEFL iBT and IELTS tests for each section and for the total score, ETS psychometric staff conducted extensive analyses between March and August 2009 using a variety of statistical methods including regression-based prediction, conditional probability-based score matching, and equipercentile linking. In reviewing the results and in consideration of the purpose of this research, we deemed it most appropriate to use the equipercentile linking method. This approach has been used extensively in a variety of settings (e.g. the linking of the SAT test with the ACT test) and allowed us to identify the corresponding TOEFL iBT scores which would restrict/pass a similar percentage of the students in the sample as would the IELTS scores. Given the need for both tests to screen out applicants in a comparable manner, we believe this to be the most defensible and straightforward approach.

Results

Descriptive Statistics of the Sample

TOEFL iBT scores range from 0 to 30 for each section and 0 to 120 for the total test; the total test score is the sum of the four sections. IELTS reports both section and total scores on a 9-band scale in one-half (0.5) band increments; the total test score is the average of the four sections. This general information about the different score scales of the two tests can help put the scores of the research sample in perspective. For example, because of the compact IELTS score scales, IELTS scales will show much smaller score variation (standard deviation values in Table 1) than the TOEFL iBT scales. Also, for some IELTS scores in Tables 3 to 7, the corresponding TOEFL iBT scores will show score ranges due to the larger TOEFL scales.

Table 1 gives the sample size, the mean scores, and the score variations (standard deviation, score range) of the two tests. On IELTS, the Reading and Listening mean scores were both 6.8, and the Speaking and Writing means were 6.3 and 6.1, respectively. The total test mean score was 6.6. On the TOEFL iBT scale, the Reading and Listening mean scores were both close to 21 and the Speaking and Writing mean scores were 20 and 21.6, respectively. The total mean score was 83.6.

To evaluate the overall performance of the students in this research sample, the 2009 test performance information on the two tests was obtained from the respective websites and is presented in the column for 2009 population mean in Table 1.

Generally speaking, the students in this sample, as a group, scored higher on both tests than their respective populations did in 2009. In other words, this sample appeared to be more able than the general populations of IELTS and TOEFL iBT test takers, respectively. This could be the outcome of self selection of the participants in that students having higher test scores might have been more willing to report their scores than those having lower scores. As a result, there were very few low scores in the data but relatively more scores at the middle to mid-high levels. Our level of confidence in the results for the low score levels would then be relatively lower than that for the middle to mid-high levels.

Table 1. Means and Standard Deviations

Test	N	Mean	Standard Deviation	Score Range	2009 Population Mean*
IELTS Total	1,153	6.6	0.8	3.5 - 9	5.8
IELTS Listening	1,153	6.8	1.1	2.5 - 9	6
IELTS Speaking	1,153	6.3	0.9	1.5 - 9	5.7
IELTS Reading	1,153	6.8	1.1	2.5 - 9	5.8
IELTS Writing	1,153	6.1	0.8	2 - 9	5.4
TOEFL Total	1,153	83.6	20.3	23 - 119	79
TOEFL Listening	1,153	20.9	7.6	1 - 30	19.4
TOEFL Speaking	1,153	20.0	3.8	5 - 30	19.7
TOEFL Reading	1,153	21.2	7.8	1 - 30	19.9
TOEFL Writing	1,153	21.6	4.6	5 - 30	20.5

*The 2009 population mean scores based on the information provided at the TOEFL and IELTS websites.

Table 2 lists the correlation coefficients between the scores from the same measures as well as the scores from the total tests. A correlation coefficient indicates the extent to which the scores on the two tests are related to each other. A correlation coefficient goes from 0 for no relation at all to 1 for a perfect relation. The correlation between total scores will always be higher than that between section scores because total scores contain all the section score information and are more reliable. As expected, the total test score correlation is 0.73, higher than any section score correlation, which varies from 0.44 for Writing to 0.68 for Reading. Generally speaking, for the scores to be linked, moderately high correlations (around 0.7) would be very desirable, but moderate correlations (around 0.5 to 0.6) are not atypical.

Table 2. Correlations between IELTS and TOEFL iBT Section and Total Scores

Score	Correlation
IELTS Listening & TOEFL Listening	0.63
IELTS Speaking & TOEFL Speaking	0.57
IELTS Reading & TOEFL Reading	0.68
IELTS Writing & TOEFL Writing	0.44
IELTS Total & TOEFL Total	0.73

The next sections provide the equipercentile linking results for each of the sections and the total score. The results are presented in tabular form. While we were able to obtain corresponding scores for the two tests at each of the IELTS score bands, the number of cases per score band varied, with generally very few cases in the very lowest and highest score bands. In general, as sample sizes increase, results become more stable and reliable. Consequently, in this study, we have more confidence in the results where there are relatively larger case counts (i.e., 5% of the sample) per score band and we recommend caution in interpreting the results where the case counts are relatively small. We have noted, by shading the area of the range of scores on the tables, where sample sizes per score band are greater than 5% of the total.

Listening Section Score Comparison Results

Table 3 presents the equipercentile linking method based score comparison results on the two Listening sections. The shaded area in Table 3 indicates that most students had scores from 5.5 to 8.5 on IELTS and 7 to 29 on TOEFL iBT test, with very few students (less than 5% or so) at the other score levels. As was mentioned earlier, due to the larger scale of TOEFL scores than IELTS scores, some IELTS Listening scores are found to correspond to a range of TOEFL scores instead of a single TOEFL score. For example, an IELTS Listening score of 5.5 would correspond to TOEFL iBT Listening scores of 7 to 11. In summary, the tabulated results indicate that an IELTS Listening score of 6.0 would correspond to TOEFL Listening scores of 12 to 19; an IELTS Listening score of 6.5 would correspond to TOEFL Listening scores of 20 to 23; an IELTS Listening score of 7.0 would correspond to TOEFL Listening scores of 24 to 26; and so forth. In all cases, the IELTS scores and their corresponding TOEFL scores (or score intervals) would pass similar percentages of test takers.

Table 3. Listening Section Score Comparisons

IELTS Score	TOEFL Score
9	30
8.5	29
8	28
7.5	27
7	24-26
6.5	20-23
6	12-19
5.5	7-11
5	4-6
4.5	3
<=4	0-2

Speaking Section Score Comparison Results

Table 4 presents the score comparison results for the Speaking sections. The information can be interpreted in the same way as described above for Listening. As shown by the shaded area in Table 4, most scores were within the range of 5.0 to 8.0 on the IELTS test and 14 to 27 on the TOEFL iBT test. Again, a number of IELTS scores correspond to TOEFL score ranges instead of single scores due to scale differences between the two tests. For example, the table shows that an IELTS score of 5.0 would correspond to TOEFL Speaking scores of 14 to 15; an IELTS Speaking score of 5.5 would correspond to TOEFL Speaking scores of 16 to 17; an IELTS Speaking score of 6.0 would correspond to TOEFL Speaking scores of 18 to 19; an IELTS Speaking score of 6.5 would correspond to TOEFL Speaking score of 20 to 22; and so forth. In all cases, the IELTS scores and their corresponding TOEFL scores (or score intervals) would pass similar percentages of test takers.

Table 4. Speaking Section Score Comparisons

IELTS Score	TOEFL Score
9.0	30
8.5	28-29
8.0	26-27
7.5	24-25
7.0	23
6.5	20-22
6.0	18-19
5.5	16-17
5.0	14-15
4.5	12-13
<=4	0-11

Reading Section Score Comparison Results

Table 5 presents the score comparison results for the Reading sections. The information can be interpreted in the same way as described above for Listening and Speaking. As shown by the shaded area in Table 5, most scores were within the range of 5 to 8.5 on the IELTS test and 4 to 29 on the TOEFL iBT test. The table shows that an IELTS Reading score of 5.0 would correspond to TOEFL Reading scores of 4 to 7; an IELTS Reading score of 5.5 would correspond to TOEFL Reading scores of 8 to 12; an IELTS Reading score of 6.0 would correspond to TOEFL Reading scores of 13 to 18; an IELTS Reading score of 6.5 would correspond to TOEFL Reading scores of 19 to 23; an IELTS Reading score of 7.0 would correspond to TOEFL Reading scores of 24 to 26; and so forth. Note that a TOEFL Reading score of 29 would correspond to both IELTS Reading scores of 8 and 8.5. This occurred because we wanted all IELTS band score points to be represented in the table. Since the highest TOEFL iBT score (30) has to map to the highest IELTS band score (9) and percentile data suggest that an TOEFL iBT score of 29 maps to an IELTS band score of 8, then it has to be the case that an TOEFL iBT score of 29 also maps to 8.5, i.e., there is no TOEFL iBT score point between 29 and 30 to map to the IELTS band score of 8.5. In all cases, the IELTS scores and their corresponding TOEFL scores (or score intervals) would pass similar percentages of test takers.

Table 5. Reading Section Score Comparisons

IELTS Score	TOEFL Score
9	30
8.5	29
8	29
7.5	27-28
7	24-26
6.5	19-23
6	13-18
5.5	8-12
5	4-7
4.5	3
<=4	0-2

Writing Section Score Comparison Results

Table 6 presents the score comparison results for the Writing sections. The information can be interpreted in the same way as described above for Listening, Reading and Speaking. As shown by the shaded area in Table 6, most scores were within the range of 5 to 7 on IELTS and 14 to 28 on the TOEFL iBT test. The table shows that an IELTS Writing score of 5.0 would correspond to TOEFL Writing scores of 14 to 17; an IELTS Writing score 5.5 would correspond to TOEFL Writing scores of 18 to 20; an IELTS Writing score of 6.0 would correspond to TOEFL Writing scores of 21 to 23; an IELTS Writing score of 6.5 would correspond to TOEFL Writing scores of 24 to 26; and so forth. In all cases, the IELTS scores and their corresponding TOEFL scores (or score intervals) would pass similar percentages of test takers.

Table 6. Writing Section Score Comparisons

IELTS Score	TOEFL Score
9.0	30
8.5	30
8.0	30
7.5	29
7.0	27-28
6.5	24-26
6.0	21-23
5.5	18-20
5.0	14-17
4.5	12-13
<=4.0	0-11

Total Score Comparison Results

Table 7 presents the comparison results for the total scores on the two tests. The comparison of total scores can be interpreted in the same way as described above for the section scores. As shown by the shaded area Table 7, most scores were within the range of 5.5 and 7.5 on the IELTS test and 46 and 109 on TOEFL iBT test. Less than 5% of test takers at each score or score interval received scores outside of this range. Because of the much larger TOEFL total score scale (0 to 120) than IELTS (0 to 9), the TOEFL score range corresponding to an IELTS score tends to be very broad. For example, an IELTS score of 5.5 would correspond to TOEFL scores of 46 to 59; an IELTS score of 6.0 would correspond to TOEFL scores of 60 to 78; an IELTS score of 6.5 would correspond to TOEFL scores of 79 to 93; an IELTS score of 7.0 would correspond to TOEFL scores of 94 to 101; and so forth. In all cases, the IELTS scores and their corresponding TOEFL scores (or score intervals) would pass similar percentages of test takers.

Table 7. Total Score Comparisons

IELTS Score	TOEFL Score
9	118-120
8.5	115-117
8	110-114
7.5	102-109
7	94-101
6.5	79-93
6	60-78
5.5	46-59
5	35-45
4.5	32-34
0-4	0-31

Discussion

To better understand the score comparison results on both the section and the total test levels, it is important to evaluate the results in the context of the characteristics of the tests, the research sample employed, the intended use of the results, and the limitations of the study. It is also worthwhile to think about what further research can be conducted in the future to enhance the quality of the score comparisons.

Evaluation of the Score Comparison Results and Limitations

There are two considerations regarding the interpretation of the results. The first is the characteristics of the two tests. It is recognized that the TOEFL and IELTS tests, while they have the same number of sections measuring similar skills and underlying constructs, were built from different frameworks and different test blueprints. For example, the TOEFL iBT test is the latest version of the TOEFL test and is built on the framework of communicative competency based on current second language acquisition theory and practice. The TOEFL iBT test contains numerous innovative task types that can only be presented on a computer-based test and tap into a wider array of English proficiency skills than a paper-based test. As a result, TOEFL iBT scores do not mean exactly the same thing as IELTS scores, and moderate correlations² between the scores, as seen in Table 2, can be expected. As previously noted, for the scores to be linked, moderately high correlations (around 0.7) would be very desirable, but moderate correlations (around 0.5 to 0.6) are not atypical.

The second consideration is the sample. As in many other studies of this nature, a self-selected group of participants volunteered to report their IELTS and TOEFL scores to ETS, which would limit generalizing the results to samples that differ significantly from this research sample. In particular, the numbers of participants with relatively low and very high scores were small. Future samples that include more cases with lower and higher scores would be recommended if an understanding of the mapping of scores at these score levels is desired. While the present case counts are sufficient to evaluate the mapping in the middle of the score scales, caution should be exercised when comparing the lower and higher scores. As described previously, the tables include shading to indicate those score levels where we have relatively more confidence in the score comparability.

² Correlation can be used as an index that is used in evaluating the potential for linking the scores from different tests that are intended to measure similar constructs but are constructed according to different frameworks.

Future Research

The score comparison results from this research were grounded in the empirical information that was obtained from the students who took both IELTS and TOEFL iBT tests, and were based on an appropriate psychometric procedure that used scores from the two tests to establish the relationships between the two sets of scores. As in any other score comparison research, the current research and the findings for the score comparison between the IELTS and TOEFL iBT tests will benefit from more research in the future, especially from future research studies based on more controlled samples in terms of sample size, sample selection, data collection, and student background information such as nationality, native language, years of learning English, etc. With a larger and better controlled sample, the data collection quality will be improved and a sufficient number of students across all scores will be provided to enhance the score comparison results.

Reference

- 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.
- ETS. (2002). *ETS Standards for quality and fairness*. Princeton, NJ: ETS.
- Holland, P. W. (2007). A framework and history for score linking. In N. J. Dorans, M. Pommerich, & P. W. Holland (Eds.), *Linking and aligning scores and scales* (pp. 5-30). New York, NY: Springer.
- Holland, P. W., & N. J. (2006). Linking and equating. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 187-220). Westport, CT: Praeger Publishers.
- International Language Testing Association. (June 2007). ILTA Guideline for Practice. Retrieved May 12, 2010 from http://www.iltaonline.com/index.php?option=com_content&view=article&id=122&Itemid=133
- Kolen, M. J., & Brennan, R. L. (1995). Test equating methods and practices. New York, NY: Springer.
- Pommerich, M., Hanson, B. A., Harris, D. J., & Sconing, J. A. (2000). *Issues in creating and reporting concordance results based on equipercentile methods* (ACT Research Rep. No. 2000-1). Iowa City, IA: ACT.
- Pommerich, M. (2007). Concordance: The Good, the Bad, and the Ugly. In N. J. Dorans, M. Pommerich, & P. W. Holland (Eds.), *Linking and aligning scores and scales* (pp. 199-216). New York, NY: Springer.
- Wang, L., Eignor, D.R., & Enright, M.K. (2008). A Final Analysis. In C. A. Chapelle, M. K. Enright, & J. Jamieson (Eds.), *Building a Validity Argument for a new Test of English as a Foreign Language* (pp. 259-318). New York, NY: Routledge.

TOEFL iBT Standard Setting for the Health Professions

The importance of establishing score requirements in the Health Professions

English proficiency is an important part of determining a student's readiness to enter an academic degree program. Students who have insufficient English ability will not benefit fully from their courses, may slow down the rest of the class, or could ultimately fail in the program. For a university, as well as for students and their families, this is an undesirable outcome, resulting in financial loss, lost opportunities and disappointment.

In the health professions, however, English proficiency plays a critical role as errors caused by insufficient language ability can be life-threatening, a far higher level of consequence than making a poor admission decision. It is for this reason that ETS strongly encourages the health professions to determine minimum score requirements for the TOEFL iBT test via standard setting.

What is standard setting?

Standard setting does not ask concern itself with establishing score equivalents from one test to another. Instead, it asks the more important question, "What level of English proficiency is necessary to carry out the responsibilities of this profession?" Using a rigorous, structured process over the course of two days, representatives of the profession determine the minimally acceptable levels of English needed and the associated TOEFL scores. Every profession has unique language needs. A standard setting meeting ensures the minimum language requirements suit the unique needs of the profession.

Panel composition

The composition of the standard setting panel is important, as it should represent a variety of stakeholders in the process of certification and licensing for the profession. For example, during a standard setting meeting for nurses, the panel included both native and non-native speakers of English, RNs, LPNs, a regulator and nurses from a wide variety of fields of practice. Participants need to be familiar with the English language demands of licensed professionals in their field. There should be 12-20 individuals on the panel.

Pre-meeting assignment

Prior to the meeting, panel members complete an assignment to help them begin thinking about the tasks that licensed professionals in their field need to perform and the English

language skills they need. For each skill (Reading, Writing, Listening, and Speaking) panel members identify tasks that are an important part of job responsibilities. For example, for Speaking, obtaining a patient's health history is a task nurses can be expected to do.

The panel member then is asked to think of individuals they have met or have worked with whose native language is not English. Some individuals may have very strong speaking skills while others may have difficulty expressing their ideas orally. The panel member should think of one person who was a **strong speaker** and one person who was a **weak speaker** and write the reasons why he/she is a strong or a weak speaker.

Developing a definition of minimally acceptable language skills

The standard setting meeting takes two full days to conduct, with the skills of Writing and Speaking considered on Day 1 and Listening and Reading discussed on Day 2. Using the information from the pre-meeting assignment, panel members consider the tasks a professional is expected to perform for each skill, and create (for example, for Speaking) a definition of a weak speaker, a strong speaker and a minimally acceptable speaker. With this definition in mind, panel members then review sample questions with responses at various levels and judge whether test takers at particular score levels display sufficient proficiency to practice in their field. Panel members make their judgments; the facilitators calculate the results and then lead a discussion of the results. After the discussion, a second round of voting takes place. The tally from this second review becomes the panel recommendation for the minimally acceptable level of English for that skill.

Arranging a standard setting meeting

In the US, the TOEFL Program has conducted standard setting meetings for a variety of health professions. Catherine Moore, ELT consultant to the TOEFL Program in Australia, is trained to lead standard setting meetings, and participated in the meeting for the US National Association of Boards of Pharmacy.

A Minimum English Proficiency Standard for The Test of English as a Foreign Language™ Internet-Based Test (TOEFL® iBT)

Anne Wendt, PhD, RN, CAE, PMP
Director, NCLEX® Examinations

Ada Woo, PhD
Psychometrician, NCLEX® Examinations



National Council of State Boards of Nursing

Introduction

In 2009, one of the initiatives of the National Council of State Boards of Nursing (NCSBN®) was to determine a passing standard for nurses on the Test of English as a Foreign Language Internet Based Test (TOEFL® iBT). Previously, NCSBN had recommended an English proficiency passing standard for TOEFL and the International English Language Testing System (IELTS®). This study provides a specific recommendation for TOEFL iBT because the Internet-based test adds a speaking component to TOEFL. In keeping with the *Standards for Educational and Psychological Testing*, the procedures and rationale used to set a passing standard should be documented (American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 1999). This research brief explains the process and the NCSBN passing standard recommendation to member boards of nursing.

TOEFL® iBT Examination

TOEFL, produced by Educational Testing Service (ETS), is an examination designed to measure English language proficiency among non-native English speakers in the academic setting. It is administered internationally in over 180 countries and is accepted as an English proficiency test among academic institutions (ETS, 2009). Two versions of the TOEFL examination are administered in the U.S.: the Paper-based Test (PBT) and the Internet-based Test (iBT). Currently, a majority of TOEFL tests administered are Internet-based, however, the PBT is available as a supplement to TOEFL iBT in a limited number of locations (ETS, 2009).

The TOEFL iBT measures reading, listening, speaking and writing skills. Test takers are allowed a total of four hours to complete the test. For the reading section, the examinee reads three to five passages of approximately 700 words each and then answers 12 to 14 questions for each passage. The reading section takes 60 to 100

minutes. For the listening section, the examinee listens to four to six classroom lectures and two to three conversations that are three to five minutes in length. After listening to a segment, examinees answer five to six questions about segment content. The listening section is administered over 60 to 90 minutes. The speaking section consists of six tasks, including two independent tasks for the examinee to express an opinion on a familiar topic and four integrated tasks based on what the examinee reads or hears on the test. The examinee has up to 30 seconds to prepare each response and up to one minute to respond. The entire speaking section takes 20 minutes. Lastly, the writing section consists of one integrated task and one independent task. The examinee is allowed 20 minutes to complete the integrated task and 30 minutes for the independent task. The examinee can receive a maximum of 30 scaled points on each of the four sections. The section-level subscores (zero to 30 points each) and a total TOEFL iBT score (zero to 120) are reported.

Selection of Standard Setting Panelists

The composition of a standard setting panel is a key element in determining the validity of the standard setting process. Twenty-two subject matter experts from the four NCSBN geographic areas served on the standard setting panel. The panel was selected to include: (1) nurses who speak a primary language other than English and have taken the TOEFL; (2) nurses who work with clients that speak a primary language other than English; (3) clinical nurse supervisors who speak a primary language other than English; (4) a nursing regulator; (5) a nursing educator; (6) a consumer of nursing services; and (7) registered nurse (RN) and licensed practical/vocational nurse (LPN/VN) representatives. In all, the standard setting panel was quite diverse, representing 17 states, seven foreign languages and various nursing specialty and practice settings. Detailed panelist demographic characteristics are listed in Table 1.

Table 1. Demographic Characteristics of Standard Setting Panel Members			
Panelist Characteristic	Category	%	(N)
Gender	Female	86%	19
	Male	14%	3
License	RN	86%	19
	LPN/VN	14%	3
NCSBN geographic region	Area I	18%	4
	Area II	41%	9
	Area III	18%	4
	Area IV	23%	5
Language spoken other than English	Spanish	18%	4
	Igbo	14%	3
	Filipino	14%	3
	Chinese	9%	2
	Arabic	5%	1
	Russian	5%	1
	Korean	5%	1
Has taken TOEFL	Yes	50%	11
	No	50%	11
Job title	Staff nurse	59%	13
	Supervisor	14%	3
	Board of nursing staff	5%	1
	ESL instructor	5%	1
	Family nurse practitioner	5%	1
	Instructor	5%	1
	Nurse education manager	5%	1
	Recruitment specialist	5%	1
Primary specialty ¹	Medical/Surgical	33%	7
	Obstetrics	19%	4
	Pediatrics	10%	2
	Child/Adolescent psychiatric	5%	1
	Critical Care	5%	1
	Geriatric	5%	1
	Hospice	5%	1
	Neonatal intensive care	5%	1
	Not reported	14%	3
Practice setting ¹	Hospital	57%	12
	Outpatient facility	14%	3
	Home health	10%	2
	Academic program	5%	1
	Administration	5%	1
	County health department	5%	1
	Day care	5%	1

¹The standard setting panel contained one public member who was not a nurse. The percentages in this category were calculated based on 21 panelists.

Panelist Orientation and Training

Another important element in the standard setting process is to ensure that the members of the panel are adequately prepared for the standard setting task. Thus, a set of preparatory materials was sent to panelists approximately two weeks prior to the standard setting workshop. The materials included a description of the TOEFL iBT, the agenda of the standard setting workshop, the NCLEX-RN® and NCLEX-PN® test plans, and a preworkshop homework assignment. Panelists were asked to read the materials and complete the homework assignment prior to attending the standard setting workshop. The homework was divided into four sections with each addressing a different language skill measured by TOEFL iBT: reading, listening, speaking and writing. For each section, panelists considered entry-level nursing activities that require the language skill being measured. In addition, panelists were instructed to reflect on their experiences with non-native English speakers and to think of an example of someone with good English skills and someone with poor English skills. The panelists then recorded skills that their examples illustrated as part of their preparation for the workshop.

Workshop facilitators oriented panelists to the goals and purpose of the three-day workshop by describing the TOEFL iBT and the standard setting process. Additionally, the minimally competent candidate concept was explained to panelists. In the presentation, facilitators emphasized that the responsibility of the panelists was to set a minimum English proficiency standard for nurses to practice at entry level in the U.S., not a minimum English proficiency standard for passing the NCLEX®. Panelists were encouraged to ask questions throughout the orientation. The panelists' evaluation at the conclusion of the workshop was positive and indicated that they felt prepared for their standard setting task (Wendt, Woo, & Kenny, in Press).

The facilitators also noted that it was important to come to consensus on the language demand that was being considered before the start of the standard setting process. After the orientation, facilitators asked panelists to discuss whether the English language requirements are the same for entry-level RNs and LPN/VNs. The panel, consisting of both RNs and LPN/VNs, agreed that the language demand is the same for RNs and LPN/VNs, despite the differences in scope of practice.

Standard Setting Procedures

Listening and Reading Sections

There were two standard setting procedures used in the standard setting workshop: the Yes/No method and the

Examinee Paper Selection method. To apply the Yes/No method each panelist was asked to imagine a nurse with minimally acceptable English skills and estimate whether such a nurse would be able to correctly answer a sample of items. This method was introduced by Impara and Plake (1997) as a variation of the modified Angoff method. In a modified Angoff standard setting, panelists estimate the probability that a minimally competent candidate would be able to obtain a correct answer to a particular question. Unlike the modified Angoff, the Yes/No method lessens the demand on panelists by asking them for a Yes/No response rather than to estimate probabilities. The Yes/No method was used to set standards for the listening and reading sections.

At the start of the listening section, panelists discussed and developed a consensus on minimally acceptable English listening skills for entry-level nurses in the U.S. As part of the discussion, the panelists defined a list of listening tasks that entry-level nurses are required to perform in order to practice nursing safely and effectively (Appendix A). This discussion was facilitated by the homework assignment that panelists completed prior to attending the workshop. After the minimally required listening skills and nursing tasks were defined, a sample listening segment and associated items were presented to the panelists. Panelists practiced the Yes/No rating procedure on the sample items. After the panelists became familiar with the TOEFL iBT listening section and the Yes/No method, they provided independent judgments on an actual TOEFL iBT listening section. For each item that was reviewed, facilitators asked panelists to consider whether an entry-level nurse with a minimally acceptable listening ability would answer the item correctly. The items for this section were drawn from retired TOEFL iBT tests. Panelists were provided two rounds of ratings with group discussion between each round.

The procedures used to set a standard for the reading section were very similar to those used for the listening section except that the items reviewed were written instead of verbal.

Writing and Speaking Sections

Unlike the listening and reading sections where examinees answer closed-ended items with definite right and wrong answers, the writing and speaking sections require examinees to generate open-ended responses that are graded by raters. Therefore, instead of using the Yes/No Method throughout the standard setting workshop, the Examinee Paper Selection method was used to set standards for the speaking and listening sections (Hambleton, Jaegar, Plake & Mills, 2000). The Examinee Paper Selection method requires panelists to read scoring

rules for each sample item that describe the elements that constitute a point for that item. Panelists then read examinee responses that correspond to each point of the scoring rules. After reviewing examinee responses that span different score points for an item, panelists chose a response that they believed would most likely be given by an internationally-educated, entry-level nurse with minimum English language skills. If the panelists believed that the minimally competent candidate would likely respond somewhere between two score levels, they were allowed to use half points.

At the start of the speaking section, panelists discussed and developed a consensus regarding their concept of the minimally competent speaker. They also defined a list of speaking tasks entry-level nurses need to perform in order to safely and effectively practice nursing in the U.S. (Appendix A). Preworkshop homework assignments facilitated the discussion and development of the task list shown in Appendix A. As previously noted, after the minimally acceptable speaking skills and tasks were developed, panelists were oriented to the TOEFL iBT speaking section. Facilitators familiarized the panelists with the section's format and score levels. To set a minimally acceptable standard for the speaking section, panelists reviewed six speaking tasks and their respective scoring rules. Twelve response profiles of varying speaking score levels from past TOEFL iBT candidates were presented to panelists as examples. Panelists were allowed to playback any response profile during the standard setting process. After the presentation of each speaking task and the corresponding scoring rules, panelists listened to responses from the 12 examinees on each of the six speaking tasks. At the conclusion of reviewing all six tasks, panelists were asked to independently consider whether each of the 12 example examinees demonstrated the minimally acceptable speaking skills required for safe and effective entry-level nursing practice. The lowest score level at which panelists thought an example examinee had the minimal speaking skills then represented their recommended standard for the speaking section. Panelists were allowed to provide two rounds of recommendations, with discussion between each round.

Procedures for the writing and speaking sections are similar except the writing section included writing skills rather than speaking skills.

Panel Recommendations

For all four TOEFL iBT sections, panelists were able to come to a consensus on a recommended standard. It is important to note that the social dynamic among panelists and facilitators was one of collaboration. After the workshop, all participants noted that they found the experience valuable and appreciated the ample opportunities they were given to express their opinion. In all TOEFL iBT skill categories, round two recommendations from the panel were more similar than those of the round one. This increased similarity suggested convergence of individual recommendations as the workshop progressed. No drastic outliers were found among second round recommendations for all four sections. This research brief will report these recommendations in aggregate form.

As indicated in the score interpretation guidelines provided by ETS and presented in Table 2, each TOEFL iBT section measures a range of English skill levels (ETS, 2009). Panel recommendations for the four sections are summarized in Table 3. In addition to analyzing recommendations from the panel as a whole, recommendations from panelist subgroups were also analyzed. These subgroups included: (1) panelists who have taken the TOEFL; (2) panelists who work with non-English speaking clients; (3) panelists who are clinical supervisors; and (4) panelists who are regulators, educators or members of the public. As can be seen in Table 3, the minimally acceptable skill levels recommended by the four panelist subgroups were remarkably consistent. In three out of the four TOEFL iBT skills, all subgroups recommended the same level of minimally acceptable performance. Panelists who have taken the TOEFL previously and those who work with non-English speaking clients recommended a slightly lower writing skill level than the other subgroups. Overall, the standard setting panel recommended intermediate levels for the reading and listening sections, and a fair level for the speaking and writing sections. The average of the panel's recommended English proficiency levels are shaded gray in Table 3.

Table 2. TOEFL iBT Score Scale

Skill	Score Range	Level
Reading	0-30	Low (0-14)
		Intermediate (15-21)
		High (22-30)
Listening	0-30	Low (0-14)
		Intermediate (15-21)
		High (22-30)
Speaking	0-4 points converted to 0-30 score scale	Weak (0-9)
		Limited (10-17)
		Fair (18-25)
		Good (26-30)
Writing	0-5 points converted to 0-30 score scale	Limited (0-16)
		Fair (17-23)
		Good (24-30)

Table 3. Summary of Panel Recommendations by Subgroups

Test Section	Mean recommendation by all panelists (N=22)	Mean recommendation by panelists who have taken TOEFL (N=11)	Mean recommendation by panelists who work with non-English speaking clients (N=8)	Mean recommendation by panelists who are clinical supervisors (N=3)	Mean recommendation by panelists who are regulators, educators or public members (N=3)
Reading	Intermediate	Intermediate	Intermediate	Intermediate	Intermediate
Listening	Intermediate	Intermediate	Intermediate	Intermediate	Intermediate
Speaking	Fair	Fair	Fair	Fair	Fair
Writing	Fair	Limited	Limited	Fair	Fair

NCLEX® Examination Committee (NEC) Deliberation

The NCSBN Board of Directors charged the NEC with developing a recommended minimum TOEFL iBT passing standard for entry-level nurses. As noted, the recommended passing standards from panelists were consistent across the four English language skills: reading, listening, speaking and writing. The average panel recommendations provided critical information for committee deliberation. In addition to

results of the standard setting process, the committee considered information from a variety of sources including existing U.S. visa-screening requirements for allied health professionals (Table 4) and normative TOEFL iBT performance data from applicants for professional licensure.

Table 4. Current TOEFL iBT Standards for VisaScreen® (CGFNS, 2005)

Total Score	Speaking Score	Professional Group
79	26	Current standard for clinical laboratory technicians, (medical technicians) and LPN/VNs
83	26	Current standard for RNs , clinical laboratory scientists (medical technologists), speech language pathologists, audiologists and physician assistants
89	26	Current standard for occupational therapists and physical therapists

As seen in Table 4, the current VisaScreen® TOEFL requirements for other allied health professions are consistent with recommendations from the standard setting panel. The committee also reviewed 2007 test score information from the TOEFL iBT in order to gain information on the impact of a recommended passing score. The test scores of TOEFL examinees who reported that they took the test to obtain professional licensure were considered. The committee chose this sub-population of TOEFL examinees because their educational level would be similar to that of entry-level nurses. After weighing all available evidence, including their in-depth knowledge about the nursing profession, the NEC recommended that the minimally acceptable TOEFL iBT scores for entry-level nurses be an overall score of 84 with a minimum speaking score requirement of 26. The recommended minimum overall score is one point higher than the TOEFL passing standard endorsed by NCSBN in 2005.

Comparison with Minimum IELTS® Passing Standard for Entry-level Nurses

In 2005, NCSBN recommended a minimally acceptable English proficiency standard for entry-level nursing using the IELTS test (NCSBN, 2005). Similar to the TOEFL, the IELTS is designed to assess English skill levels of non-native English speakers. IELTS is mostly used in Great Britain, Ireland, New Zealand and Australia. This test is designed to measure general English, not just North American English. IELTS covers listening, reading, writing and speaking and measures an examinee's proficiency levels using band scores that range from 0 to 9 (Table 5). Using a standard setting process similar to the one described in the present paper, the NEC recommended a 6.5 overall IELTS band score with a minimum of 6.0 in any one IELTS module for entry-level nurses (O'Neill, Buckendahl, Plake & Taylor, 2007).

Table 5. IELTS Band Scores

Band 9 – Expert User
Has fully operational command of the language: appropriate, accurate and fluent with complete understanding.
Band 8 – Very Good User
Has fully operational command of the language with only occasional and unsystematic inaccuracies and inappropriateness. Misunderstandings may occur in unfamiliar situations. Handles complex detailed argumentation well.
Band 7 – Good User
Has operational command of the language, though with occasional inaccuracies, inappropriacies, and misunderstandings in some situations. Generally handles complex language well and understands detailed reasoning.
Band 6 – Competent User
Has generally effective command of the language despite some inaccuracies, inappropriacies, and misunderstandings. Can use and understand fairly complex language, particularly in familiar situations.
Band 5 – Modest User
Has partial command of the language, coping with overall meaning in most situations, though is likely to make many mistakes. Should be able to handle basic communication in own field.
Band 4 – Limited User
Basic competence is limited to familiar situations. Has frequent problems in understanding and expression. Is not able to use complex language.
Band 3 – Extremely Limited User
Conveys and understands only general meaning in very familiar situations. Frequent breakdowns in communication occur.
Band 2 – Intermittent User
No real communication is possible except for the most basic information using isolated or short formulae in familiar situations and to meet immediate needs. Has great difficulty in understanding spoken and written English.
Band 1 – Non User
Essentially has no ability to use the language beyond possibly a few isolated words.
Band 0 – Did not attempt the test
No assessable information provided

Common European Framework of Reference

The developers of IELTS and TOEFL have mapped both tests onto the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR) for ease of interpretation. The CEFR is a set of guidelines used to describe language proficiency of English speakers. It is a six-level framework that describes English speaker's achievements on listening, writing, speaking and reading. Levels A1 and A2 represent basic users, levels B1 and B2 represent independent users and levels C1 and C2 are proficient users. The CEFR also provides descriptive context by which to interpret the six performance levels.

According to the University of Cambridge ESOL Examinations, the developer of IELTS, a 6.5 overall IELTS band score corresponds to level C1 (proficient user) in the CEFR (IELTS, 2009). This is slightly higher than the NCSBN recommended standard for the TOEFL iBT. According to preliminary data provided by ETS, an overall TOEFL iBT score of 84 corresponds to level B1 (independent user) in the CEFR and the recommended speaking score of 26 falls somewhere between levels B2 and C1 on the CEFR scale (ETS, 2008).

Conclusion

NCSBN's recommendation regarding the English proficiency passing standard of 84 with a minimum speaking score of 26 for TOEFL iBT is consistent with the previous passing standard recommended for TOEFL, IELTS and the level of English proficiency needed to perform entry-level nursing safely and effectively. Additionally, this recommendation is consistent with the NCSBN policy position on International Nurse Immigration which states that "domestic and international nurses need to [be] proficient in written and spoken English skills" (NCSBN, 2009). Boards of nursing can use this information to make decisions regarding the level of English proficiency needed for entry-level RNs and LPN/VNs in their jurisdictions.



National Council of State Boards of Nursing, Inc.
111 E. Wacker Drive, Suite 2900, Chicago, IL 60601-4277
312.525.3600 | International Calls: +1.312.525.3688
Testing Services: 312.525.3750 | www.ncsbn.org

References:

- American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (1999). Standards for educational and psychological testing. Washington, D.C.: American Educational Research Association.
- CGFNS International. (2005). 2005 news. Retrieved June 26, 2009, from http://www.cgfns.org/sections/announcements/news/2005/11-22-05_english.shtml
- Educational Testing Service. (2008). Mapping TOEFL® iBT on the common European framework of reference. Retrieved June 26, 2009, from http://www.ets.org/Media/Research/pdf/CEF_Mapping_Study_Interim_Report.pdf
- Educational Testing Service. (2009). Test of English as a Foreign Language™ 2009-2010 information and registration bulletin for Internet-based Testing (iBT). Retrieved June 26, 2009, from http://www.ets.org/Media/Tests/TOEFL/pdf/toefl_ibt_bulletin_09_10.pdf
- Hambelton, R.K., Jaeger, R.M., Plake, B.S., & Mills, C. (2000). Setting performance standards on complex educational assessments. *Applied Psychological Measurement*, 24, 355-366.
- International English Language Testing System. (2009). Common European framework. Retrieved June 26, 2009, from http://www.ielts.org/researchers/common_european_framework.aspx.
- Impara, J.C., & Plake, B.S. (1997). Standard setting: An alternate approach. *Journal of Educational Measurement*, 34, 353-366
- National Council of State Boards of Nursing. (2005). *NCSBN recommends English language proficiency standards for internationally-educated nurses*. Retrieved June 26, 2009, from <https://www.ncsbn.org/positionlangTemplate.pdf>
- National Council of State Boards of Nursing. (2009). *National Council of State Boards of Nursing (NCSBN) position statement international nurse immigration*. Retrieved July 31, 2009 from <https://www.ncsbn.org/1272.htm>
- O'Neill, T.R., Buckendahl, C.W., Plake, B.S., & Taylor, L. (2007). Recommending a nursing-specific passing standard for the IELTS examination. *Language Assessment Quarterly*, 4, 295-317.
- Wendt, A., Woo, A., & Kenny, L. (in Press). Setting a passing standard for English proficiency on Internet-Based Test of English as a Foreign Language. *JONA's Healthcare Law, Ethics and Regulation*, 11(3).

©2009 The National Council of State Boards of Nursing (NCSBN®) is a not-for-profit organization whose members include the boards of nursing in the 50 states, the District of Columbia and four U.S. territories — American Samoa, Guam, Northern Mariana Islands and the Virgin Islands. The College of Registered Nurses of British Columbia is an associate member.

Mission: The National Council of State Boards of Nursing (NCSBN), composed of Member Boards, provides leadership to advance regulatory excellence for public protection.

Appendix A: Skills and Tasks Identified by TOEFL® iBT Standard Setting Panel Members

1. Reading Section:

Reading Tasks of Entry-Level Nurses:

- Read patient history/charts
- Read directions for medications
- Read procedures
- Read consultation notes
- Read bulletin boards
- Read shift to shift notes

Characteristics of an Entry-Level Nurse with Minimally Acceptable Reading Ability:

- Can identify main idea(s) all the time even if they need to seek help in understanding
- Can understand enough vocabulary and grammar to get and infer meaning
- Can distinguish important from less important information most of the time
- Can connect what is read to what they know (makes appropriate connections and inferences)
- Can identify main idea(s) all the time even if they need to seek help in understanding
- Understands tone of writer most of the time

2. Listening Section:

Listening Tasks of Entry-Level Nurses:

- Listen to and understand patients, families and their issues
- Understand verbal and telephone orders
- Understand telephone calls
- Understand shift report
- Evaluate patient's knowledge based on what they say
- Understand verbal paging system messages
- Understand staff at meetings and other interactions
- Understand lectures and service training

Characteristics of an Entry-Level Nurse with Minimally Acceptable Listening Ability:

- Can identify main points all the time
- Responds appropriately most of the time to the point of discussion
- Understands enough vocabulary (including numbers) and grammar to identify main points and important details
- Can understand meaning from context most of the time
- Most of the time can make appropriate connections
- Sometimes can recognize underlying emotion or intent from tone
- Can differentiate between important and less important information most of the time

3. Writing Section:

Writing Tasks of Entry-Level Nurses:

- Document response to treatment
- Take notes on patient assessment
- Document verbal orders
- Document patient's understanding of progress/condition
- Document care plans
- Document incident reports
- Document medication-averse reactions
- Document all aspects of care
- Send e-mails and texts of documents to supervisors

- Write transfer reports
- Write discharge instructions
- Give instructions for ancillary staff
- Write personalized notes
- Document initial assessments
- Document requests for release
- Communicate with police and other social services

Characteristics of an Entry-Level Nurse with Minimally Acceptable Writing Ability:

- Spelling is mostly correct; errors don't interfere with meaning
- Grammar is mostly correct; errors don't interfere with meaning
- Stays on topic; provides relevant details
- Order of documented events doesn't interfere with meaning
- Able to identify objective vs. subjective (fact vs. opinion)
- Appropriate word choice for audience

4. Speaking Section:

Speaking Tasks of Entry-Level Nurses:

- Communicate shift to shift report orally
- Introduce self to patient
- Speak to pharmacy to provide clarification
- Explain rights to patient
- Redirecting and reorienting patients
- Ask questions for self-learning/nurse learning
- Discuss interdisciplinary care plans
- Explain procedures
- Explain care to families and patients
- Speak to doctors
- Speak on the telephone
- Explain patient progress
- Give instruction to ancillary workers
- Directions for, and conversations with, families and visitors
- Teach other nurses
- Provide feedback on patient progress

Characteristics of an Entry-Level Nurse with Minimally Acceptable Speaking Ability:

- Grammar errors don't interfere with listener's understanding
- Pronunciation errors don't interfere with listener's understanding
- Can stay on topic, maintain focus
- Intonation, clarity, volume, pace and mannerisms do not interfere with listener understanding
- Vocabulary is appropriate for listener's context/situation
- Flow of information does not impede listener's understanding
- Able to express thoughts in at least simple, structured sentences
- Speech does not require excessive effort on the part of the listener