PARAGON RESEARCH REPORTS

Linking CAEL CE Scores to IELTS - Academic Scores: Full Report

December 27, 2017
Taylor Asbury

TESTING ENTERPRISES Paragon | UBC a subsidiary of THE UNIVERSITY OF BRITISH COLUMBIA

Contact Information

Paragon Testing Enterprises

110-2925 Virtual Way, Vancouver BC V5M 4X5

research@paragontesting.ca

Contents

List of Tables

1 — Introduction

The Canadian Academic English Language Assessment- Computer Edition (CAEL CE) is designed to measure the English language proficiency of students planning to study in Canadian post-secondary institutions. The CAEL CE Test provides an authentic representation of language use in a Canadian academic context. As would be expected in a first-year Canadian university or college classroom, test takers read articles, listen to a lecture, answer questions, and write a short essay. Each test taker receives a score report showing their performance on each component as well as an overall score that is the unweighted average of the four individual component scores.

CAEL CE scores are accepted by more than 170 post-secondary institutions in Canada as proof of the English language proficiency of applicants. These institutions also accept other English language proficiency tests, and there are occasions when comparisons need to be made between scores on different tests. To help university admissions offices use CAEL CE scores to evaluate applicants, Paragon conducted score-comparison research between CAEL CE and an alternative test, the International English Language Testing System - Academic (IELTS - Academic).

The aim of this research was to identify which IELTS - Academic component score(s) are comparable to each CAEL CE component score and which IELTS - Academic Overall score(s) are comparable to each CAEL CE Overall score. This was accomplished by what is known as score linking. The Method section provides more details about score linking as well as information about the dataset and the analyses. The Results section will present score comparisons for the four skill components and the overall scores. The final section, Discussion, considers the results in terms of their applicability to the test taking population as a whole, evaluates the limitations of the sample, and makes suggestions for future research.

2 — Method

2.1 Score Linking

Linking scores is basically a transformation from a score on one scale to a score on the other scale and can be accomplished in a variety of ways (Holland & Dorans, 2006). Two methods were considered: regression-based prediction and equipercentile linking. The equipercentile linking method with loglinear pre-smoothing (R package *equate*; Albano, 2016) was preferred in this context. Equipercentile equating, as the name suggests, identifies the scores for each measure which would restrict/pass a similar percentage of the participants in the sample. It has been used successfully in a variety of settings (cf. Dorans, 1999; Educational Testing Service, 2010) and was selected for the following reasons:

- Regression has direction; the functional relationship is one-way and the results are only predicted. With equipercentile linking, on the other hand, the relationship holds in both directions (e.g., a CAEL CE 80 is most closely related to an IELTS - Academic 8, and an IELTS - Academic 8 is most closely related to a CAEL CE 80).
- Predicted scores (via regression) are the same as concordance scores only when the two sets of scores are perfectly related. This condition cannot be assumed to be true in the case of two different measurement instruments and was not true in the case of this dataset.
- Equipercentile equating is preferred to regression when the sample size is small, another feature of this study.

2.2 Data Collection

Data collection took place in August and September, 2017. Scores were provided by 91 participants who had taken both tests and who submitted their IELTS - Academic score reports for verification. To minimize the effect of test order, some participants took IELTS - Academic followed by CAEL CE, and other participants took CAEL CE followed by IELTS - Academic.

Before analysis, the data was cleaned by removing records with missing scores on one or more CAEL CE components. An example of this would be a test taker who received scores for the Reading, Listening, and Writing components but who did not provide a rateable Speaking performance. After this data cleaning procedure, 88 complete records remained in the dataset.

Table 1 shows that, of the remaining 88 test takers, all were over the age of 17 and the majority were under 40 (78.41%). 54.55% were female and 45.45% were male. While the test takers represented 25 different first-language backgrounds, the sample primarily consisted of test takers whose first language was Chinese, Farsi, Arabic, English, Korean, and Spanish. Overall, the majority of the sample represents individuals seeking entry to post-secondary institutions for both undergraduate and graduate study.

Table 1: Distribution (in %) of Test Takers by Age

Age	Percentage
17–19	5.68
20–29	37.50
30–39	35.23
≥40	21.59

2.3 Data Analysis

To understand the statistical characteristics of the test takers' scores on the two tests, descriptive statistics were computed. This enabled us to evaluate the test takers' average scores on each test at the component and overall score level, as well as the variation (standard deviation) observed in the scores. The correlation between the test takers' scores on each test was also computed to check the relationship between the test scores.

Subsequently, the equipercentile linking method with loglinear pre-smoothing (R package *equate*; Albano, 2016) was applied to identify the IELTS - Academic scores that correspond to each CAEL CE score. This procedure was carried out for each component individually and for the overall score.

3 — Results

3.1 Descriptive Statistics of the Sample

Table 2 presents the mean scores and the score variations (standard deviation and score range) of the two tests. These figures should be interpreted in relation to the reporting scale for each test. The CAEL CE reporting scale ranges from 10 to 90 for each component and for the overall score (in increments of 10). IELTS - Academic reports both component and overall scores on a 9-band scale in half-band (0.5) increments. Since the IELTS - Academic score scales are compact (a maximum range of 8), IELTS - Academic scales will show smaller score variation (standard deviation values in Table 2) than the CAEL CE scales (a maximum range of 80).

Some truncation of the sample was observed, particularly at lower levels of proficiency. This is to be expected; test takers typically do not attempt an academic purposes test until they are at a level of language proficiency that is close to the level required for admission to post-secondary institutions.

Table 2: Score Means and Variations (*N* = 88)

Test	Mean	Standard Deviation	Score Range
IELTS - Academic Listening	6.6	1.1	4–9
IELTS - Academic Reading	6.5	1.1	4–9
IELTS - Academic Writing	5.9	0.7	4.5–7.5
IELTS - Academic Speaking	6.6	0.9	4–8.5
IELTS - Academic Overall	6.5	0.8	5–8
CAEL CE Listening	59.1	16.7	20–90
CAEL CE Reading	59.2	17.0	10–90
CAEL CE Writing	51.4	10.4	20–80
CAEL CE Speaking	53.5	13.0	30–70
CAEL CE Overall	57.3	12.7	30–80

Correlation coefficients were computed between the components of each test (for instance, the CAEL CE Reading scores were correlated with the IELTS - Academic Reading scores) as well as the overall scores. A correlation coefficient can range from 0 for no relation at all to 1 for a perfect relation (and to -1 for perfect negative relation).

A correlation coefficient can be used to confirm the appropriateness of linking the scores from different tests. When two measures are very alike (such as when they are two versions of the same test), we can expect the correlation to be very high, if not perfect. In the case of two tests that have been created independently of one another (as is the case for CAEL CE and IELTS - Academic), we can expect the differences in test design to result in a lower

correlation. Other factors such as test familiarity, fatigue, anxiety, and illness can also affect test takers' performance and the correlation between the tests. Generally speaking, when the scores from two different tests are being correlated, moderate correlations (around 0.5 to 0.6) are not atypical. However, if the scores on two tests are to be linked, moderately high correlations (around 0.7) are preferred.

Table 3 presents the correlation between each IELTS - Academic component and the corresponding CAEL CE component as well as the correlation for the overall scores of both tests. The component score correlations vary from 0.57 for Writing, to 0.72 for both Reading and Speaking. The correlation between the overall scores for both tests is 0.81. Note that the correlation between overall scores will always be higher than the correlation between component scores. This is because overall scores contain all the component score information and are therefore more stable.

Table 3: Correlations between IELTS - Academic and CAEL CE Component and Overall Scores	
Scores Compared	Correlation
IELTS - Academic Listening & CAEL CE Listening	0.67
IELTS - Academic Reading & CAEL CE Reading	0.72
IELTS - Academic Writing & CAEL CE Writing	0.57
IELTS - Academic Speaking & CAEL CE Speaking	0.72
IELTS - Academic Overall & CAEL CE Overall	0.81

3.2 Score Comparison Results

The next sections present the equipercentile linking results for each of the components and the overall score. The dataset provides corresponding scores for the two tests at the majority of both tests' score bands. For all components and the overall score, the CAEL CE scores and their corresponding IELTS - Academic scores would pass similar percentages of test takers. However, it is important to note that the number of cases per score band varied, with generally very few cases in the lowest and highest score bands. If a CAEL CE score was not observed in the sample, we could not identify the corresponding IELTS - Academic score. In the tables that follow, these scores are represented as a double dash (--). Special attention has been given to CAEL CE scores of 50–70 and IELTS - Academic scores of 6–7, as these scores are commonly required for admission to post-secondary institutions.

3.2.1 Listening Component

Table 4 presents the score comparison results for the two Listening components. More than one-third of participants (37.5%) had Listening component scores from 50 to 70 on CAEL CE and 6 to 7 on IELTS - Academic. The CAEL CE Listening score of 10 was not observed in this study, so the corresponding IELTS - Academic Listening score may not be identified. The table indicates that a CAEL CE Listening score of 50 would correspond to an IELTS - Academic Listening score of 6; a CAEL CE Listening score of 60 would correspond to an IELTS - Academic score of 6.5; a CAEL CE Listening score of 70 would correspond to an IELTS - Academic Listening score of 7; and so forth.

Table 4: Listening Component Score Comparisons	
CAEL CE Score	IELTS - Academic Score
90	9
80	8
70	7
60	6.5
50	6
40	5
30	5
20	4.5
10	--

3.2.2 Reading Component

Table 5 presents the score comparison results for the two Reading components. More than one-third of the Reading component scores (39.8%) were within the range of 50 to 70 on CAEL CE and 6 to 7 on IELTS - Academic. The table shows the IELTS - Academic Reading score that would correspond to a particular CAEL CE Reading score. Note that, for a CAEL CE Reading score of 10, the corresponding IELTS - Academic Reading score is 3.5; this value of 3.5 was not observed in this study but was extrapolated.

Table 5: Reading Component Score Comparisons	
CAEL CE Score	IELTS - Academic Score
90	9
80	8
70	7
60	6.5
50	6
40	5
30	5
20	4.5
10	3.5

3.2.3 Writing Component

Table 6 presents the score comparison results for the two Writing components. More than half of the Writing scores (51.1%) were within the range of 50 to 70 for CAEL CE and 6 to 7 for IELTS - Academic. The table shows the IELTS - Academic Writing score that would correspond to a particular CAEL CE Writing score. CAEL CE Writing scores of 10 and 90 were not observed in this study, so the corresponding IELTS - Academic Writing scores may not be identified. Note that, for a CAEL CE Writing score of 80, the corresponding IELTS - Academic score is reported to be 8.5; an IELTS - Academic Writing score of 8.5 was not observed in this study but was extrapolated.

Table 6: Writing Component Score Comparisons	
CAEL CE Score	IELTS - Academic Score
90	--
80	8.5
70	7.5
60	6.5
50	6
40	5
30	5
20	4.5
10	--

3.2.4 Speaking Component

Table 7 presents the score comparison results for the two Speaking components. Nearly one-half of the Speaking scores (47.7%) were within the range of 50 to 70 for CAEL CE and 6 to 7 for IELTS - Academic. The table shows the IELTS - Academic Speaking score that would correspond to a particular CAEL CE Speaking score. The CAEL CE Speaking scores of 10, 20, 80, and 90 were not observed in this study, so the corresponding IELTS - Academic scores may not be identified.

Table 7: Speaking Component Score Comparisons	
CAEL CE Score	IELTS - Academic Score
90	--
80	--
70	8
60	7
50	6.5
40	5.5
30	5
20	--
10	--

3.2.5 Overall Score

Table 8 presents the comparison results for the overall scores on the two tests. The comparison of overall scores can be interpreted in the same way as described above for the component scores. Nearly two-thirds of the Overall scores (62.5%) were within the range of 50 to 70 for CAEL CE and 6 to 7 for IELTS - Academic. The table shows the IELTS - Academic Overall score that would correspond to a particular CAEL CE Overall score. Overall scores of 10, 20, and 90 were not observed in this study, so the corresponding IELTS - Academic Overall scores may not be identified.

Table 8: Overall Score Comparisons	
CAEL CE Score	IELTS - Academic Score
90	--
80	8
70	7
60	6.5
50	6
40	5.5
30	4.5
20	--
10	--

4 — Discussion

This report has shown that, despite the differences in their designs and their operationalization of the English for academic purposes construct, the CAEL CE and IELTS - Academic overall scores of the participants in this study are correlated to a high degree (r = .81). A CAEL CE overall band score of 60 corresponds to an IELTS - Academic overall band score of 6.5. A CAEL CE overall band score of 70 corresponds to an IELTS - Academic overall band score of 7.

There are, however, a number of issues to consider with respect to the interpretation of the results. The first is test design. Though the CAEL CE and IELTS - Academic tests have the same number of components measuring similar skills and underlying constructs, these constructs have been operationalized differently, resulting in markedly different test designs. For instance, IELTS - Academic does not integrate skills while CAEL CE adopts a strong interpretation of skills integration. Consequently, CAEL CE scores do not mean exactly the same thing as IELTS - Academic scores—even after a link has been established.

Related to these differences in test design, we can expect differences in how test takers with the same underlying proficiency will perform on each test. Indeed, if test takers have prepared for one test but not the other, we expect that this will impact their performance. In the case of this study, we expect that test takers had primarily prepared for IELTS - Academic. As a result, they probably performed relatively better on IELTS - Academic than they did on the less well-known CAEL CE.

Other considerations concern the sample. As is common in studies of this nature, we relied on a self-selected sample (i.e., participants volunteered to report their IELTS - Academic score to Paragon). Notably, the numbers of participants with relatively low and very high scores were small. This limits generalizations of the results, particularly with respect to test takers that differ significantly from the research sample. The inclusion of cases with lower and higher scores is recommended. Additionally, since the CAEL CE target population is students intending to study at post-secondary Canadian institutions, more participants aged 17–29 would be desired.

Finally, since the sample size for this study was modest (N = 88), we recommend caution in interpreting the results. Results become more stable and reliable as sample sizes increase. As in any score comparison research, the score comparison between the CAEL CE and IELTS - Academic tests will benefit from more research. A larger and better-controlled sample (in terms of data collection constraints, number of cases at each score point for each test, and participant background information) will improve our confidence in the score comparison results. To confirm the findings from this preliminary study, Paragon intends to continue collecting data from test takers who have attempted both CAEL CE and IELTS - Academic.

References

Albano, A. D. (2016). equate: An R Package for Observed-Score Linking and Equating. Journal of Statistical Software, 74(8), 1–36. doi:10.18637/jss.v074.i08

Dorans, N. J. (1999). *Correspondences between ACT and SAT I Scores*. New York, NY: College Entrance Examination Board.

Educational Testing Service (2010). *Linking TOEFL-iBT Scores to IELTS Scores: A Research Report*. Princeton, NJ: Educational Testing Service.

Holland, P. W., & Dorans, N. J. (2006). Linking and equating. In R. L. Brennan (Ed.), *Educational Measurement* (4th ed., pp. 187–220). Westport, CT: Praeger Publishers.