

# **The Occupational English Test and IELTS: A Benchmarking Report**

Gad S Lim

Cambridge Assessment English

May 2016, updated October 2017

## **Introduction**

The Occupational English Test (OET) is an international language test specifically designed to assess the language communication skills of healthcare professionals who seek to register and practise in an English-speaking environment. Distinct versions of the test are available for twelve healthcare professions. OET is recognised by regulatory healthcare bodies and councils in the UK, Australia, New Zealand, Singapore, Dubai and Namibia as well as by the Australian Department of Immigration and Border Protection and Immigration New Zealand for a number of visa categories.

Originally developed in the 1980s, OET is backed by over 30 years of research by the Language Testing Research Centre at the University of Melbourne, and has gone through a continuous cycle of research, validation and evaluation to ensure it remains relevant and fit for purpose. In 2013, ownership of the test passed to Cambridge Boxhill Language Assessment (CBLA).

CBLA is committed to continually enhancing the quality and accessibility of the test. In this report, with data provided by CBLA, we detail the findings of a study into the relationship between candidates' performance on OET and on IELTS Academic, another test used to evaluate the language abilities of prospective migrant health professionals, as well as steps taken subsequent to the study.

## **The Occupational English Test (OET) and IELTS**

Test users often need to compare scores on different tests for various reasons. However, making comparisons is not a straightforward enterprise; according to the Standards for Educational and Psychological Testing (AERA, APA, NCME, 1999), scores can be considered 'equivalent' only when the tests' features are identical or closely similar to each other. In this case, the comparison is between a specific purpose test of English for health contexts and a more general test of English for academic purposes. It should therefore be expected that performance on the two tests will differ to some extent.

### Data and Method

CBLA invited OET candidates who took the test in 2013 and who had also taken IELTS to submit their scores on the two exams. In total, 359 candidates representing 25 nationalities responded to the invitation. The number of candidates reported upon in this analysis was reduced somewhat because some respondents provided incomplete information, and because the data needed to be counterbalanced to account for expected bias in sampling (Lim, Geranpayeh, Khalifa & Buckendahl, 2013). Candidates who did better on OET than on IELTS are more likely to respond to a call from OET which, left uncorrected, would provide a skewed picture of the relationship between the two tests. Such candidates (who took IELTS first and then OET second<sup>1</sup>, because they did not get the desired result on the first test) indeed outnumbered those who took OET then IELTS in this data. The final sample was therefore adjusted to include an equal number of candidates who had taken the tests in each order.

---

<sup>1</sup> Candidates provided the date they took each test.

Given the nature of the data, scores on the two exams were linked using the equipercentile method (Kolen & Brennan, 2004). This method was also employed by previous studies with similar data and purpose (e.g. ETS, 2010; Lim et al., 2013; Pommerich, Hanson, Harris & Scoring, 2000).

## Results

OET reports grades from A to E whereas IELTS reports band scores from 0 to 9 in half band increments. Table 1 shows means, standard deviations and ranges for each subtest of each exam, with OET grades converted to a scale of 1 to 5 (E=1...A=5). Because OET reports fewer possible grades, score variation is expectedly smaller.

Table 1. Means, Standard Deviations, and Ranges

	Mean	SD	Range	Population Mean
OET Listening	3.8	0.6	2 - 5	3.7
OET Reading	3.9	0.6	2 - 5	3.7
OET Speaking	3.8	0.6	1 - 5	3.6
OET Writing	3.8	0.7	2 - 5	3.8
IELTS Listening	7.2	0.9	5.0 - 9.0	-
IELTS Reading	6.9	0.9	3.5 - 9.0	-
IELTS Speaking	7.0	0.8	3.0 - 9.0	-
IELTS Writing	6.4	0.7	3.0 - 9.0	-
IELTS Overall	6.9	0.8	4.0 - 9.0	6.8

In general, this sample of candidates is just slightly stronger than the overall population. OET candidates in this sample were on average a tenth of a grade stronger than the total population of test takers in 2013. The IELTS website reports an average Overall score of 6.8 for candidates who took the Academic version of the test in 2012 for the purpose of registering as doctors (IELTS, 2014), the latest year for which information is available, whereas the average for this sample was 6.9. Few candidates in this sample obtained OET grades of D and E, mirroring the wider population. The sample can therefore be taken as fairly representative of the population.

Table 2 shows the correlations (all significant at  $p < 0.01$ ) between scores on OET and IELTS. Correlations were on the whole moderate but somewhat weak in the case of Writing. This relationship between the two exams is not unexpected given that they are measuring fairly different constructs.

Table 2. Correlations Between OET and IELTS Scores

	Correlation
Listening	0.48
Reading	0.52
Speaking	0.48
Writing	0.36

Having considered the nature of the data, the results of the equipercentile linking for each of the four subtests are presented in Figure 1. Only OET grades A to C are shown, as there was insufficient data to make a clear determination of what the lower-bound IELTS band score is for grades D and E. As neither test was designed to measure test takers below the equivalent of IELTS band 4, that represents a safe lower bound for concordance purposes.

Because OET reports fewer grades, each OET grade level naturally covers several IELTS (half) bands. CBLA recognises that a finer-grain score reporting system would better serve

users who want to further distinguish candidates of different ability within each grade, and has included this among the improvements in modernisation plans.

Figure 1. Indicative Relationship Between OET and IELTS Scores, 2013

Listening		Reading		Speaking		Writing	
OET	IELTS	OET	IELTS	OET	IELTS	OET	IELTS
C	5.5	C	5.5	C	5.5	C	5.5
	6.0		6.0		6.0		6.0
	6.5		6.5		6.5		6.5
B	7.0	B	6.5	B	7.0	B	6.5
	7.5		7.0		7.5		7.0
	8.0		7.5		8.0		7.5
A	8.5	A	8.0	A	8.5	A	7.5
	9.0		8.5		9.0		8.0
			9.0		9.0		8.5
					9.0		9.0

It can also be seen that the range covered by each grade/band level is not the same across subtests. For example, an OET B covers IELTS bands 7.0 – 8.0 in Listening, whereas the same grade covers 6.5 – 7.0 in Writing. This reflects, on the one hand, the fact that different exams divide up the ability range in different ways. On the other hand, this could also be due to the fact that IELTS is not naturally 'flat', as the average score candidates get on Academic Writing is approximately half a band lower than the score they get for Listening (IELTS, 2014).

## **Discussion**

Many test users accept a grade of B on OET at the same time that they accept a band score of 7.0 on IELTS. As this data shows, the two are not entirely comparable, for entirely expected reasons.

As previously noted, IELTS is a test of English for academic contexts, whereas OET is a test of English for healthcare contexts. The cut scores on OET were originally arrived at by consulting a group of healthcare professionals, who had experience supervising international healthcare professionals, on what level of English is necessary for them to perform their job safely (Lumley, Lynch & McNamara, 1994; McNamara, 1996). While the challenges of oral communication across the two contexts are comparable, the requirements for textual communication are quite different: reading and writing of extended texts features less in healthcare settings (Macqueen, et al, 2012; Vidakovic & Khalifa, 2013). With that in mind, the cut scores for Reading and Writing being set closer to IELTS 6.5 are appropriate.

Even so, there is a fairness issue involved when OET Bs and IELTS 7s are being accepted as comparable outcomes. For that reason, taking all things into account, CBLA has since the conclusion of the study reported here adjusted the cut scores for OET Reading and Writing grade B, so that they are closer in line with accepted IELTS equivalents. A more recent standard setting exercise (Knoch, Elder, Flynn, Manias, McNamara, Zhang, & Huisman, 2017) provides further confirmatory evidence in support of the adjustments made.

Something that has become clear to CBLA as a result of the study reported here is that the current OET score reporting system (A-E) could use greater granularity, so that candidates within each grade can be further distinguished. A grade of C, for instance, spans the ability range covered by IELTS bands 5.5 to 6.5, and the need has indeed arisen among some users of the test to identify people who are at IELTS 6.5. With that in mind, based on the information this study provides, a grade of C+ has been introduced that is comparable to that IELTS band score.

Thus, following this study, the best approximate overall IELTS score ranges covered by each OET grade at present are as follows:

OET	IELTS
A	8.0 – 9.0
B	7.0 – 7.5
C+	6.5
C	5.5 – 6.0
D	4.0 – 5.0*

\*see Results section

## 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.
- Angoff, W.H. (1971). Scales, norms, and equivalent scores. In R.L. Thorndike (Ed.), *Educational Measurement* (2<sup>nd</sup> ed., pp. 508-560). Washington, DC: American Council on Education.

- Buckendahl, C.W., Russell, S., & Lim, G.S. (2014). *Aligning the Occupational English Test to the Common European Framework of Reference for Languages: A standard setting study*. Melbourne, Australia: Cambridge Boxhill Language Assessment.
- Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge, UK: Cambridge University Press.
- Educational Testing Service. (2010). *Linking TOEFL iBT scores to IELTS scores: A research report*. Princeton, NJ: Author.
- Hammond, D., & Buckendahl, C.W. (2006). Do portfolio assessments have a place in dental licensure? Against the proposition. *Journal of the American Dental Association*, 137(1), 30-34.
- IELTS. (2014). Percentile ranks 2014. Available at:  
[http://ielts.org/researchers/analysis\\_of\\_test\\_data/percentile\\_ranks\\_2012.aspx](http://ielts.org/researchers/analysis_of_test_data/percentile_ranks_2012.aspx).
- Impara, J.C. & Plake, B.S. (1997). An alternative approach to standard setting. *Journal of Educational Measurement*, 34(4), 355-368.
- Knoch, U., Elder, C., Flynn, E., Manias, E., McNamara, T., Zhang, Y., & Huisman, A. (2017). Towards improved quality of written patient records: Development and validation of language proficiency standards for writing for non-native English speaking health professionals. Melbourne: Language Testing Research Centre, University of Melbourne.
- Kolen, M.J., & Brennan, R.L. (2004). *Test equating, scaling and linking: Methods and practices*. New York: Springer.
- Lim, G.S., Geranpayeh, A., Khalifa, H., & Buckendahl, C.W. (2013). Standard setting to an international reference framework: Implications for theory and practice. *International Journal of Testing*, 13(1), 32-49.
- Lumley, T., Lynch, B., & McNamara, T. (1994). A new approach to standard-setting in language assessment. *Melbourne Papers in Language Testing*, 3(2), 19-40.
- Macqueen, S., Yahalom, S., Kin, H., & Knoch, U., Writing demands of healthcare professionals. Melbourne: Language Testing Research Centre.
- McNamara, T. (1996). *Measuring Second Language Performance*. London: Longman.

- O'Neill, T.R., Buckendahl, C.W., Plake, B.S., & Taylor, L. (2007). Recommending a nurse-specific passing standard for the IELTS examination. *Language Assessment Quarterly*, 4(4), 295-317.
- Plake, B.S. (1998). Setting performance standards for professional licensure and certification. *Applied Measurement in Education*, 11(1), 65-80.
- 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 Report 2000-1). Iowa City, IA: ACT.
- Vidakovic, I., & Khalifa, H. (2013). Stakeholder perceptions of Occupational English Test: An exploratory study. *Research Notes*, 54, 29-32.