ETS TOEFL

Quality Beyond Measure.

MyBest™ Scores: A Rationale for Using *TOEFL iBT*® Superscores

A Research Report

ETS
March 2019

Starting in August 2019, *TOEFL iBT*® test score reports will include a new set of scores called *MyBest*™ scores, in addition to the scores from a single test date. *MyBest* scores — sometimes generically referred to as superscores — are the highest scores on each section (Reading, Listening, Speaking and Writing) from all of the test taker's valid TOEFL iBT test scores from tests taken in the last two years, and a total score that is the sum of the *MyBest* section scores.

The use *of MyBest* scores reflects an increasingly popular approach in university admissions of combining the highest section scores from multiple test dates. This practice provides opportunities for test takers to leverage their best performance on each section, demonstrate growth in areas of relative weakness and compensate for suboptimal performance caused by factors unrelated to the test.

MyBest scores are valid indicators of performance. They are built from valid section scores and their validity is supported by research.

Here is an example of a test taker who took the TOEFL iBT test three times in two years and earned the following scores (*MyBest* scores in bold):

Section	Test Date One	Test Date Two	Test Date Three	*MyBest*™ Scores
Reading	**24**	23	23	**24**
Listening	19	20	**22**	**22**
Speaking	20	20	**21**	**21**
Writing	21	**23**	22	**23**
Total Score	84	86	88	**90**

The test taker's highest Reading section score was 24, from Test Date One; the highest Listening section score was 22, from Test Date Three; the highest Speaking section score was 21, from Test Date Three; and the highest Writing section score was 23, from Test Date Two. These are the test taker's *MyBest* section scores. The resulting *MyBest* total is the sum of the *MyBest* section scores, which is 90. As can be seen, the *MyBest* total score is higher than the total score for any single test date.

Rationale

There are several factors that support the validity of *MyBest* scores:

- *MyBest* scores are valid indicators of performance as they are drawn from all valid test scores within a two-year period.
- Studies have shown that superscores have similar score validity to other commonly used scoring methods, such as use of the most recent test scores or use of the average scores across all testing occasions. Results of these research studies indicated that observed increases between superscores and scores from a single administration reflect the positive effects of retesting (Mattern, Radunzel, Bertling, & Ho, 2018; Roszkowski, & Spreat, 2016).

- ETS researchers conducted a study that compared TOEFL iBT test takers' *MyBest* scores and their most recent scores, using all test takers who repeated the test within a recent two-year window. Study results indicated that:
  - The average score difference between the *MyBest* total score and most recent total test score was approximately four points. Score differences were generally lower when there was a shorter duration between tests.
  - On average, the difference between the *MyBest* scores and the most recent section scores were greater for Reading and Listening than for Speaking and Writing.
  - Those with greater differences had lower proficiency levels overall and had, on average, retested a greater number of times.

Thus, based on study results, most repeat test takers can anticipate small to moderate differences between *MyBest* scores and their most recent test scores. These small differences may be important if the test taker is retesting to meet a specific score requirement.

In practice, use of *MyBest* scores in admissions considerations may increase the number of eligible candidates. For example, consider a university that requires a minimum score of 18 for each section and 80 for the total score.

Below are three examples — drawn from the ETS study — of candidates who would be borderline ineligible for admission based on one or two of the section scores or the total score, but would be eligible for admission based on *MyBest* scores. Italics are used to indicate scores that did not meet criteria. *MyBest* scores are bolded.

Candidate One met criteria for three sections on Test Dates One and Two. At Test Date Three all four section scores met criteria, but the total score did not meet criteria. Using the *MyBest* scores, the section and total score criteria were met.

Section	Test Date One	Test Date Two	Test Date Three	*MyBest* Scores
Reading	19	18	**24**	**24**
Listening	**20**	19	19	**20**
Speaking	*15*	*15*	**18**	**18**
Writing	18	**19**	18	**19**
Total Score	72	71	79	**81**

Candidate Two met criteria for three of the sections at Test Date One, and two sections on Test Date Two. Using the *MyBest* scores, the section and total score criteria are met.

Section	Test Date One	Test Date Two	*MyBest* Scores
Reading	18	**22**	**22**
Listening	*17*	**20**	**20**
Speaking	**19**	*17*	**19**
Writing	**21**	*17*	**21**
Total Score	*75*	*76*	**82**

Candidate Three met criteria for two sections at Test Date One, and all section criteria on Test Date Two, but the total score was too low. Using the *MyBest* scores, the section and total score criteria are met.

Section	Test Date One	Test Date Two	*MyBest* Scores
Reading	**19**	18	**19**
Listening	**22**	19	**22**
Speaking	*17*	**20**	**20**
Writing	*15*	**21**	**21**
Total Score	*73*	*78*	**82**

Thus, as can be seen with the examples, use of *MyBest* scores provides greater flexibility for admissions decisions through the option for inclusion of those candidates with *MyBest* scores that meet the admissions requirements who would otherwise be disqualified based on results of one or more test occasions.

After *MyBest* scores are launched, ETS will continue to monitor their validity to ensure that *MyBest* scores remain a valid way to meet score requirements.

References

Mattern, K., Radunzel, J., Bertling, M., & Ho, A. D. (2018). How should colleges treat multiple admissions test scores? *Educational Measurement: Issues and Practice, 37*(3), 11–23

Roszkowski, M., & Spreat, S. (2016). Retaking the SAT may boost scores but this doesn't hurt validity. *Journal of the National College Testing Association, 2*(1), 1–16.