

TOEFL[®] *Research* **INSIGHT**

Guidelines for Setting Useful Score Requirements for the *TOEFL iBT*[®] Test

VOLUME 9

TOEFL® Research Insight Series, Volume 9:

Guidelines for setting useful score requirements for the TOEFL iBT® test

Preface

The TOEFL iBT® test is the world's most widely respected English language assessment and used for admissions purposes in more than 150 countries, including Australia, Canada, New Zealand, the United Kingdom, and the United States (see test review in Alderson, 2009). Since its initial launch in 1964, the TOEFL® test has undergone several major revisions motivated by advances in theories of language ability and changes in English teaching practices. The most recent revision, the TOEFL iBT test, was launched in 2005. It contains a number of innovative design features, including integrated tasks that engage multiple skills to simulate language use in academic settings and test materials that reflect the reading, listening, speaking, and writing demands of real-world academic environments.

In addition to the TOEFL iBT test, the TOEFL® Family of Assessments was expanded to provide high-quality, English proficiency assessments for a variety of academic uses and contexts. The TOEFL® Young Students Series features the TOEFL Primary® and TOEFL Junior® tests, which are designed to help teachers and learners of English in school settings. In addition, the TOEFL ITP® program offers colleges, universities, and others affordable tests for placement and progress monitoring within English programs as a pathway to eventual degree programs.

At ETS, we understand that scores from the TOEFL Family of Assessments are used to help make important decisions about students, and we would like to keep score users and test takers up-to-date about the research results that help assure the quality of these scores. Through the publication of the TOEFL® Research Insight Series, we wish to communicate to the institutions and English teachers who use the TOEFL tests the strong research and development base that underlies the TOEFL Family of Assessments and demonstrate our continued commitment to research.

Since the 1970's, the TOEFL test has had a rigorous, productive, and far-ranging research program. But why should test score users care about the research base for a test? In short, it is only through a rigorous program of research that a testing company can substantiate claims about what test takers know or can do based on their test scores, as well as provide support for the intended uses of assessments and minimize potential negative consequences of score use. Beyond demonstrating this critical evidence of test quality, research is also important for enabling innovations in test design and addressing the needs of test takers and test score users. This is why ETS has established a strong research base as a fundamental feature underlying the evolution of the TOEFL Family of Assessments.

This portfolio is designed, produced, and supported by a world-class team of test developers, educational measurement specialists, statisticians, and researchers in applied linguistics and language testing. Our test developers have advanced degrees in fields such as English, language education, and applied linguistics. They also possess extensive international experience, having taught English on continents around the globe. Our research, measurement, and statistics teams include some of the world's most distinguished scientists and internationally recognized leaders in diverse areas such as test validity, language learning and assessment, and educational measurement.

To date, more than 300 peer-reviewed TOEFL Family of Assessments research reports, technical reports, and monographs have been published by ETS, and many more studies on the TOEFL tests have appeared in academic journals and book volumes. In addition, over 20 TOEFL test-related research projects are conducted by ETS's Research & Development staff each year and the TOEFL Committee of Examiners — comprising language learning and testing experts from the global academic community — funds an annual program of TOEFL Family of Assessments research by independent external researchers from all over the world.

The purpose of the *TOEFL Research Insight Series* is to provide a comprehensive, yet user-friendly account of the essential concepts, procedures, and research results that assure the quality of scores for all products in the TOEFL Family of Assessments. Topics covered in these volumes feature issues of core interest to test users, including how tests were designed; evidence for the reliability, validity, and fairness of test scores; and research-based recommendations for best practices.

The close collaboration with TOEFL test score users, English language learning and teaching experts, and university scholars in the design of all TOEFL tests has been a cornerstone to their success and worldwide acceptance. Therefore, through this publication, we hope to foster an ever-stronger connection with our test users by sharing the rigorous measurement and research base, as well as solid test development, that continues to help ensure the quality of the TOEFL Family of Assessments.

John Norris, Ph.D.

Senior Research Director
English Language Learning and Assessment
Research & Development Division
ETS

The following individuals contributed to this volume (in alphabetical order): Sandy Bhangal, Marian Crandall, John Norris, Spiros Papageorgiou (lead author), Jonathan Schmidgall, and Richard J. Tannenbaum.

Guidelines for setting useful score requirements on the TOEFL iBT test

Test scores are used to facilitate various decisions, such as admission into a degree program, placement into classes, or certification and licensure. Depending on the context in which a test is used, score-based decisions can have significant impact on individual students, educational institutions, and society. One of the main purposes of the TOEFL iBT test is to measure the ability of international students to use English in an academic environment. Therefore, TOEFL iBT test scores are primarily used to facilitate decisions about student admission into higher education programs and courses where instruction takes place in English, placement into English language classes, and decisions about the language proficiency level of international graduate students who undertake responsibilities as teaching assistants. To make such decisions, a minimum score — typically called the “cut score” — needs to be defined. A cut score on the TOEFL iBT test is essentially the score that a student needs to achieve to meet requirements for admission and placement.

A standard setting study is typically organized to set cut scores; however, such a study might not be practical for many score users, as it requires a considerable amount of resources (e.g., recruiting experienced facilitators and panelists who meet for several days, preparing materials for the panelists to review, collecting data related to test scores). Therefore, this volume in the *TOEFL Research Insight Series* aims to help score users, such as admissions officers and English language program directors, set reasonable and useful cut scores using available TOEFL iBT test resources. This volume builds upon the discussion of the interpretation of TOEFL iBT test scores and their use in making decisions about students’ English language proficiency in Volume 5: *Information for TOEFL iBT® Score Users, Teachers, and Learners*. However, the focus of this volume is on topics related to score requirements, as summarized in Figure 1.

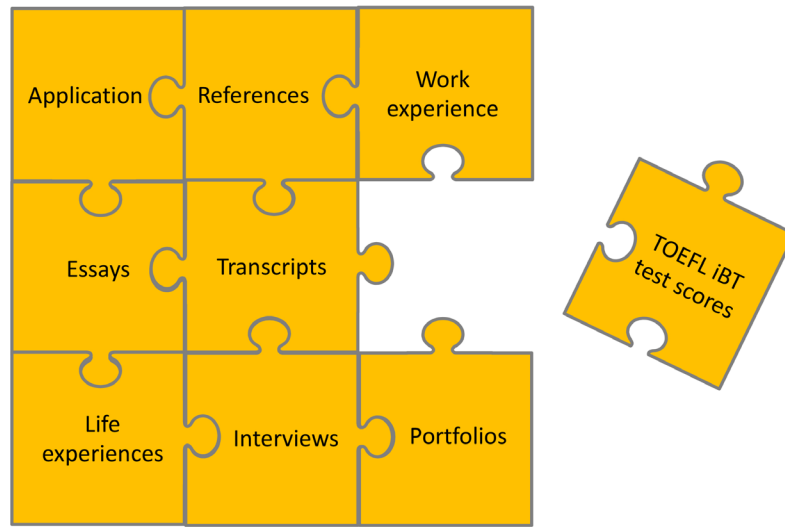
Figure 1 Content of this volume

- The role of language proficiency in academic success
- Consequences of false classifications resulting from cut scores
- Available resources to facilitate setting cut scores on the TOEFL iBT test
- Critical steps in setting TOEFL iBT test score requirements
- Baseline recommendation for TOEFL iBT test cut scores for college admissions

Language proficiency and academic success

When it comes to evaluating students’ ability to use English in an academic context, language proficiency becomes part of a holistic admission policy for international students. A holistic approach will evaluate multiple criteria in addition to TOEFL iBT test scores as shown in Figure 2.

Figure 2 Components of a holistic admission policy for international students



While a high score on the TOEFL iBT test indicates high ability in using English in academic contexts, it cannot guarantee on its own that a student will be successful academically. Bridgeman, Cho, and DiPietro (2016) explain why: “English language skills are a necessary but not sufficient condition for success in academic study for international students at a university in which English is the only or dominant language of instruction” (p. 308). They point out that other factors, beyond language proficiency, can affect success. These factors include quantitative skills, content knowledge, and various noncognitive attributes such as motivation and persistence. A language test is intended to measure language proficiency, not abilities beyond language proficiency. Therefore, no matter how carefully cut scores are set, some students whose language skills were deemed sufficient for studying in English might still fail academically for reasons unrelated to their language proficiency.

Carefully applied minimum score requirements on the TOEFL iBT test can help admissions officers feel confident in decisions about international students’ applications. At the same time, additional insights into a student’s English language proficiency can complement TOEFL iBT test scores as part of a holistic application policy (see also *Volume 5: Information for TOEFL iBT® Score Users, Teachers, and Learners*). For example, if a student fails to meet TOEFL iBT test score requirements by a few score points, other sources of information — such as those shown in Figure 2 — can help admissions staff feel more confident in their decision to reject the candidate or make an exception and admit them.

This volume emphasizes the need to apply score requirements in a principled manner, as discussed in subsequent sections. However, no matter how carefully cut scores have been decided, their usefulness for decision making depends on two important principles: relevance of test design for a given purpose and empirical evidence supporting the validity of the test scores. The TOEFL iBT test addresses these principles in the following ways:

- **Principle 1: The test design reflects the demands of real-life academic tasks.** The design of the TOEFL iBT test is based on years of research and it comprehensively evaluates the language skills and

abilities that English language learners need to succeed in academic environments where English is the medium of instruction. To do so, the test includes language tasks that reflect those that students need to perform in class (see Volume 1: *TOEFL iBT® Test Framework and Test Development*).

- **Principle 2: The usefulness of the test scores for making decisions is supported empirically by ongoing research.** As explained in the preface and other volumes in the *TOEFL Research Insight Series* (Volume 2: *TOEFL® Research*; Volume 4: *Validity Evidence Supporting the Interpretation and Use of TOEFL iBT® Scores*), the TOEFL iBT test is supported by a unique, comprehensive research program — with hundreds of peer-reviewed publications authored by ETS staff and non-ETS researchers. TOEFL test-related research provides compelling evidence of the validity of the test scores and the usefulness of these scores for making important decisions about students' English language proficiency.

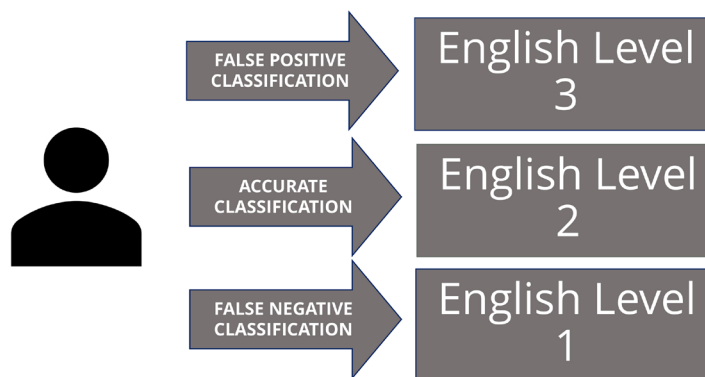
The above principles have important implications for score users who need to develop useful and relevant score requirements related to the academic language proficiency of international students:

- If a test does not evaluate the relevant language skills and abilities, then there is little value in investing in the process of setting cut scores because classification of students into “meeting” and “not meeting” the language requirements will be meaningless.
- Cut scores on a language proficiency test are likely to be useful if there is strong empirical evidence of the usefulness of the test scores for decision making; conversely, the lack of such empirical evidence threatens the usefulness of the cut scores.

Classification decisions and their consequences

When setting cut scores, test takers are classified into two or more categories. Figure 3 illustrates the case of a test taker who needs to be placed into the appropriate English language support class (from Level 1 — the lowest class to Level 3 — the most advanced class).

Figure 3 Type of classification decisions when setting score requirements



Assume in the example illustrated in Figure 3 that the Level 2 class is the accurate placement for this student. If a language test is used to facilitate the placement decision and cut scores for these classes are reasonable — that is, not too high nor too low — the student will be accurately classified as a “Level 2” student. But if cut scores are not reasonable, then two types of false classification are possible. A false positive classification would place the student into the Level 3 class. In this case, the student is assumed to have sufficient language proficiency for this class when, in fact, this is not the case (the student’s language proficiency is suitable for the lower, Level 2 class). A false negative classification would place the student into the Level 1 class. In this case, the student is assumed to lack language proficiency for the Level 2 class when, in fact, the student’s language proficiency is adequate for the Level 2 class.

The expectation is that decisions about the classification of students based on cut scores will be accurate. However, in practice, false classifications for some students are expected. While it is not possible to fully eliminate false classifications, the likelihood of one type of false classification can be reduced at the cost of increasing the likelihood for the other type of false classification. Score users need to decide which type of false classification is more important to avoid when setting cut scores, after considering the consequences of false classifications in their own context. Figure 4 shows the possible consequences of stringent score requirements (a high cut score) on the TOEFL iBT test in the context of university admissions, while Figure 5 shows the possible consequences of lenient score requirements (a low cut score).

Figure 4 Possible consequences of stringent TOEFL iBT test score requirements

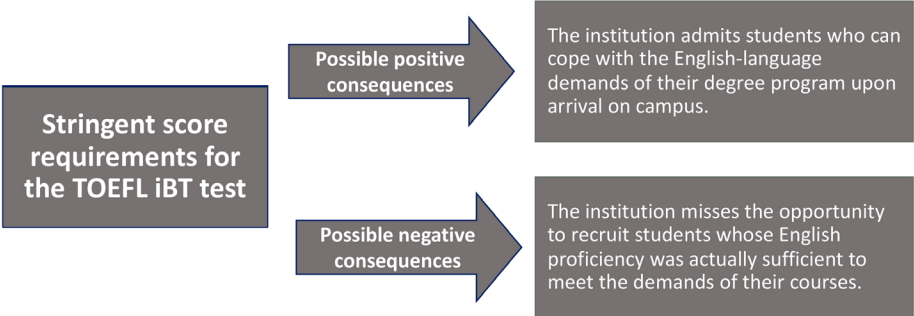
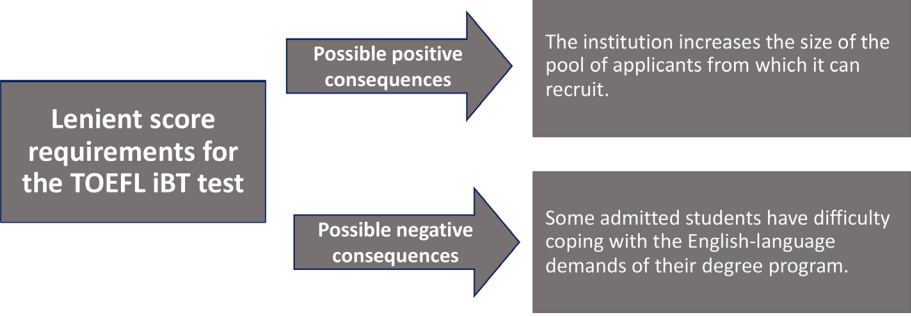


Figure 5 Possible consequences of lenient TOEFL iBT test score requirements



As shown in Figure 4, an institution might decide to apply stringent score requirements by setting high cut scores on the TOEFL iBT test, thus reducing the likelihood of false positive classifications. Consequently, the institution can have high confidence in recruiting international students with the ability to use English in an academic environment when they arrive on campus. However, high cut scores raise the likelihood for false negative classifications, as some students might be denied admission when they can actually cope with the English language demands of their degree program. In this case, the institution misses the opportunity to recruit qualified students. In the opposite case, as shown in Figure 5, an institution might decide to set lower cut scores on the TOEFL iBT test, thus reducing the likelihood of false negative classifications. In this case, the institution will be able to recruit from a larger pool of international students than the institution in the previous example. However, setting lower cut scores also raises the likelihood for false positive classifications, as some students might be admitted who subsequently have difficulty coping with the English language demands in their degree programs.

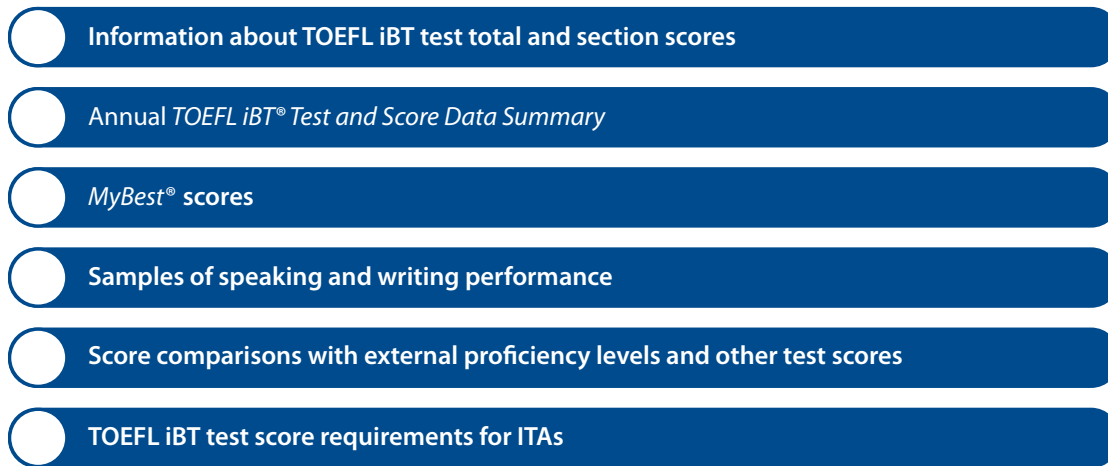
The previous examples are set in the context of university admissions; however, false classifications might have negative consequences in other contexts where English language test scores are used to facilitate decisions about students. For example, when placing students into classes at different language levels, misplacing them in inappropriately difficult courses might lead them to feel frustrated and unmotivated and they might decide to drop out. Misplacing students in courses that are too easy might make them feel bored. Ultimately, learning is less likely to happen in either situation. When using TOEFL iBT test scores to screen graduate students' language proficiency before they undertake the role of an international teaching assistant (ITA), inappropriately low TOEFL iBT test cut scores might mean that some ITAs are not understood by their students when they teach in English. Conversely, inappropriately high TOEFL iBT test cut scores might mean that some qualified graduate students cannot undertake the role of ITA, thus missing the opportunity to gain teaching experience and financially support their own studies.

Irrespective of the context, when setting TOEFL iBT test cut scores to facilitate decisions about students' English language proficiency, the starting point should be to consider the consequences of applying stringent or lenient score requirements as well as identifying the type of classification errors — false positive or false negative — that should be minimized. We return to this issue later, when we discuss critical steps in finalizing score requirements and evaluating their effectiveness.

Resources to facilitate the setting of cut scores for the TOEFL iBT test

The *TOEFL*® program provides resources that can inform decisions about test-taker performance and the setting of cut scores. These resources, shown in Figure 6, are discussed in this section.

Figure 6 Resources to facilitate the setting of cut scores for the TOEFL iBT test



Information about TOEFL iBT test total and section scores

TOEFL iBT test scores are reported on a score scale of 0–30 for each of the four test sections — reading, listening, speaking, and writing. To facilitate score interpretation, the section scores are grouped into levels and performance descriptors illustrate the meaning of these levels (see Appendix). The total score is reported on a scale of 0–120, which is the sum of the four section scores.

It is good practice to set cut scores on some, if not all, of the separate section scores as well as the total score, so that admission decisions are based on a nuanced understanding of how the language profile of a student aligns with the language profile needed. Considering the student’s complete language profile is important because two students might receive the same total score, but their abilities across language skills might vary. Research also shows that decisions about English language proficiency can be better informed when considering both TOEFL iBT test total and section scores rather than the total score in isolation (Bridgeman et al., 2016; Ginther, & Yan, 2016).

Annual TOEFL iBT Test and Score Data Summary

The annual *TOEFL iBT® Test and Score Data Summary* is a report that provides useful statistical information about the performance of TOEFL iBT test takers during the previous calendar year. The most recent version can be found at www.ets.org/toefl/score-users/resources-services/. Score users can use information from the annual report to evaluate the reasonableness of the cut scores they have set, specifically:

- Mean (arithmetic average) for total and section scores for the overall test taking population as well as various subgroups, such as gender, reason for taking the test, native language, and native country. The mean total and section scores of these groups can offer an indication of how strict or lenient score requirements might be.

- Percentile ranks for the total and section scores, which show the percentage of test takers at or below a score, for the overall test taking population and various subgroups. Table 1 lists the percentile ranks for three TOEFL iBT test total scores in 2019. The table also illustrates the use of percentile ranks to evaluate the strictness or leniency of three hypothetical score requirements for the total TOEFL iBT test score (cut score or 80, 92, or 100). The last column shows how many students who took the test that year would be eligible to apply if each of the three cut scores were used.

Table 1 Example of percentile ranks and their interpretation

TOEFL iBT test total score	Percentile rank for all test takers in 2019	A cut score at this level means that
100	78	22% of test takers in 2019 would have been eligible to apply
92	61	39% of test takers in 2019 would have been eligible to apply
80	38	62% of test takers in 2019 would have been eligible to apply

MyBest® scores

MyBest scores, sometimes called “superscores,” were introduced in 2019 as one of several improvements to the experience of TOEFL iBT test takers. In addition to the student’s total and section scores from a given test administration, TOEFL iBT test score reports also include the highest total and section scores from all of the test administrations within the past two years. (For the rationale behind the two-year expiration policy of test scores, see Powers, & Lall, 2013.)

A growing body of educational research suggests that superscores are helpful for making university admission decisions (see ETS, 2019). Institutions wishing to increase the pool of test takers who meet their score requirements might want to consult *MyBest* scores to establish the highest scores an applicant has achieved across multiple test administrations.

Samples of speaking and writing performance

Score users can access audio recordings of test taker responses to the speaking tasks and the responses to the writing tasks. The speaking and writing responses might be helpful to score users when reviewing applications of students who failed to meet TOEFL iBT test score requirements by a few score points and might help increase confidence in deciding whether to admit the student.

Score comparisons with external proficiency levels and other test scores

Two types of information on score comparisons can facilitate the setting of cut scores and are available on the TOEFL iBT test website <https://www.ets.org/toefl/score-users/scores-admissions/compare>

- **TOEFL iBT test score mapping onto the language proficiency levels of the Common European Framework of Reference (CEFR).** TOEFL iBT test total and section scores have been mapped on the CEFR levels, based on analysis of informed expert judgments, test scores, and academic performance data. Score mapping to the CEFR levels might help institutions already familiar with the framework set relevant cut scores. However, direct comparisons between TOEFL iBT test scores and the scores of other tests based on CEFR mapping should be avoided. As widely reported in language assessment literature, there is no independent authority certifying how various test providers map the scores of their tests to the CEFR levels (Papageorgiou, Tannenbaum, Bridgeman, & Cho, 2015). In addition, interpretations of the CEFR levels across test providers can vary considerably (Green, 2018).
- **Comparison with IELTS® test scores.** To help ensure that applicants are treated fairly, irrespective of the test they took, score users can consult the comparison tables and the interactive widget from the score concordance study between TOEFL iBT test scores and IELTS band levels. The score concordance study examined scores from 1,153 test takers who took both tests. Institutions should note that while the two tests assess similar constructs and may be used for similar purposes—and this evidence justifies the concordance study — they are not identical and cannot be considered interchangeable. For example, the two tests use different score reporting mechanisms. With a range from 0 to 30 for the section scores and 0 to 120 for the total score, the TOEFL iBT test score scales are more refined than the IELTS 9-band reporting scale. Therefore, multiple TOEFL iBT test scores will be equivalent to the same IELTS band. That is, there is no one-to-one mapping of scores from one test to the other. Because of notable differences between the TOEFL iBT test and other tests in terms of design, score reporting, and psychometric quality — and because of the complexity of designing a score concordance study — ETS does not endorse other test providers' score concordance tables comparing TOEFL iBT test scores to the scores of these providers' tests.

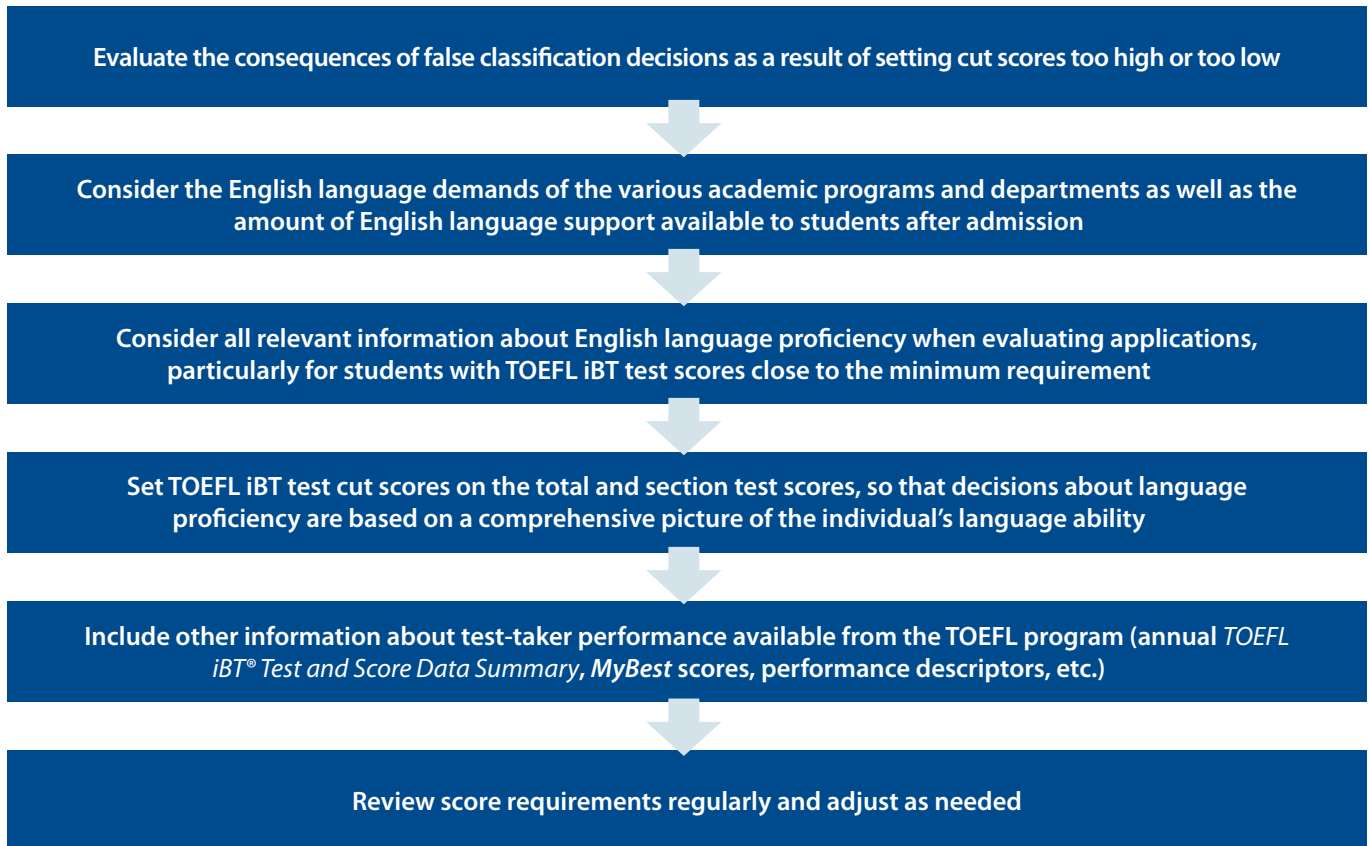
TOEFL iBT score requirements for ITAs

English language proficiency is one among several important factors to consider when determining the ability of international graduate students to accomplish various teaching tasks associated with their role as ITAs. A growing body of research supports the use of TOEFL iBT test scores, in particular the speaking section score, to evaluate the language proficiency of prospective ITAs (see Volume 10: *Using TOEFL iBT® Test Scores for Selecting International Teaching Assistants*). Research findings point to a speaking score of at least 23 for conditional ITA appointment (e.g., an international graduate student is offered the opportunity to teach on the condition that other language-related requirements are met) and a speaking score of at least 26 for unconditional ITA appointment (e.g., an international graduate student is offered the opportunity to teach without the need to fulfill other language-related requirements). Institutions might find information about score requirements for ITAs relevant to their admission policy, including when cut scores need to be set for international students applying to graduate degree programs.

Critical steps in setting score requirements

This volume in the *TOEFL Research Insight Series* discussed available resources that can support the setting of reasonable and useful cut scores on the TOEFL iBT test. Regardless of the process followed to set TOEFL iBT test cut scores, score users should follow a series of critical steps, which are summarized in Figure 7:

Figure 7 Steps in setting score requirements for the TOEFL iBT test



Baseline recommendation for TOEFL iBT test cut scores for admissions purposes

Cut scores on the TOEFL iBT test are likely to differ across institutions because of various contextual characteristics, such as those discussed earlier in this volume (language demands of the instruction, amount of language support offered, etc.). Table 2 provides baseline recommendations for total score requirements for college admissions purposes based on the experience of the TOEFL program working with score users around the world. The cut score recommendations are only intended as a starting point; score users who wish to establish TOEFL iBT test score requirements that are relevant and useful in their context should consider all relevant information about TOEFL iBT test scores presented in this volume. For example, English language programs might want to consider the TOEFL iBT test performance descriptors mentioned earlier to facilitate decisions about placement of students into classes and set learning objectives.

We conclude this volume by reiterating the importance of considering the consequences of classification errors and acting to rectify them. If false positive classifications are of concern, score users should explore if any students are being admitted without the needed language skills and provide support to help them improve their language proficiency. If false negative classifications are of concern, score users should explore whether there is indeed evidence of qualifying students being discouraged to apply.

Table 2 Baseline recommendation for TOEFL iBT test total score requirements

Total score requirement	Language demands of instruction and language support
90–120	<ul style="list-style-type: none"> • Usually appropriate when the English language demands of courses are high. • Most students typically don't need language support classes upon arrival on campus, although some students might benefit from it in regard to academic writing for specific disciplines.
70–89	<ul style="list-style-type: none"> • Usually appropriate for most courses with moderate English language demands. For courses with high English language demands, students might need to attend language support classes, especially for specific skills where their TOEFL iBT test section scores indicated weakness. • Many students may benefit from language support classes upon arrival on campus in regard to academic writing for specific disciplines.
60–69	<ul style="list-style-type: none"> • Usually appropriate for courses with low English language demands or when a substantial amount of language support is available to students upon arrival on campus. • Students typically benefit from most types of language support classes (e.g., general English, academic English).
Below 60	<ul style="list-style-type: none"> • Students are unlikely to be able to cope with the English language demands of their courses. • A year or more of full-time, intensive English language classes is typically needed.

Note: Information in this table is provided for guidance only. Score users should consider factors relevant to their context and set requirements for the total TOEFL iBT test score, as well as the four section scores, accordingly.

References

- Alderson, J. C. Test review: Test of English as a Foreign Language™: Internet-based Test (TOEFL iBT®). *Language Testing*, 26(4), 621–631. doi:10.1177/0265532209346371
- Bridgeman, B., Cho, Y., & DiPietro, S. (2016). Predicting grades from an English language assessment: The importance of peeling the onion. *Language Testing*, 33(3), 307–318.
- ETS (2019). MyBest™ scores: A rationale for using TOEFL iBT® superscores. Retrieved from https://www.ets.org/s/toefl/pdf/mybest_su.pdf
- Ginther, A., & Yan, X. (2018). Interpreting the relationships between TOEFL iBT scores and GPA: Language proficiency, policy, and profiles. *Language Testing*, 35(2), 271–295.
- Green, A. (2018). Linking tests of English for academic purposes to the CEFR: The score user's perspective. *Language Assessment Quarterly*, 15(1), 59–74.
- Papageorgiou, S., Tannenbaum, R. J., Bridgeman, B., & Cho, Y. (2015). *The association between TOEFL iBT® test scores and the Common European Framework of Reference (CEFR) levels* (Research Memorandum No. RM-15-06). Princeton, NJ: Educational Testing Service.
- Powers, D. E., & Lall, V. (2013). *Supporting an expiration policy for English language proficiency test scores* (Research Memorandum No. RM 12-03). Princeton, NJ: Educational Testing Service.

Appendix

Performance Descriptors for the TOEFL iBT Test (available at www.ets.org/s/toefl/pdf/pd-toefl-ibt.pdf)

Level	Reading Section
<p>Advanced Score range 24–30 CEFR Level C1</p>	<p>Test takers who receive a Reading section score at the Advanced level typically understand academic passages in English at the introductory university level. These passages are dense with propositions and information and can include difficult vocabulary; lengthy, complex sentences and paragraphs; and abstract or nuanced ideas that may be presented in complex ways.</p> <p>Test takers who score at the Advanced level typically can</p> <ul style="list-style-type: none"> • Understand a range of academic and low-frequency vocabulary as well as less common meanings of words. • Understand explicit connections among pieces of information and make appropriate inferences, even when the passage is conceptually dense and the language is complex. • Recognize the expository organization of a passage and the purpose that specific information serves within the larger context, even when the purpose of the information is not marked, and the passage is conceptually dense. • Follow a paragraph-length argument involving speculation, qualifications, counter-evidence, and subtle rhetorical shifts. • Synthesize information in passages that contain complex language and are conceptually dense.
<p>High-Intermediate Score range 18–23 CEFR Level B2</p>	<p>Test takers who receive a Reading section score at the High-Intermediate level typically understand the main ideas and important details of academic passages in English at the introductory university level, but they may have an incomplete or incorrect understanding of parts of passages that are especially dense with propositions and information, or complex in their presentation of ideas and information.</p> <p>Test takers who score at the High-Intermediate level typically can</p> <ul style="list-style-type: none"> • Understand common academic vocabulary, but sometimes have difficulty with low-frequency words or less common meanings of words. • Understand explicit connections among pieces of information and make appropriate inferences, but may have difficulty in parts of a passage that contain low-frequency vocabulary or that are conceptually dense, rhetorically complex, or abstract. • Distinguish important ideas from less important ones. • Often recognize the expository organization of a passage and the purpose of specific information within a passage, even when such information is not explicitly marked. • Synthesize information in a passage, but may have difficulty doing so when the passage is conceptually dense, rhetorically complex, or abstract.
<p>Low-Intermediate Score range 4–17 CEFR Level B1</p>	<p>Test takers who receive a Reading section score at the Low-Intermediate level typically understand some main ideas and important information presented in academic passages in English, but their overall understanding is limited. They are able to understand connections across two or more sentences when the relationships are clear and simple, such as a claim followed by a supporting example. However, they have difficulty following denser or more complex parts of a passage.</p> <p>Test takers who score at the Low-Intermediate level typically can</p> <ul style="list-style-type: none"> • Understand texts with basic grammar, but have inconsistent understanding of texts with complex grammatical structures. • Understand high-frequency academic vocabulary, but often have difficulty with lower-frequency words. • Locate information in a passage by matching words or relying on high-frequency vocabulary, but their limited ability to recognize paraphrases results in incomplete understanding of the connections among ideas and information. • Identify an author’s purpose when that purpose is explicitly stated or easy to infer from the context. • Recognize major ideas in a passage when the information is clearly presented, memorable, or illustrated by examples but have difficulty doing so when the passage is more demanding.
<p>Below Low-Intermediate Score range 0–3</p>	<p>Test takers with a Reading section score below 4 have not yet demonstrated proficiency at the Low-Intermediate level.</p>

Level	Listening Section
<p>Advanced Score range 22–30 CEFR Level C1</p>	<p>Test takers who receive a Listening section score at the Advanced level typically understand conversations and lectures that take place in academic settings. The conversations and lectures may include difficult vocabulary, abstract or complex ideas, complex sentence structures, various uses of intonation, and a large amount of information, possibly organized in complex ways.</p> <p>Test takers who score at the Advanced level typically can</p> <ul style="list-style-type: none"> • Understand main ideas and explicitly stated important details, even if not reinforced. • Distinguish important ideas from less important points. • Keep track of conceptually complex, sometimes conflicting, information over extended portions of a lecture. • Understand how information or examples are being used (for example, to provide evidence for or against a claim, to make comparisons or draw contrasts, or to express an opinion or a value judgment) and how pieces of information are connected (for example, in a cause-effect relationship). • Understand different ways that speakers use language for purposes other than to give information (for example, to express an emotion, to emphasize a point, to convey agreement or disagreement, or to communicate an intention). • Synthesize information, even when it is not presented in sequence, and make appropriate inferences on the basis of that information.
<p>High-Intermediate Score range 17–21 CEFR Level B2</p>	<p>Test takers who receive a Listening section score at the High-Intermediate level typically understand the main ideas and important details of conversations and lectures that take place in academic settings. The conversations and lectures may include difficult vocabulary, abstract or complex ideas, complex sentence structures, various uses of intonation, and information that must be tracked across sequences of utterances.</p> <p>However, lectures and conversations that are dense with information may present difficulty if the information is not reinforced.</p> <p>Test takers who score at the High-Intermediate level typically can</p> <ul style="list-style-type: none"> • Understand main ideas and explicitly stated important details that are reinforced (by repetition, paraphrase, or indirect reference). • Distinguish main ideas from other information. • Keep track of information over an extended portion of an information-rich lecture or conversation, and recognize multiple, possibly conflicting, points of view. • Understand how information or examples are being used (for example, to provide support for a claim), and how pieces of information are connected (for example, in a narrative explanation, a compare-and-contrast relationship, or a cause-effect chain). • Understand, though perhaps not consistently, ways that speakers use language for purposes other than to give information (for example, to emphasize a point, express agreement or disagreement, express opinions, or convey intentions indirectly), especially when the purpose is supported by intonation. • Synthesize information from adjacent parts of a lecture or conversation, and make appropriate inferences on the basis of that information, but may have difficulty synthesizing information from separate parts of a lecture or conversation.
<p>Low-Intermediate Score range 9–16 CEFR Level B1</p>	<p>Test takers who receive a Listening section score at the Low-Intermediate level typically understand the main idea and some important details of conversations and lectures that take place in academic settings. These conversations and lectures can include basic academic language, abstract or complex ideas that are significantly reinforced, complex sentence structures, certain uses of intonation, and a large amount of information that is repeated or significantly reinforced.</p> <p>Test takers at the Low-Intermediate level typically can</p> <ul style="list-style-type: none"> • Understand main ideas, even in complex discussions, when the ideas are repeatedly referred to, extensively elaborated on, or illustrated with multiple examples. • Understand explicitly stated important details, but may have difficulty understanding details if they are not reinforced (such as through repetition or with an example) or marked as important, or if they are conveyed over several exchanges among different speakers. • Understand some ways that speakers use language to express an opinion or attitude (for example, agreement, disagreement, surprise), especially when the opinion or attitude is related to a central theme, clearly marked as important, or supported by intonation. • Understand connections between important ideas, particularly if the ideas are related to a central theme or are repeated, and can make appropriate inferences from information expressed in one or two sentences, especially when that information is reinforced.
<p>Below Low-Intermediate Score range 0–8</p>	<p>Test takers with a Listening section score below 9 have not yet demonstrated proficiency at the Low-Intermediate level.</p>

Level	Speaking Section
<p>Advanced Score range 25–30 CEFR Level C1</p>	<p>Test takers who receive a Speaking section score at the Advanced level are typically able to communicate fluently and effectively on a wide range of topics with little difficulty.</p> <p>Test takers who score at the Advanced level typically can</p> <ul style="list-style-type: none"> • Speak clearly and use intonation to support meaning so that speech is generally easy to understand and follow; any minor lapses do not obscure meaning. • Speak with relative ease on a range of general and academic topics, demonstrating control of an appropriate range of grammatical structures and vocabulary; any minor errors may be noticeable, but do not obscure meaning. • Convey mostly well-supported summaries, explanations, and opinions, including both concrete and abstract information, with generally well-controlled organization and cohesion; lapses may occur, but they rarely impact overall comprehensibility.
<p>High-Intermediate Score range 20–24 CEFR Level B2</p>	<p>Test takers who receive a Speaking section score at the High-Intermediate level are typically able to communicate effectively on most general or familiar topics, and to make themselves understood when discussing more complex or academic topics.</p> <p>Test takers who score at the High-Intermediate level typically can</p> <ul style="list-style-type: none"> • Speak clearly and without hesitancy on general or familiar topics, with overall good intelligibility; pauses and hesitations (to recall or plan information) are sometimes noticeable when more demanding content is produced, and any mispronunciations or intonation errors only occasionally cause problems for the listener. • Produce stretches of speech that demonstrate control of some complex structures and a range of vocabulary, although occasional lapses in precision and accuracy may obscure meaning at times. • Convey sufficient information to produce mostly complete summaries, explanations, and opinions, but some ideas may not be fully developed or may lack elaboration; any lapses in completeness and cohesion may at times affect the otherwise clear progression of ideas.
<p>Low-Intermediate Score range 16–19 CEFR Level B1</p>	<p>Test takers who receive a Speaking section score at the Low-Intermediate level are typically able to talk about general or familiar topics with relative ease.</p> <p>Test takers who score at the Low-Intermediate level typically can</p> <ul style="list-style-type: none"> • Speak clearly with minor hesitations about general or familiar topics; longer pauses are noticeable when speaking about more complex or academic topics, and mispronunciations may obscure meaning at times. • Produce short stretches of speech consisting of basic grammatical structures connected with “and,” “because” and “so”; attempts at longer utterances requiring more complex grammatical structures may be marked by errors and pauses for grammatical planning or repair; use vocabulary that is sufficient to discuss general or familiar topics, but limitations in range of vocabulary sometimes result in vague or unclear expression of ideas. • Convey some main points and other relevant information but summaries, explanations, and opinions are sometimes incomplete, inaccurate, and/or lack detail; long or complex explanations may lack coherence.
<p>Basic Score range 10–15 CEFR Level A2</p>	<p>Test takers who receive a Speaking section score at the Basic level are typically able to communicate limited information about familiar, everyday topics.</p> <p>Test takers who score at the Basic level typically can</p> <ul style="list-style-type: none"> • Speak slowly and carefully so that they make themselves understood, but pronunciation may be strongly influenced by the speaker’s first language and at times be unintelligible; speech may be marked by frequent pauses, reformulations, and false starts. • Produce mostly short utterances, connecting phrases with simple linking words (such as “and”) to make themselves understood; grammar and vocabulary are limited, and frequent pauses may occur while searching for words. • Convey some limited information about familiar topics; supporting points and/or details are generally missing, and main ideas may be absent, unclear, or not well connected.
<p>Below Basic Score range 0–9</p>	<p>Test takers with a Speaking section score below 10 have not yet demonstrated proficiency at the Basic level.</p>

Level	Writing Section
<p>Advanced Score range 24–30 CEFR Level C1</p>	<p>Test takers who receive a Writing section score at the Advanced level are typically able to write in English on a wide range of academic and nonacademic topics with confidence and clarity.</p> <p>Test takers who score at the Advanced level typically can</p> <ul style="list-style-type: none"> • Produce clear, well-developed, and well-organized text; ungrammatical, unclear, or unidiomatic use of English is rare. • Express an opinion on a controversial issue, and support that opinion with appropriate details and explanations in writing, demonstrating variety and range of vocabulary and grammatical structures. • Select important information from multiple sources, integrate it, and present it coherently and clearly in writing, with only occasional minor imprecision in the summary of the source information.
<p>High-Intermediate Score range 17–23 CEFR Level B2</p>	<p>Test takers who receive a Writing section score at the High-Intermediate level are typically able to write in English well on general or familiar topics. When writing about complex ideas or ideas on academic topics, they can convey most of the main ideas.</p> <p>Test takers who score at the High-Intermediate level typically can</p> <ul style="list-style-type: none"> • Produce summaries of multiple sources that include most of the main ideas; some important ideas from the sources may be missing, unclear, or inaccurate. • Express an opinion on an issue clearly; some ideas and explanations may not be fully developed and lapses in cohesion may at times affect a clear progression of ideas. • Write with some degree of facility; grammatical mistakes or vague/incorrect uses of words may make the writing difficult to follow in some places.
<p>Low-Intermediate Score range 13–16 CEFR Level B1</p>	<p>Test takers who receive a Writing section score at the Low-Intermediate level are typically able to produce simple written texts in English on general or familiar topics.</p> <p>Test takers who score at the Low-Intermediate level typically can</p> <ul style="list-style-type: none"> • Produce a simple text that expresses some ideas on an issue, but the development of ideas is limited because of insufficient or inappropriate details and explanations. • Summarize some relevant information from multiple sources, but important ideas from the sources are omitted or significantly misrepresented, especially ideas that require unfamiliar vocabulary or are complex. • Write with limited facility, with language errors obscuring connections or meaning at key junctures between ideas in the text.
<p>Basic Score range 7–12 CEFR Level A2</p>	<p>Test takers who receive a Writing section score at the Basic level are typically able to communicate very basic information in written English.</p> <p>Test takers who score at the Basic level typically can</p> <ul style="list-style-type: none"> • Produce some text that is related to the topic, but with little detail and/or lack of organization. • Convey some information from the sources or some ideas on an issue, but grammatical errors, unclear expressions, and/or poor sentence structure make their writing difficult to comprehend.
<p>Below Basic Score range 0–6</p>	<p>Test takers with a Writing section score below 7 have not yet demonstrated proficiency at the Basic level.</p>