Understanding the Recent Changes to the TOEFL iBT

Zachary Kelly, Asia University

Abstract

On August 1st, 2019 Educational Testing Service (ETS) made the first changes to the Test of English as a Foreign Language (TOEFL) Internet-based Test (iBT) since its international rollout in 2006. These changes affect test length, question types, and score reporting options. This paper will contextualize and describe these changes by: outlining the history of the TOEFL; discussing the uses of the TOEFL in the Japanese market; going over its recent changes; and examining the test in its current form. Readers who work in the educational field should walk away with a clear understanding of the current iteration of the TOEFL iBT, which may be helpful to their pedagogy, student advising, and general professional development.

Introduction

The TOEFL is created and administered by ETS, an educational non-profit organization based in Princeton, NJ, USA, which is the maker of several high-stakes standardized tests, including: the Test of English for International Communication (TOEIC); the Graduate Record Examination (GRE); and, in partnership with the College Board, the Scholastic Assessment Test (SAT) (ETS Products and Services, n.d.; Lawrence, et al., 2003).

On May 22, 2019, ETS announced via official press release that the TOEFL iBT would change in three ways, effective August 1, 2019 (Educational Testing Service, 2019). This was newsworthy, as the TOEFL iBT had not undergone any significant revisions since its inception in 2006. The TOEFL continues to be a popular test in Japan and this paper will consider the use of the TOEFL currently and in the future from a Japanese perspective. However, the primary aim of this paper is to familiarize readers with the TOEFL both in a general sense and with an emphasis on its recent changes.

A Brief History of the Test of English as a Foreign Language

The TOEFL was developed in the early 1960s to measure the English proficiency of non-native speakers planning to study at institutions where English is the medium of instruction. Norris (2020) identifies three stages in the test construct and content since its first administration in 1964, as outlined in Table 1.

Table 1.The Evolution of the TOEFL Test Construct and Content Over Three Stages

Stages	Construct	Content
1. The First TOEFL Test 1964-1979	Discrete components of language skills and knowledge	Multiple-choice items assessing vocabulary, reading comprehension, listening comprehension,
		knowledge of correct English structure and grammar
2. A Suite of TOEFL Tests 1979-2005	Original constructs (listening, reading, structure, and grammar) retained but two additional constructs added: writing and speaking ability	In addition to multiple- choice items assessing the original constructs, separate constructed-response tests of writing and speaking: the TWE test and the TSE test

The TOEFL iBT Test	Communicative	Academic tasks that require
2005-Present	competence—the ability to put language knowledge to	the integration of receptive and productive skills, such
	use in relevant academic	as listening, reading, and
	contexts	writing or speaking, as well as multiple-choice items for
		listening and reading

(p. 5)

Norris attributes the evolution of the test to changes in second language acquisition (SLA) theory, as well as technological advances.

The first test in Table 1 (Stage 1) was a product of prevailing ideas about SLA that claimed that language proficiency could be effectively measured by judging competency in discreet components, such as listening and reading comprehension, and grammar knowledge. The test was paper-based, multiple-choice, and did not provide opportunities to demonstrate competence in productive skills (speaking and writing).

Stage 2 saw two significant changes to the TOEFL test. The first change, as indicated in Table 1 (under Content) was the inclusion of separate constructs to measure productive skills: The Test of Spoken English (TSE) and, in 1986, The Test of Written English (TWE). The TWE and TSE were optional add-ons to the original test that were offered to a relatively small number of test-takers. For example, the TSE was originally requested of international graduate students being considered for teaching assistant positions; by contrast, the TWE was "only offered at five select administrations per year" (Norris, 2020, p. 6). The second significant change not only involved content, but also the method of delivery. Originally, the TOEFL had been administered as a paper and pencil, multiple-choice exam, but in 1998, the TOEFL computer-based test (TOEFL CBT) was introduced, although the test still replicated the earlier format and questions. Further, an essay was now a required and integrated component of all administrations of the exam.

Table 1 reveals that in 2005 ETS introduced the TOEFL iBT (Stage 3) and began its international rollout in 2006. This third stage was a radical revision, barely recognizable from the original version. An emphasis was put on communicative competence, including *integrated tasks* that required test-takers to synthesize information from multiple sources. Additionally, six speaking tasks and two essays were included. The length of the listening passages was increased, but note-taking was now permitted. Discrete components of language skills were de-emphasized. For example, previous versions of the test included grammar error correction questions. The TOEFL iBT eliminated such questions that focused

on syntactical structure, and instead refocused attention on assessing whole language acquisition and productive competency.

In addition to the shifts in content reflecting prevailing SLA research, evolutions in technology changed the administration and eventually the scoring of the test. For example, although the test started to be administered on individual computer workstations, using software programs, in 1998 with the TOEFL CBT, the current TOEFL iBT streams data through the internet. The administration and data collection are between ETS and the test-taker. One unexpected recent advantage of this method is the ability, due to the current COVID emergency, for test-takers to take the exam at home without entering a test center (TOEFL iBT Home Edition, 2020). In addition, regarding scoring changes, the Speaking and Writing sections of the TOEFL test, historically scored by human raters, are now scored with a combination of human raters and AI scoring (Understanding Your TOEFL iBT Scores, n.d.).

The TOEFL in Japan

The TOEFL has been administered in Japan for nearly 40 years. The International Educational Exchange Council (CIEE) was commissioned by ETS to be the Japan Secretariat beginning in 1981, and by 1990, there were over 100,000 Japanese examinees (International Educational Exchange Council [CIEE], n.d.). In 2005 and 2006, the last year ETS publicly released data on the number of test-takers by country (Labi, 2010), there were 78,635 examinees in Japan between June 2005 and June 2006 (Test and Score Data Summary for Computer-based and Paper-based Tests, 2007). This number is dwarfed by the 1,680,000 people who took the ubiquitous TOEIC in 2009, as reported in Takahashi (2012). However, it is important to keep in mind that these tests measure different language proficiencies and are used for different purposes. As stated by ETS in What are the differences between TOEIC and TOEFL tests? (2019), "The TOEIC tests measure proficiency in English relevant to the global workplace whereas the TOEFL tests measure the academic communication skills in English" (para. 1). Traditionally, TOEFL examinees were current or prospective students intending to study abroad, while TOEIC examinees were students, or those already in the workforce, who needed to show English language proficiency to be hired or promoted. This delineation is not completely tidy. As Takahashi (2012) notes, "The TOEIC test, which was initially developed for measuring general English ability of corporate workers, has come into wide use in educational institutions in Japan including universities, high schools and junior high schools" (p.127). At the same time, some Japanese professionals submit TOEFL iBT

scores to prospective employers. Since 2012, the TOEFL iBT has been increasingly accepted in application packets by some *domestic* universities as a demonstration of English proficiency (CIEE, n.d.-b). Finally, it is important to distinguish between the TOEFL iBT and the institutionally based TOEFL (TOEFL ITP), which is popular in Japan. According to CIEE, the TOEFL ITP is used by over 500 groups and 220,000 individuals throughout Japan (n.d.-a). It is administered by companies or educational institutions including high schools, universities, graduate programs, and corporations for "in-house" decisions involving admissions, placement, progress monitoring, and scholarships (CIEE, n.d.-a).

Recent Changes to the TOEFL iBT

On May 22, 2019, ETS announced via official press release, the first major revision to the TOEFL iBT since its establishment in 2006. Effective August 1, 2019, the test was shortened from approximately 4 to 3.5 hours. The changes affect three of the four sections:

- The Reading section has fewer questions per passage (now 10, from 12-14);
- The Listening section has one or two fewer lectures;
- The Speaking section now has four questions instead of six: Questions 1 and 5
 were removed, and the remaining questions were renumbered 1 to 4.

(Frequently Asked Questions About the Shorter TOEFL iBT Test, 2019)

Table 2 offers a visual representation of the contrast between the older and most recent version of the TOEFL iBT. Changes to the Reading section reduced the number of questions per passage. This should be seen as advantageous to test-takers as they will have slightly more time: 1.8 minutes per question compared to 1.4 minutes per question on the older TOEFL iBT (54 minutes/ 30 questions = 1.8 min./ Q). The question types have not changed, nor have the nature of the passages.

 Table 2.

 Comparison of the traditional and current versions of the TOEFL iBT

	Traditional TOEFL iBT (2006- July 2019)	Current TOEFL iBT (after August 1, 2019)
Reading	3-4 passages; 12-14 questions per passage; 60-80 minutes	3-4 passages; 10 questions per passage; 54-72 minutes

Listening	4-6 lectures with 6 questions each; 2-3 conversations with 5 questions each; 60-90 minutes	3-4 lectures with 6 questions each; 2-3 conversations with 5 questions each; 41-57 minutes
Speaking	6 speaking tasks: 2 independent, 4 integrated	4 speaking tasks: 1 independent, 3 integrated
Writing	2 writing tasks: 1 integrated (20 minutes), 1 independent (30 minutes)	2 writing tasks: 1 integrated (20 minutes), 1 independent (30 minutes) No change
Scoring	120 points; 30 points per section	120 points; 30 points per section No change
Time	4 hours	3.5 hours

Changes to the Listening section may be seen as a mixed bag for test-takers. On one hand, slightly less time is allotted: 1.5 minutes per question on the new test compared to 1.8 minutes on the older test. However, the reduced length of the listening section may be interpreted positively in that it may reduce the likelihood of test-taker fatigue. There are no changes to the types of listening material or questions.

Here it is worth pointing out the relationship between the Reading and Listening sections on any given test date. According to ETS's *Official Guide to the TOEFL Test*, 5th Edition (2017), each administration of the TOEFL iBT includes "dummy passages" on either the reading or listening section of the test. These passages are unscored and are being tested for inclusion on future exams. If a test-taker has a long reading section (four passages; 72 minutes), he or she will have a short listening section (two conversations, three lectures; 51 minutes) and vice versa. Whichever section is longer will include some unscored "dummy passages," so, using the above example, one reading passage would be a "dummy" and the accompanying 10 questions would be unscored.

The most significant changes announced in the recent update are to the speaking section. As noted above, prior to August 1, 2019, test-takers had six speaking tasks; now they only have four. At first glance, it appears this would make the updated test easier since there are fewer tasks to prepare for; however, because of which tasks were eliminated and which are still included, some test-takers may find the new test more challenging in this regard. This is because the remaining speaking tasks are more heavily skewed towards academic content (2/4 tasks vs. 2/6 tasks on the previous version of the test). Test-takers now have fewer

chances to demonstrate speaking proficiency; also, half of their chances require them to summarize academic content, which may be taxing for some.

In addition, effective August 1, 2019, ETS has also updated how scores are reported, by introducing a system called *MyBest* scores, which will present the highest scores test-takers have earned on each of the four sections in the past two years. Per ETS, "This practice provides opportunities for test takers to leverage their best performance on each section(.)" (*My Best Scores, A Rationale for Using TOEFL iBT Superscores*, 2019, p.2). In other words, if a test-taker attempted the TOEFL three times with the example scores in Table 3 (shown below), the reported score would be 100, the combined high scores from each section.

 Table 3.

 Example scores for MvBest score reporting

Test section	Attempt 1	Attempt 2	Attempt 3	High score
Reading	18	24	22	24
Listening	24	24	26	26
Speaking	26	22	24	26
Writing	18	22	24	24

Reported score: 100

There is no limit to how many times the test can be taken as long as it is not taken multiple times in a 3-day period, and test-takers average a nearly 10-point score increase after three attempts according to ETS 5 Tips for retaking the TOEFL test (2019). However, it is worth noting that not all universities accept the MyBest score and will instead consider only the highest total score on a given test day, so it is important for test-takers to apprise themselves of the score-reporting options available for their chosen institutions(s).

A Closer Look at the New TOEFL iBT

This section aims to provide the reader with a succinct understanding of the layout of the new TOEFL iBT by referencing the most recent *ETS Official Guide to the TOEFL Test* (2017).

Reading

The first section of the TOEFL iBT is Reading. This section includes three or four academic passages, approximately 700 words in length. The passages are slightly edited excerpts from college-level textbooks that are introductions to a discipline or topic and assume no specialized knowledge on the part of the test-taker. Topics are wide-ranging and

include the Humanities, Science, and Social Sciences; they "are classified into three basic categories based on author purpose: (1) Exposition, (2) Argumentation, and (3) Historical" (p. 37). Each passage is followed by ten multiple-choice questions, all of which are worth one point, except the final question which is worth two to three points. There are ten question types that measure various reading skills, including the test-takers' ability to: comprehend information directly stated in the passages; make inferences; paraphrase; understand rhetorical purpose; and determine organizational scheme. Test-takers may skip questions and skip between passages within the allotted time (54 minutes in the case of three passages; 72 minutes for four passages). The most significant change to the Reading section is the reduction in the number of questions per passage, and the fact that test-takers now have slightly more time to answer each question (as noted previously). The total number of correct responses is converted to a scaled score between one and 30.

Listening

In the Listening section of the TOEFL iBT, there are two formats: conversations and lectures. Each conversation or lecture is between three to six minutes, and test-takers are expected to take notes for reference when answering the questions. Conversations take place between one male and one female, one of whom is a student. The conversations are meant to mimic encounters that are typical of college life on American campuses, such as when a student visits a professor during office hours to get clarification about an assignment, or a service encounter where a student visits the financial aid office to inquire about a payment. Each conversation is followed by five multiple-choice questions.

Lectures are approximately five minutes and cover various topics in the Arts, Life Science, Physical Science, and Social Sciences. Specialized, prior knowledge is not assumed; the information needed to answer the questions comes from the lecture. The lecture may feature a professor only, or a professor interacting with one or more students who are asking or answering a question related to the lecture topic, or interjecting for the purpose of further clarification. Each lecture is followed by six multiple-choice questions.

There are three types of listening comprehension questions: basic comprehension, pragmatic understanding, and connecting information questions. Most of the questions are multiple-choice, with four answer choices and one correct answer. However, a test-taker may also encounter the following variables:

 multiple-choice questions with more than one correct answer (for example, two answers out of four choices, or three answers out of five choices)

- questions that require the test-taker to put in order events or steps in a process
- questions that require the test-taker to match objects or text to categories in a table (p. 121)

Unlike in the Reading section, test-takers cannot skip between passages or questions. The total number of correct responses is scaled and assigned a score on a 30-point scale. The most significant change to the Listening section is the reduction in the number of lectures.

At the end of the Listening section, test-takers are given a ten-minute break before continuing with the Speaking section.

Speaking

There are four speaking tasks (formerly six), and they always appear in the order indicated in Table 4.

Table 4.Speaking Section Overview

Task Type	Task Description	Timing
Independent Task		
Choice Integrated Tasks	This question asks the test-taker to make and defend a personal choice between two contrasting behaviors or courses of action.	Preparation time: 15 seconds Response time: 45 seconds
2. Campus Situation Topic: Summarize and Explain	 A reading passage (80-110 words) presents a campusrelated issue. A listening passage (60-80 seconds) comments on the issue in the reading passage. The question asks the test-taker to summarize the speaker's opinion within the context of the reading passage. 	Preparation time: 30 seconds Response time: 60 seconds
3. Academic Course Topic: General/Specific	A reading passage (80-110 words) broadly defines a term, process, or idea from	Preparation time: 30 seconds Response time: 60 seconds

	•	term, process, or idea from the reading passage. The question asks the test- taker to combine and convey important information from the reading passage and the lecture excerpt.	
4. Academic Course Topic: Summary	•	The listening passage is an excerpt from a lecture (90-120 seconds) that explains a term or concept and gives concrete examples to illustrate that term or concept. The question asks the test-taker to summarize the lecture and demonstrate an understanding of the relationship between the examples and the overall topic.	Preparation time: 20 seconds Response time: 60 seconds
Total time: 17 minutes		-	

Adapted from ETS Official Guide to the TOEFL Test (2017), p. 18

Speaking Task 1 is considered an *independent task* since it asks test-takers to only respond to a prompt using their own ideas to support their position. An example prompt may be, "Some students study for classes individually. Others study in groups. Which method of studying do you think is better for students and why?" (p. 169). Test-takers have 15 seconds to prepare and 45 seconds to respond. Like all speaking tasks on the TOEFL iBT, responses are given into a headset microphone. The recordings are assessed by human and machine raters. Scores are given on a scale of one to four for each task based on: *delivery* (pronunciation, pacing, intonation); *language use* (grammar and vocabulary accuracy and complexity); and *topic development* (completeness and cohesion of response, effective use of time). Scores are totaled and a combined scaled score of one to 30 is given for the speaking section.

Speaking Tasks 2 - 4 are *integrated tasks* where test-takers synthesize information from a listening, or a combined reading and listening, and summarize the provided content; they are not to give their opinion or include their outside knowledge on these tasks. Sixty seconds are allotted for Speaking Tasks 2 - 4. Test-takers may take notes during all portions of the speaking section.

Writing

There are no changes to the Writing section of the TOEFL iBT. The Writing section consists of two essays: integrated writing and independent writing. Essay responses are typed without the assistance of word-processing software such as spell check. The essays are assessed by human and machine raters and assigned a score between one and five based on development, organization, and language use. Combined essay scores are transferred to a 30-point scale. As shown in Table 5, the first essay asks test-takers to integrate information from an academic reading and a lecture on the same topic. Test-takers are allowed to take notes. Twenty minutes are allotted for this task. The second essay offers a prompt where test-takers have 30 minutes to respond to and support with evidence based on their personal knowledge and experience.

Table 5.Writing Section Overview

Task Type	Task Description	Timing and Word Count
Integrated Writing		
1. Read/ Listen/Write	Read a short text (250-300 words) on an academic topic. Listen to a professor discuss the same topic from a different perspective. Write a summary connecting key points in the listening passage to key points in the reading.	Reading time: 3 minutes Listening time: about 2 minutes Writing time: 20 minutes Suggested word count: 150-225 words
Independent Writing		
2. Writing from Knowledge and Experience	Test-takers write an essay that states, explains, and supports their opinion on an issue. Typical essay questions begin with statements such as: "Do you agree or disagree with the following statement? Use reasons and specific details to support your answer." Or "Some	Writing time: 30 minutes Suggested word count: 300+ words

you prefer/ agree with? Give reasons and specific details."

Adapted from ETS Official Guide to the TOEFL Test (2017), p. 20

Looking Ahead

Will the TOEFL iBT continue to be popular in Japan? For the foreseeable future, yes. While competing exams such as the International English Language Testing System (IELTS) may continue to consume some market share previously enjoyed by the TOEFL iBT (Abe et al., 2018), and a decline in the number of potential university-aged test-takers is imminent due to the well-documented demographic shift in Japan (Ministry of Education, Culture, Sports, Science and Technology- Japan [MEXT], n.d.), there is little reason to believe that the test will disappear any time soon. It is still the preferred test for international applicants to most American universities, and the United States continues to be the top destination for Japanese students who study abroad according to Japan Association of Overseas Students (2017). Perhaps more significant, though, is the recent decision by MEXT to change the admission guidelines for domestic universities. As soon as 2020, universities will begin accepting the TOEFL iBT (as well as a number of other English exams including IELTS) as a demonstration of English proficiency. While it is not clear that all universities will begin this transition immediately (McCrostie, 2017), over time the policy is certain to increase the number of TOEFL iBT test-takers as this introduces the test to a whole new demographic. As a result, it is not unreasonable to assume that an increasing number of incoming university freshmen will be familiar with the exam. As language professionals, it is incumbent on us to stay abreast of updates and changes in assessment that affect our field.

References

- 5 Tips for retaking the TOEFL test. (2019, October 21). ETS. https://www.toeflgoanywhere.org/blog/5-tips-retaking-the-toefl-test
- Abe, O., Matsuzaki, A., Wakita, M., & Koizumi, R. (2018). Which Should Be Dominant in Japan, the TOEFL iBT or the IELTS? (Course name: What Does a Test Measure?): Problem-Based Learning to Encourage Active Learning and Teamwork Among First Year Medical Students - Student Reports -. *Juntendo Medical Journal*, 64(2), 101– 104. https://doi.org/10.14789/jmj.2018.64.JMJ16-WN13
- CIEE International Education Exchange Council. (n.d.). *TOEFL テストの歴史* | *TOEFL iBT テストとは* | *受験者の方へ* | *TOEFL テスト日本事務局(CIEE Japan*). TOEFL テスト 受験者の方へ | TOEFL テスト日本事務局(CIEE Japan). https://www.toefl-ibt.jp/test_takers/toefl_ibt/history.html
- CIEE International Education Exchange Council. (n.d.-a). *TOEFL ITP テスト(団体向け TOEFL テスト) | TOEFL テスト日本事務局*. TOEFL ITP テスト-TOEFL テスト日本事務局. September 29, 2020, http://www.cieej.or.jp/toefl/itp/
- CIEE International Education Exchange Council. (n.d.-b). 活用状況全般| 日本の大学入試での活用| 受験者の方へ | TOEFL テスト日本事務局 (CIEE Japan) . TOEFL テスト 受験者の方へ | TOEFL テスト日本事務局 (CIEE Japan) . 2020, https://www.toefl-ibt.jp/test_takers/nyushi/overview.html
- Educational Testing Service. (2017). The Official Guide to the TOEFL Test with DVD-ROM, Fifth Edition (5th ed.). McGraw-Hill Education.
- Educational Testing Service. (2019, May 22). The TOEFL Test Experience Just Got Better [Press release]. https://news.ets.org/press-releases/toefl-test-experience-just-got-better/
- ETS Products and Services. (n.d.). ETS.

 https://www.ets.org/products/?WT.ac=etshome_products_menu1_180417
 Frequently Asked Questions About the Shorter TOEFL iBT Test. (2019). ETS.

 https://www.ets.org/s/toefl/pdf/faqs shorter toefl ibt test.pdf
- Japan Association of Overseas Students. (2017). JAOS 2017 Survey on the Number of Japanese Studying Abroad. JAOS. https://www.jaos.or.jp/wpcontent/uploads/2018/01/JAOS-Survey-2017_Number-of-Japanese-studyingabroad180124.pdf
- Labi, A. (2010, July 25). English-testing companies vie for slices of a growing market. *The Chronicle of Higher Education*. https://www.chronicle.com/article/english-testing-companies-vie-for-slices-of-a-growing-market/
- Lawrence, I. M., Rigol G.W., Van Essen T., & Jackson C. A. (2003). *A Historical Perspective of the Content of the SAT* (Research Report number 2003-3). The College Board. https://www.ets.org/Media/Research/pdf/RR-03-10-Lawrence.pdf

- McCrostie, J. (2017, July 5). Spoken English tests among entrance exam reforms Japan's students will face in 2020. *Japan Times*. https://www.japantimes.co.jp/community/2017/07/05/issues/spoken-english-tests-among-entrance-exam-reforms-japans-students-will-face-2020/
- Ministry of Education, Culture, Sports, Science and Technology-Japan. (MEXT) (n.d.).

 Current status and issues of education in Japan.

 https://www.mext.go.jp/en/policy/education/lawandplan/title01/detail01/sdetail01/137
 3809.htm
- MyBest Scores: A Rationale for Using TOEFL iBT Superscores. (2019). ETS. https://www.ets.org/s/toefl/pdf/mybest_su.pdf
- Norris, J. (2020). TOEFL Research Insight Series, TOEFL Program History (Volume 6). [white paper] Educational Testing Service. https://www.ets.org/s/toefl/pdf/toefl_ibt_insight_s1v6.pdf
- Takahashi, J. (2012). An overview of the issues on incorporating the TOEIC test into the university English curricula in Japan. *Tama University Global Studies Department Bulletin*, 4(3), 127–138.
- Test and Score Data Summary for Computer-based and Paper-based Tests (July 2005-June 2006). (2007) ETS. https://www.ets.org/Media/Research/pdf/TOEFL-SUM-0506-CBT.pdf
- TOEFL iBT Home Edition. (2020). ETS. https://www.ets.org/s/cv/toefl/at-home/?utm_campaign=AtHomeTOEFL20&utm_medium=TGA&utm_source=HeroI mage&utm_content=LearnMoreCTA&_ga=2.117470153.1377505112.1606443404-598939321.1606443404
- Understanding Your TOEFL iBT Scores (For Test-takers). (n.d.). ETS. https://www.ets.org/toefl/test-takers/ibt/scores/understanding/
- What are the differences between TOEIC and TOEFL tests? (2019, October 22). ETS Global. https://www.etsglobal.org/fr/en/help-center/score-usage/differences-between-toeic-and-toefl-tests