toefl itp

# English proficiency trends and educational policy context: Examining TOEFL ITP scores in Japan from 2012-2024

John M. Norris
Hiroshi Kosuga
Shuken Shiozaki

*July 2026*

ETS Japan Research Report

RESEARCH REPORT

# English proficiency trends and educational policy context: Examining TOEFL ITP scores in Japan from 2012-2024

John M. Norris[1], Hiroshi Kosuga[2], Shuken Shiozaki[3]

1 ETS Japan, Tokyo, Japan
2 ETS Japan, Tokyo, Japan
3 ETS Japan, Tokyo, Japan

## Executive summary

English language education in Japan has a lengthy history. The past several decades in particular have featured multiple educational policy initiatives intended to improve the English language proficiency of students in Japanese schools and universities. At the primary and secondary school levels, since 2012, English language education reform has been promoted officially through (a) new courses of study that emphasize the development of communicative competence, including productive as well as receptive abilities, (b) the development of activity-based and task-oriented teaching practices and learning materials, and (c) the potential adoption of high-stakes assessments (e.g., for university admissions) that test all four communicative skills (listening, reading, speaking, writing). At the university level, over the same period, improving English language proficiency has been promoted by (a) emphasizing the development of language ability for practical purposes and intercultural competence, (b) encouraging the internationalization of university campuses and coursework, and (c) fostering opportunities for extended study abroad.

During this time of educational policy initiatives and reform efforts, the TOEFL ITP test has been administered annually to hundreds of thousands of high school and university students, as well as other test takers in Japan. The TOEFL ITP test measures English proficiency in the skills of Listening, Structure and Written Expression, and Reading, and it reports scores for each skill and overall on scales that have been mapped to important global proficiency frameworks (e.g., the CEFR). This report analyzes the English proficiency levels exhibited by students at Japanese high schools and universities that administered the TOEFL ITP during the period of 2012-2024. The resulting patterns offer useful insights into Japanese students' English proficiency differences over time, reflecting learning outcomes that are potentially associated with large-scale educational policy initiatives.

Key findings include:

- TOEFL ITP total scores were considerably higher in 2024 than in 2012 for both high school and university undergraduate students in Japan, with both groups falling in the CEFR B1 proficiency range by the end of the time period.

- For high school students, those with Japanese as a first language (L1) generally achieved higher TOEFL ITP total and section scores than those with other first languages; Japanese L1 students scored above the CEFR B1 level on total and Listening section scores by the end of the time period, while students with other first languages did not; over the time period, high school students demonstrated the greatest positive differences in section scores for the skill of Listening.

- For university undergraduate students, those with Japanese as a first language (L1) generally

achieved higher TOEFL ITP total and section scores than those with other first languages; Japanese L1 students showed gradual, positive score differences throughout the time period investigated, with the greatest differences in sections scores for the skill of Reading.

- Comparisons of Japanese L1 students at Super Global Universities (SGU) versus other universities showed that SGU students consistently outperformed other university students on TOEFL ITP total and section scores through the time period; both total and section scores for SGU students demonstrated higher magnitude differences between 2012 and 2024 than scores for other university students.

- By 2024, comparisons of Japanese L1 students' TOEFL ITP total scores with the worldwide TOEFL ITP test-taker population showed that high school students scored at the 29th percentile, university undergraduate students scored at the 49th percentile, and SGU students scored at the 61st percentile.

These findings show considerable positive iterations in English language proficiency, as measured by the TOEFL ITP test, for Japanese learners of English at high schools and universities in Japan. These iterations occurred during a period of intensive educational policy initiatives and reform efforts intended to promote English language development and related outcomes. While causal links cannot be confirmed, the synergies of English language proficiency test data and policy context are, at a minimum, encouraging.

## Introduction

The development of English language proficiency has been promoted as a high-priority goal for Japanese learners of all kinds, and especially for school and university learners, over a considerable period of modern history. Recent decades have seen the intensification of efforts at the governmental policy level to encourage reforms in language education with the ultimate goals of improving the abilities of Japanese citizens to communicate in English (and other foreign languages). These policy initiatives have accompanied parallel goals of modernization, globalization, and social and economic stability, English proficiency being perceived as a key factor in academic and workplace competence. They have also occurred within societal circumstances that are increasingly volatile and subject to rapid changes in the roles played by technological advances in the global information economy.

Within this exciting and challenging environment, public goals for English learning have coalesced around the notion of integrated abilities across both receptive and productive skills, as well as an emphasis on development of practical abilities to use English for various communicative purposes. Accordingly, pressures have mounted for reforms in language teaching and learning, at schools and universities in particular, and intended learning outcomes encapsulate a renewed interest in measuring the proficiency of Japanese learners.

Among various possibilities for assessing the English language proficiency of Japanese learners, the TOEFL ITP (Institutional Testing Program) test has been adopted by numerous institutions over the past several decades. Following the launch of the first official TOEFL test in 1965 in the U.S., TOEFL ITP was later developed by ETS as an alternative to high-stakes, high-security English language testing (e.g., TOEFL PBT, CBT, and eventually iBT). The primary intended uses for TOEFL ITP included common purposes such as placement testing and progress monitoring, these considered to be lower stakes than the purposes of the main TOEFL test.[1] The ITP test consists of three sections covering English abilities in: Listening, Structure and Written Expression, and Reading. Scores are reported separately for each of these skills and in the form of a total score, using the original TOEFL score scale (310-677). A separate test of Speaking ability was introduced in 2022, available in digital versions of the ITP test only, but scores for Speaking are not included in the calculation of total scores for the main test.

The TOEFL ITP test has experienced widespread use across diverse institutional contexts since its introduction to Japan in 1981, including governmental agencies, private language schools, corporations, and others. By far the most frequent users of the test in Japan are universities and high schools, where placement into English language courses and monitoring of English proficiency achievements are very common uses.

The focus of the current report is on patterns in the scores of TOEFL ITP test takers at high schools and universities in Japan during the period of 2012 to 2024. In the following sections, we first review the societal and policy landscape surrounding English language education in Japan during the past several decades, paying particular attention to efforts at improving English proficiency among students in the school and university levels. We then present findings from the analysis of TOEFL ITP scores, collected as part of the regular administration of the test, from over one million test takers in Japanese high schools and undergraduate university programs. Our intention with these analyses is not to imply cause-effect relationships between educational policies and test scores—the nature of the data and varying population samples do not support such inferences. However, we believe that the overarching patterns in these very large data sets provide important insights into the potential development of aspects of English language proficiency in parallel with the intentions of educational policies. At a minimum, they offer one source of observations for reflecting on the English proficiency levels of Japanese learners of English in the recent era.

## Educational policy context in Japan

### Historical Development of English Language Education in Japan (1800s – 2010s)

The trajectory of English language education in Japan has been shaped over several centuries by broader societal demands and associated governmental actions, including modernization, university entrance examinations, and internationalization. During the Edo period, under the Tokugawa shogunate's policy of national isolation (sakoku), the use of English was highly limited. However, entering the 19th century, Japan experienced a rapid increase in arrivals of Russian and British vessels for commerce and diplomacy purposes, heightening the need for practical skills in foreign languages such as Russian and English, particularly for interpretation and translation.[2,3]

Along with other modernization efforts of the Meiji period (1868–1910s), a modern school system was established in Japan, and foreign language education was institutionalized at the secondary level and above as a means of acquiring Western knowledge and technologies. During the Taishō period through prewar Shōwa (1910s–1940), English became established as a key subject for university entrance requirements, although instruction was dominated by the grammar-translation method, resulting in limited development of speaking ability. During World War II (1940–1945), English learning was restricted due to its status as an "enemy language." [2,3]

In the postwar period, under the influence of the General Headquarters (GHQ), educational reforms such as the 6-3-3-4 system, which set expectations for years of schooling from primary through tertiary education, were implemented. Language teaching approaches shifted to some extent from a grammar-translation focus toward an emphasis on listening and speaking skills. However, during the period of rapid economic growth and increased competition for achieving entrance into selective universities (1970s–1990s), English education was pressured to prioritize reading comprehension and grammatical accuracy due to high-stakes examination emphases. At the same time, internationalization progressed through initiatives such as the ALT system and the JET Programme (established in 1987), leading to increased acknowledgement of the need for "usable English." [2,3]

From the 2000s onward, growing societal demand for practical English proficiency led to a focus on the development of English communication skills, explicitly reflected in the national Course of Study for Japan.[4] In 2008, "Foreign Language Activities" were introduced in the upper grades of elementary school as an effort to encourage an earlier start to language learning and learner motivation.[6,7] Since the 2020s, integration of the four skills—reading, listening, speaking, and writing—has become a central educational policy objective of the Japanese government. Language instruction was pushed to an even earlier introduction in elementary school, with "Foreign Language Activities" in Grades 3–4 and English as a regular subject in Grades 5–6. Additionally, discussions began regarding the possible adoption of external proficiency tests (e.g., TOEFL, Eiken) in university entrance examination systems.[5,6,7]

### Recent Societal Context of English Language Education (2012–2024)

The most salient societal demands on English education in the recent era have been practical applicability and integration of the four skills to achieve communicative competence. These demands are closely linked not only to educational policy but also to study abroad promotion and broader global developments. Accordingly, this section examines two dimensions: (1) governmental policy developments in English education; and (2) the relationship between English education, study abroad initiatives, and global circumstances.

***Recent Policy Developments for English Language Education***

In 2013, the Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT) released the "Implementation Plan for English Education Reform Corresponding to Globalization." This plan proposed activity-based language learning in middle elementary grades and formal instruction of language as a subject in upper grades. It also emphasized that English should be used as the medium of instruction for English language classes, particularly in senior high schools, and encouraged at the junior high level. The plan proposed outcomes measures including target proficiency levels for students and teachers, the use of external testing, and CAN-DO descriptors.[8]

These policies gained added prominence due to the anticipated Tokyo 2020 Olympics and subsequent revisions to the Course of Study, solidifying a consistent policy framework based on continuity of foreign language instruction across educational stages, four-skills integration, and English-medium instruction.[8,9]

It also outlined a timeline: system preparation from 2014, phased implementation beginning in FY2018, and full implementation by FY2020. This timeline underscored the reality that, in particular, expanding elementary school English required substantial human resource development, including teacher training, recruitment of specialist instructors, and the use of external personnel such as Assistant Language Teachers (ALTs).[8]

Beginning in March 2017, MEXT issued the revised Courses of Study (announced for elementary and junior high schools in 2017 and for high schools in 2018), identifying more clearly defined language learning outcomes ("What students will be able to do"), the three pillars of competency (knowledge and skills; thinking, judgment, and expression; motivation to learn and exhibit human qualities), and the concept of "active, interactive, and deep learning."[10]

The powerful role played by university entrance examinations in influencing language teaching and learning was also acknowledged during this time. Beginning in 2013, the Council on Rebuilding Education proposed that external English proficiency tests be introduced for university entrance testing purposes, in order to emphasize the development of all four skills and communicative competence. By 2016, MEXT initiated a plan to adopt external tests for evaluating English language proficiency at the transition from high school to university, leading to extended discussions around the feasibility of doing so. However, in 2019, the planned introduction of private-sector English proficiency tests into the Common Test for University Admissions—aimed at comprehensive four-skills assessment that resulted in improvement of high school English education—was postponed due to concerns regarding fairness, accessibility, cost, comparability across tests, and operational reliability.[11,12]

Meanwhile, the GIGA School Initiative, intended to address disparities in educational ICT infrastructure and promote "Society 5.0" learning, introduced the practice of "one device per student", significantly transforming the potential learning environment for English language instruction.[12,13]

By 2020, the new Course of Study was fully implemented in elementary and high schools, institutionalizing earlier and subject-based English education. In 2021, a revised Common Test was introduced without the uniform use of private tests. In educational settings, efforts were made to improve teaching methods aimed at teaching all four language skills, while also placing greater emphasis on the assessment of English proficiency through the use of digital technologies.[11] Concurrently, the phased introduction of the new high school curriculum began, further promoting integrated four-skills language development.[10]

The COVID-19 pandemic in 2020 highlighted the need to ensure continuity of learning, accelerating

discussions on digital learning infrastructure. By 2024, full-scale implementation of digital textbooks in English and the integration of AI-supported learning tools had begun.[11,12]

Overall, this period was structured around the symbolic timeline of the 2020 Tokyo Olympics international event, driving reforms related to integrated four-skills language learning, ICT expansion, and university entrance examinations for English.[9,12] However, challenges of various kinds arose and highlighted the need to ensure teacher quality amid increasing specialization, secure human resources, and resolve alignment between school education and desired learning outcomes in relation to powerful university entrance examinations.[13]

### *English Education, Study Abroad Promotion, and Globalization*

During the period of 2012-2024, English education reform in Japan became closely linked to the broader policy agenda of "global human resource development," within which foreign language education, expansion of study abroad, and university internationalization were treated as an integrated policy package with a focus on university undergraduate education.

The Second Basic Plan for the Promotion of Education (2013), adopted by Japan's executive Cabinet, identified foreign language education, bidirectional student mobility, and university internationalization as core pillars, setting numerical targets against the backdrop of demographic decline and intensifying global competition for university education.[14] Consequently, English education reform and study abroad promotion were positioned not only as educational policies but also as components of a national growth strategy for Japan, subject to KPI (Key Performance Indicators) evaluation.[9,14]

The "Japan Revitalization Strategy" (Abenomics), launched in 2012, included a goal of doubling the number of Japanese high school and university students studying abroad by 2020, linking economic strategy and education policy. This was operationalized through the "Tobitate! Study Abroad JAPAN" initiative, a public–private partnership aimed at fostering nationwide momentum for overseas study.[15] Despite the intentions of these and related efforts, an emphasis on quantitative targets led in many cases to a proliferation of short-term study abroad programs (e.g., with most students experiencing stays abroad of less than one month), which did not fully meet labor market expectations. In response to a 2017 recommendation by the Ministry of Internal Affairs and Communications, the Third Basic Plan for the Promotion of Education (2018) shifted policy emphasis from quantity to quality—prioritizing longer-term study abroad, practical English proficiency, and intercultural competence.[10,16]

Also at the higher education level, initiatives such as the 2012 "Project for Promotion of Global Human Resource Development" and the 2014 "Top Global University Project" (Super Global University; SGU) promoted study abroad opportunities, English-medium instruction, and institutional internationalization. Of particular interest, the SGU initiative sought to encourage universities in Japan to globalize and reform their approaches to English language education. Launched in 2014 with the specific objective of addressing declining global competitiveness and limited international presence, it provided various kinds of support to two groups of public and private Japanese universities: thirteen "Type A" universities aiming for world class status; and 24 "Type B" universities focused on broad internationalization and institutional reform.[9]

Despite encouraging developments in internationalization and practical language education, the COVID-19 pandemic in 2020 severely disrupted international mobility, undermining key foundations of these policies. In English education, emergency measures included a rapid transition to online instruction and increased use of digital materials; however, disparities in access to technology also became evident, raising concerns about educational equity.[11,12]

The Fourth Basic Plan (2023), in the post-pandemic context, acknowledged insufficient outcomes in terms of study abroad participation, English proficiency, and responsiveness to societal needs. While maintaining a focus on quality in global human resource development, it emphasized the promotion of educational digital transformation (e.g., online international exchange), ensuring learning continuity, and visualizing learning outcomes.[12]

In contemporary English education in Japanese universities, the role of English has expanded to include its use as a tool for inquiry-based learning, project-based learning (PBL), presentations, and discussions. Thus, English is increasingly positioned not merely as a means of communication, but as a tool for thinking, learning, and knowledge dissemination.[10,12]

## TOEFL ITP scores at high schools and universities in Japan: 2012-2024

### Test data and analytic approach

The data set for the current report is based on TOEFL ITP tests administered during the years 2012 through 2024 at high schools and universities in Japan. The tests were administered at these institutions for specific local purposes and following official test administration practices, and students' test scores are presumed to represent their English language proficiency at the time of test taking. Data collected during these regular test administrations included TOEFL ITP total test scores (ranging from 310-677) and skill section scores for Listening (ranging from 31 to 68), Structure and Written Expression (ranging from 31 to 68), and Reading (ranging from 31 to 67). To facilitate interpretation of score meaning in subsequent analyses, Table 1 shows the TOEFL ITP total and section scores mapped to the proficiency levels of the Common European Framework of Reference for Languages (CEFR).[17] Note that the cut scores indicate the minimum or threshold value at which the corresponding proficiency level can be claimed.

Table 1. TOEFL ITP scores mapped to CEFR levels

| CEFR Levels | Total Score | Listening Comprehension | Structure and Written Expression | Reading Comprehension |
|---|---|---|---|---|
| C1 | 620 | 62 | 64 | 60 |
| B2 | 543 | 55 | 53 | 55 |
| B1 | 433 | 46 | 43 | 41 |
| A2 | 343 | 38 | 32 | 33 |

Note: https://www.ets.org/toefl/itp/scoring.html#accordion-74c86af37a-item-6f39ad0428

In addition to test scores, data were collected regarding the following variables: test administration date; institution of test administration; type of institution (High School, University); test-taker gender, age, and first language (self-reported data). Note that within the University category, the focus here is on undergraduate students, with graduate student test-taker data not included in the analyses.

Regarding the size of the data set under investigation, the specific numbers of test takers per category or per year constitute proprietary and confidential business intelligence information. However, we can report the general parameters of the data set. First, the total number of high school and university undergraduate test takers, from whom scores and other data were collected over the twelve-year period, exceeded one million. Second, in the main analyses reported here, the minimum sample size examined for any category of test taker (e.g., high school test takers in a given year with Japanese as their first language) exceeded 10,000 for all years under investigation (unless otherwise indicated in secondary

analyses). Third, a substantial number of institutions contributed data for all years under examination, although the specific number varied for both high schools and universities across all twelve years. Table 2 shows the minimum, average, and maximum number of institutions of each type that contributed test-taker data between 2012 and 2024. Note that all analyses were conducted at the conglomerate group or type level; no analyses were conducted on individual institutions or individual test-taker data.

**Table 2. Number of institutions contributing annual test-taker data**

| | University | High School |
|---|---|---|
| min | 133 | 41 |
| mean | 184 | 59 |
| max | 231 | 90 |

The basic analytic approach adopted for the current report focused on calculating average annual TOEFL ITP test scores and displaying them graphically over the 12 years under investigation. These annual averages (for both total and skill section scores) were calculated for meaningful categories of test takers, including: high school students and university undergraduate students overall; high school students with Japanese as their first language versus Other first languages; university undergraduate students with Japanese as their first language versus Other first languages. These basic analyses allowed for an examination of the English language proficiency levels—in the form of TOEFL ITP test scores—and possible differences for distinct groups of test takers and institutional contexts over time.

In addition, within the university category, all institutions in the data set were subsequently classified by their status as either a Super Global University (according to the list of 37 universities reported by the Japanese Ministry of Education) or a non-Super Global University. Average TOEFL ITP scores were then calculated for the Japanese first-language students within each of these groups of universities for the twelve years under examination, to enable comparison of possible score differences between the categories and over time.[18]

Given the very large sample sizes for all categories under investigation, inferential statistical significance tests were deemed misleading and therefore not conducted for any comparisons made within the data. To be clear, with sample sizes exceeding $N = 10{,}000$ for any given observed average value, statistical tests at virtually any probability level (e.g., $p$-values of .001 or .0001) would result in any observed difference being found 'statistically significant'. Here, instead, we recommend examining the observed average test scores and drawing interpretations based on their practical meaning (e.g., the actual magnitude of the difference observed between two average scores), bearing in mind that any observed difference between two mean values will be statistically trustworthy.

In the following, we first present overall findings regarding the demographic characteristics of test takers in the high school and university undergraduate data sets, along with overall group-level TOEFL ITP test score patterns for these two categories of test takers. Subsequently, we explore finer-grained analyses according to distinct groups of test takers, specifically focusing on those who reported Japanese as their first language, and then considering universities with the Super Global University classification.

## Findings

### Test-taker demographics and overall test scores

Tables 3 and 4 show basic demographic characteristics of high school and university undergraduate test takers in the current data set, displayed for each year under investigation. Examining the high school students first, several patterns bear emphasis. First, the distribution between male and female test takers remained relatively balanced over the 12 years, although the number of self-identified Other gender students increased noticeably in the final four years.

Table 3. Demographic characteristics of TOEFL ITP high school test takers in Japan

| High School Test Takers | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 | 2020 | 2021 | 2022 | 2023 | 2024 |
| G | Male | 44% | 43% | 47% | 49% | 49% | 53% | 51% | 51% | 53% | 48% | 49% | 46% | 45% |
| | Female | 55% | 57% | 52% | 51% | 51% | 47% | 47% | 46% | 47% | 45% | 46% | 44% | 42% |
| | Other | 01% | 00% | 01% | 00% | 00% | 00% | 02% | 03% | 00% | 07% | 05% | 10% | 13% |
| L | Japanese | 92% | 91% | 89% | 82% | 83% | 79% | 83% | 84% | 90% | 85% | 87% | 80% | 76% |
| | Other | 08% | 09% | 11% | 18% | 17% | 21% | 17% | 16% | 10% | 15% | 13% | 20% | 24% |
| A | Mean | 17.1 | 17.1 | 17.1 | 17.0 | 17.0 | 17.0 | 17.0 | 16.9 | 17.0 | 16.9 | 16.8 | 16.9 | 16.8 |
| | SD | 1.35 | 1.36 | 1.38 | 1.34 | 1.33 | 1.30 | 1.42 | 1.52 | 1.43 | 1.55 | 1.58 | 1.58 | 1.54 |

Note: G = Gender; L = First Language; A = Age

Table 4. Demographic characteristics of TOEFL ITP university undergraduate test takers in Japan

| University Undergraduate Test Takers | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 | 2020 | 2021 | 2022 | 2023 | 2024 |
| G | Male | 42% | 42% | 43% | 44% | 45% | 45% | 43% | 42% | 45% | 50% | 47% | 47% | 47% |
| | Female | 55% | 54% | 53% | 53% | 52% | 50% | 50% | 48% | 48% | 48% | 48% | 49% | 48% |
| | Other | 03% | 04% | 04% | 03% | 03% | 05% | 07% | 10% | 06% | 02% | 05% | 04% | 04% |
| L | Japanese | 66% | 66% | 66% | 69% | 70% | 67% | 67% | 73% | 74% | 80% | 76% | 77% | 77% |
| | Other | 34% | 34% | 34% | 31% | 30% | 33% | 33% | 27% | 26% | 20% | 24% | 23% | 23% |
| A | Mean | 20.3 | 20.5 | 20.4 | 20.5 | 19.7 | 19.6 | 19.6 | 19.6 | 20.1 | 20.3 | 20.3 | 20.4 | 20.3 |
| | SD | 5.84 | 6.07 | 5.88 | 5.79 | 4.26 | 3.91 | 4.07 | 3.66 | 4.55 | 8.78 | 8.94 | 9.00 | 8.27 |

Note: G = Gender; L = First Language; A = Age

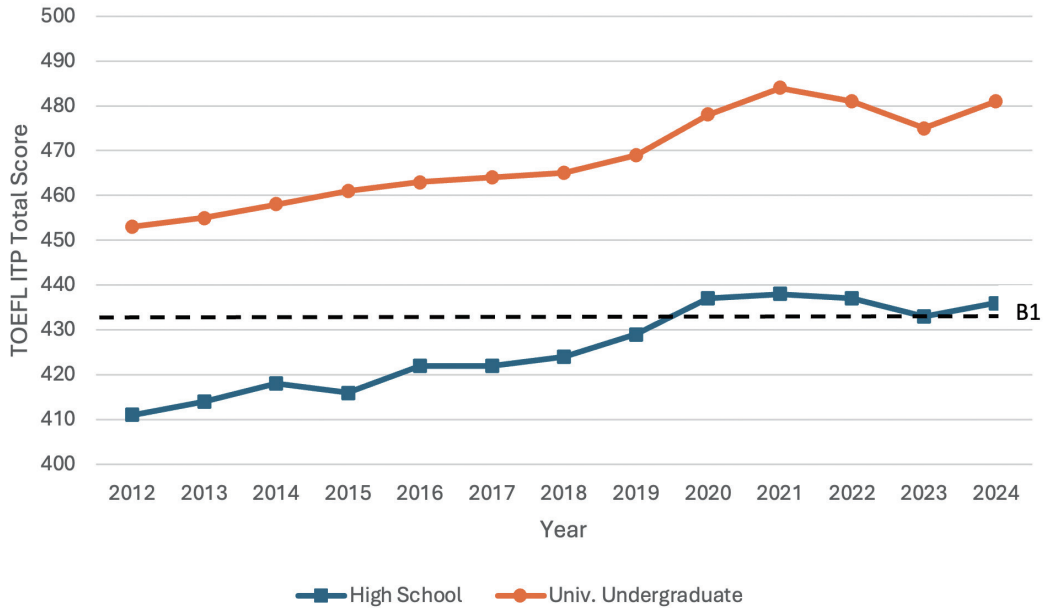
Second, average age of the students remained relatively consistent throughout the period, hovering around 17 years of age (and with low standard deviation values), suggesting that a majority of high school test takers were nearing the end of their studies at the time of testing. Third, and most critical for the current report, the proportion of students with Japanese versus Other first languages (L1) varied somewhat over the years under investigation, ranging from a high of 92% Japanese L1 in 2012 to a low of 76% Japanese L1 in 2024. These differences indicate the possibility that a substantial number of international students (of various kinds) took the TOEFL ITP test at high schools included in the current data set. It is unknown to what extent these students with Other L1s completed any/all of their English language studies within the possibility that a substantial number of international students (of various kinds) took the TOEFL ITP test at high schools included in the current data set. It is unknown to what extent these students with Other L1s completed any/all of their English language studies within

the Japan (versus external) context, while it is considerably more likely that a majority of Japanese L1 students did so. Subsequent analyses therefore address this distinction to provide a more nuanced depiction of English language proficiency among high school students in Japan.

Turning to university undergraduate test takers, the demographic patterns are somewhat distinct. Regarding gender, there was a relative balance between male and female students, with the proportion of male students rising slightly over the twelve-year period, while female students maintained a majority in nearly all years. Average age of students hovered around 20 years, although the standard deviation values suggest a substantial range of student ages within the university undergraduate group for any given year. In terms of first languages, a sizeable proportion of students reported languages other than Japanese, ranging from 34% Other L1 in 2012 to 20% in 2021. Beginning with the COVID-19 pandemic years, the proportion of students with Japanese L1 rose considerably, ranging from 73% to 80% of the test-taking population in the current sample.

An initial analysis of the full data sets for each of the high school and university undergraduate test-taker groups provides a general window into English proficiency levels and differences over the time period of interest. Figure 1 shows average annual TOEFL ITP total scores for each of these two groups. For high school students, total scores differed by +25 points from 2012 to 2024, with average scores initially in the A2 score range and then surpassing the B1 cut score of 433 in the year 2020, remaining at or above the criterion for the remaining years. For university undergraduate students, scores for all years fell in the B1 range while displaying a +28 point difference from 2012 to 2024. Average total scores for both of these groups of test takers, then, showed a generally positive difference in year over year comparisons, with the greatest differences occurring in the 2015-2016 and 2019-2020 comparisons for the high school students, and in the 2019-2020 and 2020-2021 comparisons for the university undergraduates. Bear in mind that these initial analyses include all test-taker data available in each of the two categories. In order to further refine interpretations about the English proficiency of high school and university undergraduate test takers in Japan, the following sections probe deeper into several key grouping categories within these data.

Figure 1. Average annual TOEFL ITP total scores for high school and university undergraduate test takers in Japan



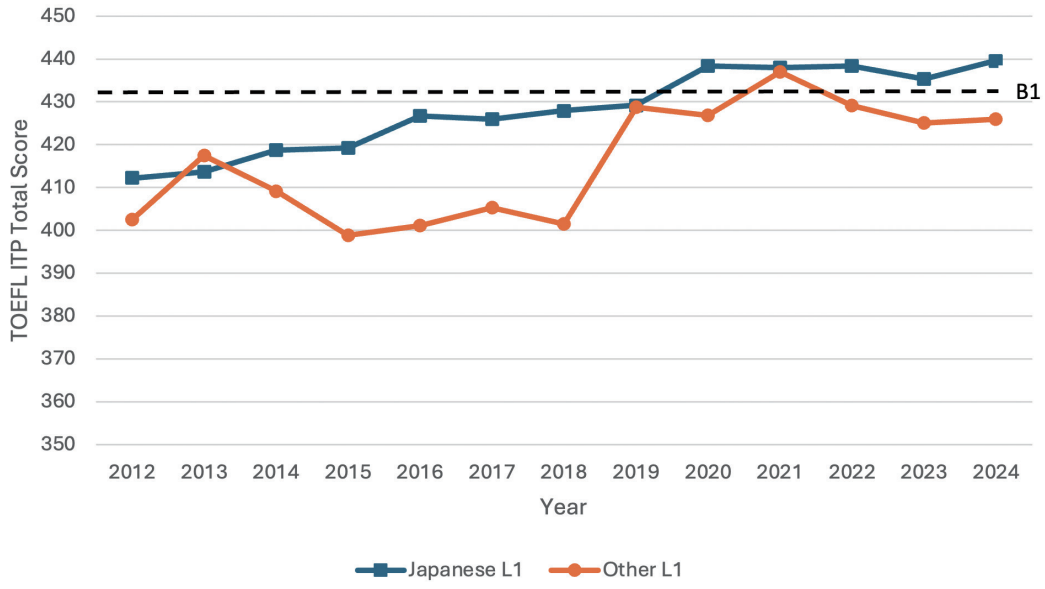
### High School Students

An initial set of group-based analyses focused on the TOEFL ITP test scores of high school students in Japan. Of particular interest were score differences over the period from 2012 to 2024, for total and skill section scores, and specifically the test performances by students self-identifying as having Japanese as their first language. Note that sample sizes for Japanese L1 high school students exceeded 10,000 data points for all years observed; sample sizes for Other L1 high school students were somewhat smaller and should be interpreted with caution.

Figure 2 displays the average annual TOEFL ITP total scores for the two groups of high school test takers: students with Japanese as their first language versus students with a first language other than Japanese. It is noteworthy that the average total scores were generally higher for the Japanese L1 students, in some cases substantially higher (e.g., 27-point difference in 2018). Over the time period examined, both groups of students showed positive score differences, with Japanese L1 students scoring 28 points higher in 2024 compared with 2012, and Other L1 students scoring 23 points higher over the same time period. In addition, from the year 2020, Japanese L1 students' average total scores fell above the CEFR B1 cut score (433 points), while Other L1 students' average scores reached that criterion only for the year 2021 but remained below the B1 cut score for all other observed years.

Beginning in 2014, ETS included score percentile equivalents in the annual test score and data analyses, allowing for a standardized basis for interpreting observed scores in comparison to the global TOEFL ITP test-taking population. While TOEFL ITP test scores are generally quite stable over time, there are slight variations in overall score distributions among test takers in any given year. The calculation of percentile scores allows for a consistent comparison of the average scores for Japanese test takers with the full population of test takers worldwide within a given year. Over the period examined, annual score percentile ranks for high school students in Japan with Other L1s were seven percentile points higher in 2024 (23rd percentile) compared with 2014 (16th percentile).[19,20] For the same period, annual score percentile ranks for Japanese L1 students were ten percentile points higher in 2024 (29th percentile) compared with 2014 (19th percentile). Another way of interpreting these findings is that, in 2024, Japanese L1 high school students in this sample scored higher on the TOEFL ITP than 29 percent of other test takers worldwide. Note that the largest between-year differences in total scores for Japanese L1 students happened in 2015-2016 and 2019-2020.

**Figure 2. Average annual TOEFL ITP total scores for high school students in Japan with Japanese L1 versus Other L1**



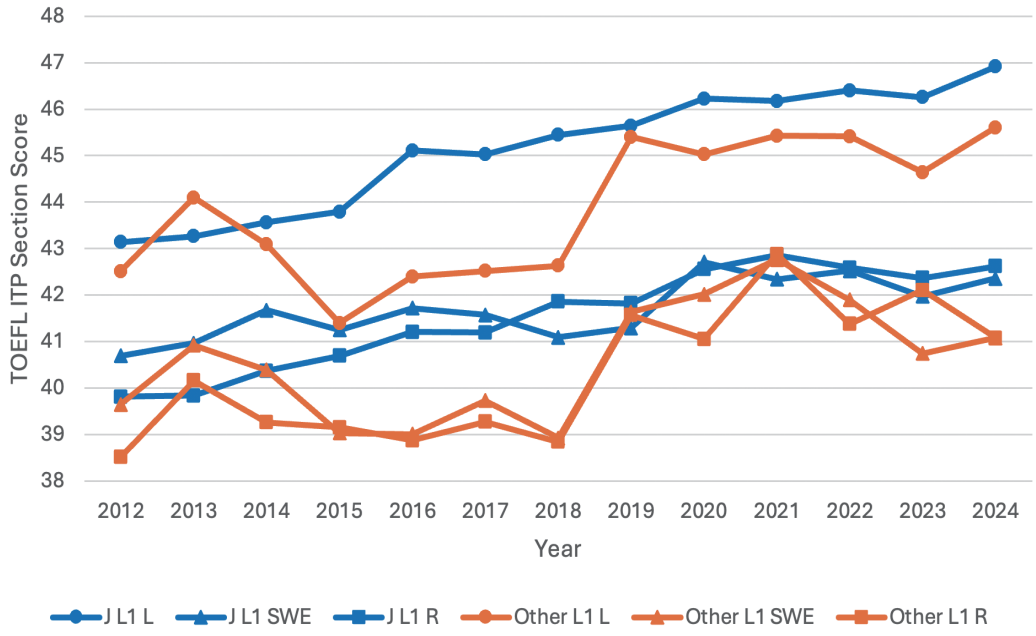
Turning to the TOEFL ITP skill section scores, Figure 3 displays average annual scores for the two groups of students on Listening, Structure and Written Expression, and Reading over the twelve year time period (note the distinct score scale for section scores versus total test scores). In general, Japanese L1 students showed gradual positive score differences across the 12-year period for the three skill sections, while Other L1 students' average scores differed in both positive and negative directions from year to year, albeit with an overall positive difference over the entire time period.

Both groups of students scored highest and demonstrated the greatest difference over time for the Listening section. Average Listening section scores for students reporting Other L1s differed by +3.10 points from 2012 to 2024, although they did not surpass the cut score for CEFR B1 in any observed year. By comparison, average Listening section scores for students reporting Japanese L1 differed by +3.77 points from 2012 to 2024, and they fell above 46 points from 2020 – 2024, thereby achieving the CEFR B1 criterion for those years. Note that the largest between-year difference in Listening scores for Japanese L1 students happened between 2015-2016.

Section score differences for Structure and Written Expression were more modest for both groups of students from 2012 to 2024, +1.44 points for Other L1 students and +1.67 points for Japanese L1 students, with Japanese L1 students generally scoring one to three points higher than Other L1 students in most years observed. Note that average annual scores for neither group surpassed the CEFR B1 criterion of 43 points for Structure and Written Expression in any year observed.

Score differences for the Reading section were more substantial for both groups of students, though not of the same magnitude as for the Listening section. Japanese L1 students' average annual Reading section scores differed by +2.81 points from 2012 to 2024, while Other L1 students' Reading scores differed by +2.56 for the same time period. Regarding the CEFR B1 cut score of 41 points, Japanese L1 students' average scores surpassed and remained above the criterion from 2016 through 2024, whereas Other L1 students' average scores did so from 2019 through 2024. Note that the largest between-year differences in Reading scores for Japanese L1 students happened in 2015-2016 and 2019-2020.

**Figure 3. Average annual TOEFL ITP section scores for high school students in Japan with Japanese L1 versus Other L1**



**University Undergraduate Students**

A next set of group-based analyses focused on the TOEFL ITP test scores of university undergraduate students in Japan. Of particular interest were score differences over the period from 2012 to 2024, for total and skill section scores, and specifically the test performances by students self-identifying as having Japanese as their first language. Note that sample sizes exceeded 10,000 for all groups of students within each year of the time period examined.

Figure 4. Average annual TOEFL ITP total scores for university undergraduate students in Japan with Japanese L1 versus Other L1

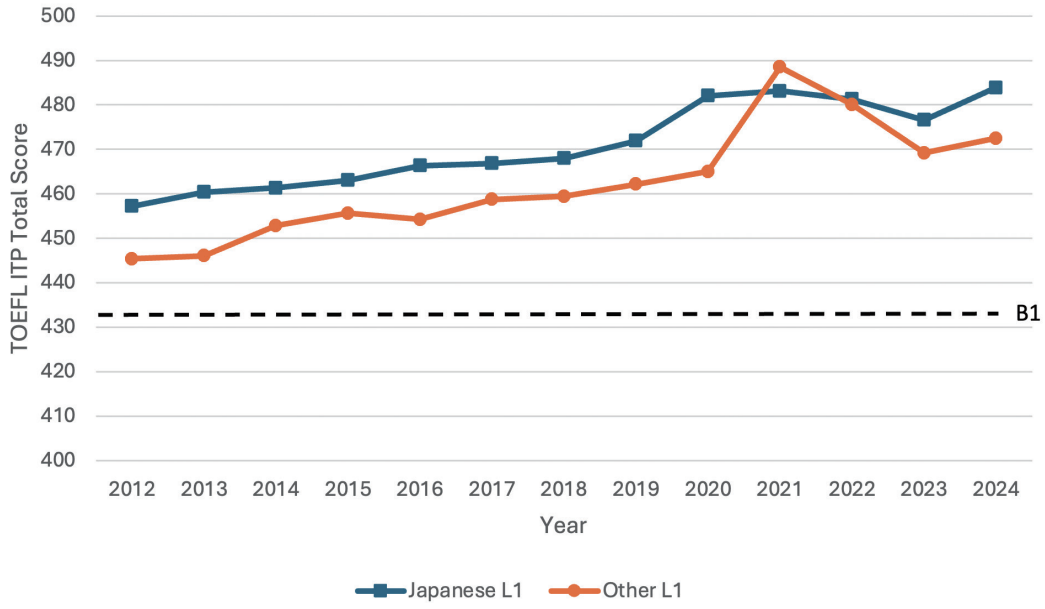



Figure 4 displays the average annual TOEFL ITP total scores for two groups of university undergraduate test takers: students with Japanese as their first language versus students with a first language other than Japanese. It is noteworthy that the average total scores were generally higher for the Japanese L1 students, on the order of approximately 10 score points, with the exception of years 2021 and 2022.

Over the time period examined, both groups of students also showed equivalent and positive score differences, with both scoring 27 points higher in 2024 compared with 2012. In addition, average scores for both groups of students fell above the CEFR B1 cut score (433 points). Over the period for which percentile scores were available, annual score percentile ranks for university students with Other L1s in Japan were six percentile points higher in 2024 (42nd percentile) compared with 2014 (36th percentile). For the same period, annual score percentile ranks for Japanese L1 university students in Japan were nine percentile points higher in 2024 (49th percentile) compared with 2014 (40th percentile). Another way of interpreting these findings is that, in 2024, Japanese L1 university undergraduate students in this sample scored higher on the TOEFL ITP than 49 percent of other test takers worldwide. Note that the largest between-year differences in total scores for Japanese L1 students happened in 2019-2020.

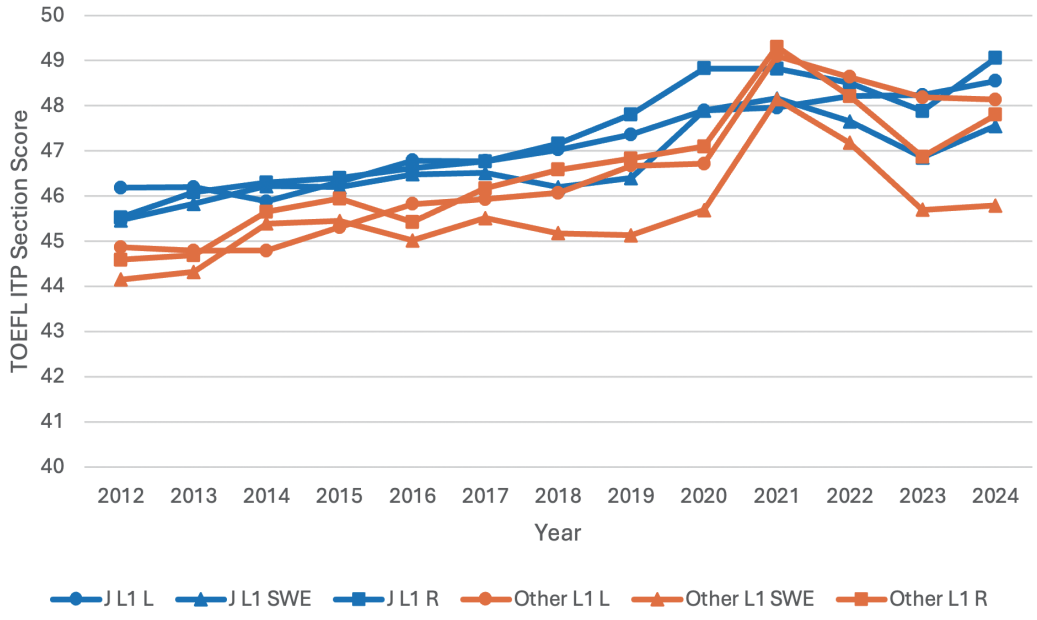
Turning to the TOEFL ITP skill section scores, Figure 5 displays average annual scores for the two groups of university students on Listening, Structure and Written Expression, and Reading over the twelve year time period (note the distinct score scale for section scores versus total test scores). In general, Japanese L1 students showed gradual positive score differences across the 12-year period for the three skill

sections, while Other L1 students' average scores differed in both positive and negative directions, with distinctly high scores in 2021 followed by lower scores through 2024.

Average Listening section scores for students reporting Other L1s differed by +3.26 points from 2012 to 2024, falling beneath the cut score for CEFR B1 from 2012-2018 and then above that criterion for the remaining years. Average Listening section scores for students reporting Japanese L1 differed by +2.37 points from 2012 to 2024, and they fell above 46 points for all years except for 2014, thereby achieving the CEFR B1 criterion for nearly the entire period. Note that the largest between-year difference in Listening scores for Japanese L1 students happened between 2019-2020.

Section score differences for Structure and Written Expression from 2012 to 2024 were +1.64 points for Other L1 students and +2.03 points for Japanese L1 students, with Japanese L1 students generally scoring around one point higher than Other L1 students, with the exception of year 2021, when scores were nearly identical. Note that average annual scores for both groups surpassed the CEFR B1 criterion of 43 points for Structure and Written Expression in all years observed.

**Figure 5. Average annual TOEFL ITP section scores for university undergraduate students in Japan with Japanese L1 versus Other L1**



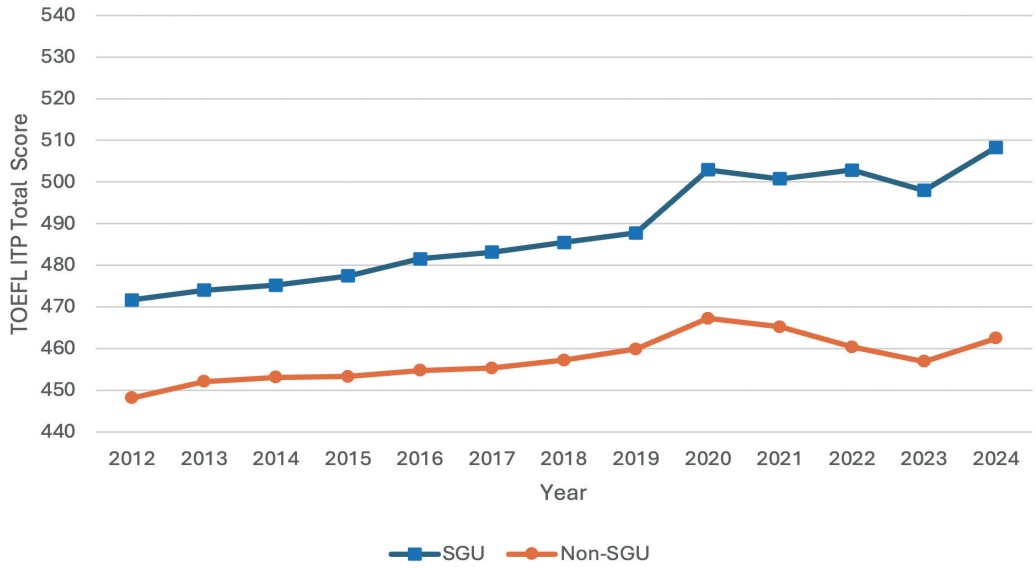

Score differences for the Reading section were similarly substantial for both groups of students. Japanese L1 students' average annual Reading section scores differed by +3.54 points from 2012 to 2024, while Other L1 students' Reading scores differed by +3.21 for the same time period. Japanese L1 students generally scored around one point higher the Other L1 students on the Reading section, with the exception of year 2021, when Other L1 students' scores slightly surpassed those of Japanese L1 students. Average Reading section scores for both groups of students surpassed the CEFR B1 cut score of 41 points in all years observed. Note that the largest between-year differences in Reading scores for Japanese L1 students happened in 2019-2020 and 2023-2024.

### Super Global Universities

A final set of analyses considered score differences over time for the set of universities classified as Super Global (officially "Top Global"). The objective was to detect whether test takers at these universities (SGU) exhibited differing levels of English proficiency compared with test takers at universities not classified as Super Global (non-SGU), and to examine the magnitude of score difference over time within each of these classifications. Here, the analyses focused only on those test takers who self-reported Japanese as their first language, in order to remove any variability in average scores introduced by changing proportions of international student test takers at any given institution and in any given year. Note that sample sizes exceeded 10,000 for all groups examined within all years of the time period.

Figure 6 displays the average TOEFL ITP total scores for Japanese L1 university test takers at SGU and non-SGU universities, for the period of 2012 through 2024. The overall score differences between test takers at the two classifications are apparent, ranging from 22 to 46 score points in favor of SGU universities for within-year comparisons. It is also apparent that the magnitude of difference in total scores became greater over time. Compared with 2012, by 2024 average total scores for test takers at SGU universities were 36 points higher, whereas scores for test takers at non-SGU universities were 15 points higher over the same time period. Note that average total scores for students at both classifications were higher than the CEFR B1 cut score.

**Figure 6. Average TOEFL ITP total scores for Japanese L1 undergraduate test takers at Super Global and other universities**

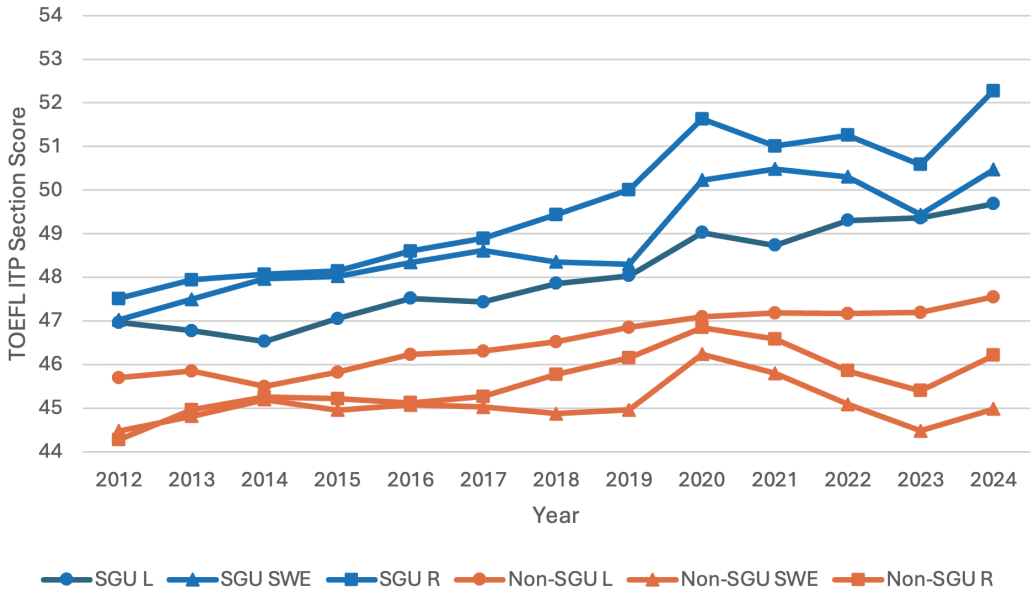



Beginning in 2014, examination of percentile score equivalencies provides another perspective on these total score patterns. For Japanese L1 test takers at non-SGU universities, average total TOEFL ITP test scores in 2014 fell at the 36th percentile rank, and they increased a modest two percentile points to 38th by 2024. The difference over time for Japanese L1 test takers at SGU universities was considerably larger, from the 47th percentile in 2014 to the 61st percentile by 2024. To restate this observation, by 2024, the average TOEFL ITP total scores for Japanese L1 test takers at SGU universities in Japan were higher than the scores of 61 percent of the total population of TOEFL ITP test takers worldwide.

It is also of interest to examine the TOEFL ITP section scores for the same groups of test takers. Figure 7 displays the average TOEFL ITP section scores (Listening, Structure and Written Expression, Reading) for

Japanese L1 undergraduate test takers at SGU and non-SGU universities. Once again, the overall difference between test takers in the two categories is clear, with SGU students scoring anywhere from one to six points higher than non-SGU students on a given skill in a given year (note that the score scale for skill sections is different from the total test score scale). The magnitude of score difference over the time period examined also reveals a notable advantage to the SGU students. For Listening, while the scores for non-SGU students were 1.86 points higher in 2024 compared to 2012, the Listening scores for SGU students were 2.67 points higher over the same time period. For Structure and Written Expression, SGU students scored 3.47 points higher in 2024 compared to 2012, while non-SGU students scored 0.50 points higher for the same time comparison. Lastly, for Reading, SGU students showed the largest difference, 4.77 points, from 2012 to 2024, and non-SGU students also showed the largest difference, 1.94 points, albeit considerably smaller in comparison with the difference score of the SGU students.

Figure 7. Average TOEFL ITP section scores for Japanese L1 undergraduate test takers at Super Global and other universities




Lastly, it is noteworthy that the score profiles differed between test takers at the two classifications. While SGU students consistently scored higher than non-SGU students on all three sections of the test, their scores were highest for the Reading section and lowest for the Listening section. By contrast, students at non-SGU universities showed consistently highest scores on Listening and lowest scores on Structure and Written Expression. Note that the average scores for students at both classifications on all three skill sections were above the CEFR B1 cut score for all observations, with the exception of Listening scores for non-SGU students in the years 2012-2015.

## Discussion and limitations

There are, of course, limitations on what can be interpreted from these analyses. Although the TOEFL ITP test provides a thorough assessment of communicative language ability for listening and reading skills, the productive skills of speaking and writing were not tested in the data analyzed for this report. It remains uncertain to what extent the patterns observed here, and especially the large positive differences for listening and reading proficiency among Japanese students, might be indicative of similar patterns in speaking or writing proficiency. It is also the case that the high school and university undergraduate students whose data were analyzed may not be reflective of the broader population of Japa-

nese learners of English at these educational levels. Indeed, many of the institutions that administered TOEFL ITP during this time period can be considered relatively or highly selective public and private high schools and universities with academically high-achieving student populations. The extent to which the test scores of students at these institutions may represent the English proficiency of broader student populations in Japan remains uncertain. Lastly, the data analyzed here consisted of varying groups of institutions and students from year to year; while many institutions administered TOEFL ITP repeatedly throughout the time period, others contributed data for only some of the years analyzed. Caution should be taken in interpreting the patterns as clearly indicative of change or development in English proficiency, given the reality that the populations of students may have differed somewhat from year to year.

Notwithstanding these limitations on interpretation, the findings offer interesting insights into the English proficiency of large numbers of Japanese high school and university students during a period of intensive educational policy initiatives and reform efforts. Japanese L1 high school students showed substantial positive differences in TOEFL ITP total and section scores over the twelve-year period, with the largest differences exhibited for the listening skill. Average scores for listening and reading, as well as total scores, surpassed the cut-scores for the CEFR B1 level during the time period examined, suggesting the achievement of an important threshold of English proficiency for high school students (B1 is the initial level considered to indicate an independent user of the language). Test scores for this group showed the biggest between year differences in 2015-2016 and 2019-2020, these changes suggestive of potential educational responses to the launch of major policy initiatives, including the 2013 "Implementation Plan for English Education Reform Corresponding to Globalization" and the planned adoption of external English proficiency tests for admissions exams by 2020.

For university undergraduate students, substantial positive score differences in TOEFL ITP total and section scores were also observed among Japanese L1 learners, with average scores falling in the middle of the CEFR B1 range. The largest score differences were observed for reading and listening skills, and Japanese L1 students consistently outperformed students reporting other first languages at the same institutions in Japan. Compared to all TOEFL ITP test takers worldwide, this large sample of Japanese L1 university undergraduate students exhibited a difference of +9 percentile points from 2012-2024, landing at the 49th percentile by the end of the time period.

Of particular note in the TOEFL ITP test score patterns of university undergraduate students were the very substantial score differences observed for students at Super Global Universities. SGU students exhibited the largest score differences during the time period, especially for the reading skill and total scores, and they clearly outperformed students at non-SGU universities in the data set, broadening their superior test score difference by 46 total score points in 2024. SGU students also scored at the 61st percentile of TOETL ITP test takers worldwide by 2024, 14 percentile points higher than their average scores in 2012, a remarkable iteration. The apparent positive trajectory in English proficiency scores was consistent over the years from the launch of the "Top Global Universities" initiative in 2014, with the largest between-year differences happening in 2019-2020 and 2023-2024.

## Conclusion

While the patterns in the current report should not be over-interpreted, it is clear that large groups of Japanese high school and university students exhibited substantial positive iterations in their English proficiency over the 2012-2024 time period. That finding is, at a minimum, consistent with the intended trajectory of language learning outcomes targeted by multiple governmental policy initiatives and educational reform efforts. It is truly fortuitous that hundreds of institutions administered a robust measure of English proficiency, the TOEFL ITP test, on an annual basis during this time. The test score

data accumulated from over a million test takers offers a particularly valuable window into the status of Japanese students' English proficiency, an important educational goal and a key competency of interest for social, economic, and personal development.

Looking to the future, we hope to encourage the ongoing and regular administration of robust measures of English proficiency at high schools and universities in Japan, as a means of gauging progress in this critical academic, professional, and life skill. The addition of productive skills assessments (e.g., the TOEFL ITP Speaking test) would also provide a useful expansion to analyses such as those presented here. We also encourage institutional test users and other stakeholders to reflect on the patterns displayed in this report as a kind of framework for situating and interpreting the English proficiency levels of their own test takers. Especially for Japanese high school and university undergraduate students, these data show the average levels of proficiency achieved by a sample of test takers—and the extent to which these differed over a meaningful time range—providing a valuable basis for comparison.

## References


1. The TOEFL ITP® Assessment Series. https://www.ets.org/toefl/itp/about.html
2. Erikawa, H. (2008). Nihonjin ha Eigo wo Dou Manandekitaka "How have Japanese people learned English?" 日本人は英語をどう学んできたか". Kenkyusha.
3. Ministry of Education, Culture, Sports, Science and Technology (MEXT). (n.d.). Japan's modern educational system. https://www.mext.go.jp/b_menu/hakusho/html/others/detail/1317220.htm
4. Ministry of Education, Culture, Sports, Science and Technology (MEXT). (2003). Action plan to cultivate "Japanese with English abilities".
5. Ministry of Education, Culture, Sports, Science and Technology (MEXT). (2017). Course of study for elementary schools (2017 notification).
6. Ministry of Education, Culture, Sports, Science and Technology (MEXT). (2025).Current status, issues, and points for consideration in foreign language education. "外国語に関する 現状・課題と検討事項" https://www.mext.go.jp/content/20250924-mxt_kyoiku01-000044934_002.pdf
7. Ministry of Education, Culture, Sports, Science and Technology (MEXT). (2019). Initiatives for elementary school foreign language education toward full implementation of the new Course of Study.
8. Ministry of Education, Culture, Sports, Science and Technology (MEXT). (2013). English education reform plan in response to globalization.
9. Ministry of Education, Culture, Sports, Science and Technology (MEXT). (2018). The third basic plan for the promotion of education.
10. Ministry of Education, Culture, Sports, Science and Technology (MEXT). (2017). Course of study for elementary and secondary schools in Japan: Revised guidelines and commentaries (2017–2019).
11. Cabinet Secretariat, Government of Japan. (2023). The 6th meeting of the Council for the Creation of Future Education: Distributed materials.
12. Ministry of Education, Culture, Sports, Science and Technology (MEXT). (2023). The fourth basic plan for the promotion of education.
13. Ministry of Education, Culture, Sports, Science and Technology (MEXT). (2019). GIGA School Program in Japan.
14. Ministry of Education, Culture, Sports, Science and Technology (MEXT). (2013). The second basic plan for the promotion of education.
15. Ministry of Education, Culture, Sports, Science and Technology (MEXT). (n.d.). Tobitate! (Leap for Tomorrow) Study Abroad JAPAN initiative.
16. Ministry of Internal Affairs and Communications. (2017). Policy evaluation on the promotion of global

human resource development: Recommendations based on evaluation results.

17. Council of Europe (2001). Common European Framework of Reference for Languages. Author. https://www.coe.int/en/web/common-european-framework-reference-languages/cefr-and-its-language-versions

18. Ministry of Education, Culture, Sports, Science and Technology (MEXT). (n.d.). Support for the creation of Super Global Universities. https://www.mext.go.jp/a_menu/koutou/kaikaku/sekaitenkai/1360288.htm

19. ETS. (2015). Test and score summary for the TOEFL ITP® test: January – December 2014 test data. Author.

20. ETS. (2025). Test and score summary for the TOEFL ITP® test: January – December 2024 test data. Author.