

広島大学学術情報リポジトリ  
Hiroshima University Institutional Repository

Title	Exploring Directions for Improving General English Writing Courses through the TOEIC® Writing Test
Author(s)	AMANO, Shuichi; HOUGHAM, Daniel; HIGA, Marshall; NAKAGAWA, Atsushi; TAKAHASHI, Yuka; FRASER, Simon
Citation	Hiroshima Studies in Language and Language Education , 27 : 1 - 15
Issue Date	2024-03-01
DOI	
Self DOI	<a href="https://doi.org/10.15027/54909">10.15027/54909</a>
URL	<a href="https://doi.org/10.15027/54909">https://doi.org/10.15027/54909</a>
Right	Copyright (c) 2024 広島大学外国語教育研究センター
Relation	



# Exploring Directions for Improving General English Writing Courses through the TOEIC® Writing Test

Shuichi AMANO  
Daniel HOUGHAM  
Marshall HIGA  
Atsushi NAKAGAWA

Institute for Foreign Language Research and Education  
Hiroshima University  
Yuka TAKAHASHI  
Global Education Center  
Tokyo Kasei University  
Simon FRASER  
Institute for Foreign Language Research and Education  
Hiroshima University

In this study, we provide an analysis of the outcomes of writing instruction within liberal arts English courses by employing the TOEIC® Speaking & Writing Tests Institutional Program (henceforth referred to as the “TOEIC® Speaking and Writing Tests IP”). Drawing from these findings, we evaluate the status of these courses and consider potential challenges that lie ahead. Hiroshima University’s Institute for Foreign Language Research and Education (FLaRE) has consistently utilized the TOEIC® Speaking & Writing Tests IP, administering it since 2016, with a notable interruption during the COVID-19 pandemic in the academic years 2020 and 2021. In its formative two years, this test was used primarily to gauge the outcomes of an English language program named Hiroshima Special Program for English Communication (HiSPEC). Following the cessation of the HiSPEC program due to broader curriculum reforms within the university, the test has been employed continuously since 2018. Its renewed purpose has been to assess the outcomes of writing instruction for the regular class student cohort, who are required to achieve eight credits (modified to six credits from 2020) from liberal arts English courses (see Amano et al., 2021 and Morita, 2020 for more detail).

Proficiency tests play a pivotal role as essential tools for realizing effective management in university English courses under the liberal arts education program. These English courses within the liberal arts framework are vital in achieving the aim of maintaining or enhancing students’ English proficiency. Proficiency tests serve as one means of objectively assessing students’ English abilities over the courses. Based on the test results, instructors and curriculum coordinators can discern student proficiency levels, strengths of the course/curriculum, and areas needing improvement. Consequently, appropriate measures can be designed in response. If the overall performance is found lacking, there might be a need to reconsider the choice of teaching materials or reevaluate class assignment methods. Alternatively, instructors and curriculum coordinators might need to customize the curriculum to facilitate easier comprehension for students or offer tailored support to specific individuals. Conversely, if the aggregate performance is good,

it suggests the course is functioning effectively. Thus, we should be cautious before introducing alterations that might threaten previously established successes. Additionally, course reforms typically demand considerable effort and financial resources. It might be administratively and financially imprudent to make hasty modifications to a course perceived as successful. In such contexts, the TOEIC® Speaking and Writing Tests IP is important for our liberal arts English courses. Note that both the Speaking and Writing tests were administered in 2018 and 2019, but only the Writing test was administered in 2022, with the Speaking test being dropped as a precautionary measure against the potential spread of COVID-19. For this reason, the sections that follow will focus solely on the Writing test administration. This paper elucidates the results of the TOEIC® Writing Test IP at Hiroshima University’s School of Education from 2022, marking its resumption following the two-year gap caused by the COVID-19 pandemic, and offers a comparative analysis with data from 2018 and 2019.

## **BACKGROUND**

The TOEIC® Writing Test evaluates the writing skills required for efficient communication within a global work environment. The impetus for our institution’s integration of the TOEIC® Writing Test IP stemmed from Hiroshima University’s recognition as a member of the “Top Global University Initiative,” a program aiming to enhance international competitiveness, championed by the Ministry of Education, Culture, Sports, Science and Technology.<sup>1</sup> As part of this program, our institute’s first initiative was the formulation of the HiSPEC program. This program, crafted as a small-group English language proficiency enhancement module, was designed to nurture students with the skills needed to succeed globally. To appraise the outcomes and ascertain the efficacy of the HiSPEC program, the TOEIC® Writing Test IP was employed as an assessment instrument (as delineated in Uenishi et al., 2017; Uenishi et al., 2018; Uenishi et al., 2019).<sup>2</sup> However, merely two years after its inception, the HiSPEC program faced an early end owing to an overarching university-wide curriculum change (comprehensively detailed in Morita, 2020). In the wake of this, the TOEIC® Writing Test IP was used in a new way: it was systematically administered in selected schools at our university. The primary objective was to gauge the effectiveness of regular classroom writing instruction. Each year, a selection was made from approximately half of the first-year student cohort.

In the academic year of 2018, the test was conducted, encompassing eight classes belonging to the Schools of Science, Applied Biological Science, and Informatics and Data Science. Each class had an enrollment ranging from 26 to 28 students. Two tests were administered on two occasions, in the months of October and January, marking the beginning and end of the writing course. The Communication IIA and IIB courses, offered during this period, were tailored to foster writing proficiency, thereby facilitating students in the acquisition of writing skills efficiently. These classes met weekly for sessions spanning a duration of 90 minutes each. Among the cohort of 172 students enrolled in these courses, a subset comprising 128 individuals undertook both the October and January tests, and responded to the accompanying questionnaire surveys. The outcomes, presented in Table 1, were described in accordance with the summary presented by Uenishi et al. (2020),<sup>3</sup> ultimately revealing a discernable average score differential of 15.4 points. This was comparable to the HiSPEC students in terms of score growth.

**TABLE 1. Scores of 2018 Students in the TOEIC® Writing Test**

Month implemented	October	January
Mean	118.0	133.4
Standard deviation	20.4	18.4
Maximum	160	160
Minimum	60	70

In the academic year 2019, students from the School of Engineering took the TOEIC® Writing Test IP. They were guided by seven instructors across ten classes, with class sizes varying from 23 to 30 students. The writing course ran from October 2019 to early February 2020, including the holiday season. Tests were administered in mid-April and late January.<sup>4</sup> Although 275 students initially enrolled in the target classes, ultimately, 166 students completed both tests. The results, presented in Table 2, are based on the summary of Amano et al. (2021). In the January test, there was an increase of 5.30 points in the average score. It is important to note that this improvement, while commendable, was somewhat lower than students from HiSPEC programs and regular classes in the 2018 test administration. Additionally, 15% of students maintained their scores, 52% experienced an increase, and 33% showed a decline in their scores.

**TABLE 2. Scores of 2019 Students in the TOEIC® Writing Test**

Month implemented	April	January	Difference
Mean	117.41	122.71	5.30
Standard deviation	19.57	20.67	22.45
Maximum	160	160	60
Minimum	70	20	-80
Median	110	120	10

In summary, our institute’s utilization of the TOEIC® Writing Test IP evolved from its initial application within the HiSPEC program, which began through our university’s involvement in the Top Global University Initiative. After the HiSPEC program was discontinued due to university-wide curricular changes, the TOEIC® Writing Test IP was repurposed to evaluate regular classroom writing instruction across select schools. The results from the academic years 2018 and 2019 provided valuable insights into the actual writing skills of students in the regular class.

After a two-year hiatus due to the COVID-19 pandemic, we turned our attention in the 2022 academic year to the next phase of our assessment project. In that academic year, we extended the examination to a broader student population, further expanding our understanding of writing proficiency development. Specifically, we focused on students from the School of Education, some of whom were English language education majors. Our research question for this phase is: To what extent do these students, who encompass a diverse range of academic interests within the School of Education, demonstrate growth in writing skills when exposed to regular classroom writing instruction? However, it should be noted that our objective is not to derive findings that can be generalized, but to discover insights that can contribute to the improvement of our liberal arts English program. This inquiry will be explored in detail in the following section.

## **PROCEDURE**

### **Data Collection**

In the 2022 academic year, students from the School of Education were selected to take the TOEIC® Writing Test IP. With the participation of these students, the rotational testing of students from all schools located at the Higashi-Hiroshima Campus came to an end. Four FLARE members were responsible for teaching six classes, each targeting first-year students.<sup>5</sup> The class sizes ranged from 20 to 26 students. The writing courses were offered over the third and fourth terms, spanning from October 2022 to February 2023, with each instructor conducting a 90-minute lesson once a week. The test was administered twice: in mid-October of 2022 and in mid-January of 2023. After the October test, students received approximately ten writing lessons before the January test. Consistent with past practice, while teachers were mindful of the TOEIC® Writing Test preparation, the content and materials of the lessons were determined at the instructor's discretion. Furthermore, it was agreed that the test results would be incorporated as part of the evaluation. Additionally, until the previous administration, the test was taken using dedicated laptop computers, but from this administration, it was taken online using test takers' own computers. However, there were no changes in the question format or the given test time.

### **Data Analysis**

In this study, we not only used the TOEIC® Writing Test IP but also incorporated a survey based on the Can-Do Statement list developed by the Institute for International Business Communication, the administrative organization for the TOEIC® program in Japan. The statements on the 5-point Likert scale used in the survey are presented in Table 4, along with the results. Employing both the standardized TOEIC® Writing Test IP and the more practical Can-Do Statements for writing skills<sup>6</sup> offers two primary advantages. First, it allows for a multifaceted evaluation of students' writing skills. While standardized tests assess students' skills based on objective criteria, Can-Do Statements capture subjective shifts in a student's actual language skills in everyday life and work contexts. Combining these two approaches enables a more comprehensive assessment of a student's capabilities. Second, the use of Can-Do Statements offers students an opportunity to reflect on and evaluate their perceived language proficiency. This, in turn, enables FLARE teachers to discern any alignment or discrepancy between a student's self-perception and the assessment from the testing organization, providing clear directions for further guidance and points of adjustment. Anticipating these benefits, we decided to detail our approach and findings from the Can-Do Statements survey in this article.

### **Student Participation**

Although the total number of students enrolled across the target classes was 132, due to students who could not take either the October or January tests and those who did not consent to provide their scores for this report, only 89 students who took both tests were considered for analysis. A major reason for the exclusion of many students was the high number of absences at the October test due to justifiable reasons related to COVID-19, especially among students majoring in English language education. Unfortunately, due to circumstances beyond anyone's control, this adversely affected an objective of this study, which was to gather data from a broad spectrum of English proficiency levels among the students at the School of Education.

## RESULTS

### Test Scores

Table 3 presents the writing test scores of students taking the test in the 2022 academic year. The pre-test score in October was 129.10. When compared with students from the previous two academic years, this average score was more than 10 points higher, suggesting that the pre-instructional skill level was the highest for students in regular classes. After approximately 10 writing lessons from October, their scores improved to 134.61 in the January test. This increase of 5.51 points is comparable to that of students from the 2019 academic year; however, it is somewhat disappointing when compared to the 2018 academic year students, who saw a score improvement of about 15 points. Nevertheless, this score remains the highest post-test result for students in regular classes over the past three academic years.

**TABLE 3. Scores of 2022 Students in the TOEIC® Writing Test (N = 89)**

Academic year 2022	October	January	Difference
Mean	129.10	134.61	5.51
Standard deviation	19.17	19.95	17.12
Maximum	170	180	60
Minimum	80	90	-40
Median	130	140	10

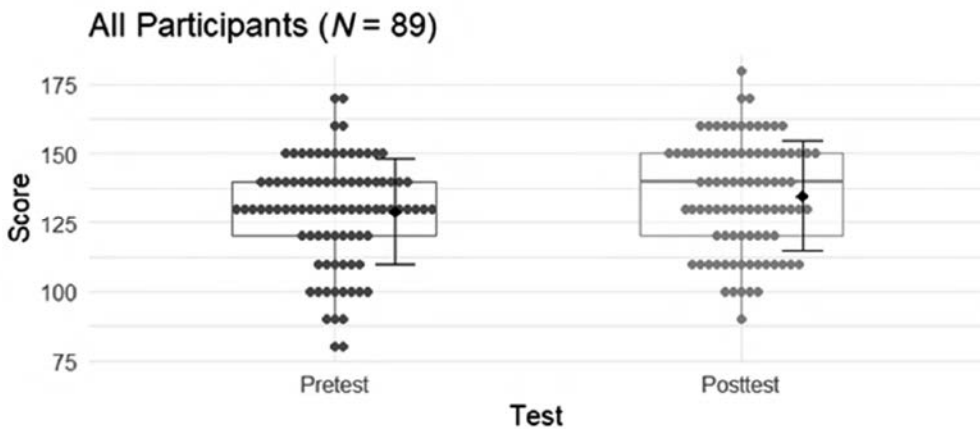
Initially, we conducted the Shapiro-Wilk test to check the normality of our data. The result yielded a p-value of 0.003, suggesting that at a 5% significance level, the data did not follow a normal distribution. Upon closer inspection, we found two outliers – two students who scored 80 points on the pretest. These outliers were not extreme, and the data distribution showed a slight left skew.

Consequently, guided by the recommendations of Larson-Hall (2015) and Wilcox (2012), we conducted a robust paired samples t-test, which allowed us to assess the significance of the difference between the pre-test and post-test scores more accurately. For this analysis, we used the “trimpb” function from the “WRS” package (Wilcox & Schonbrodt, 2017) in R (R Core Team, 2019), applying a 5% trimmed mean. This approach was chosen based on the presence of a non-extreme outlier and slight data skewness, allowing us to retain 95% of our data (see appendix for the R code we used). Although it may seem counterintuitive to discard data from the sample by using a trimmed mean, Wilcox (2001) asserted that, when there are outliers in the data, the trimmed mean gives higher power. Loewen and Plonsky (2016) pointed out that trimming a smaller percentage (e.g., 5% instead of 20%) is appropriate where possible because it allows us to retain more of our data set in the analysis.

The robust t-test with a 5% trimmed mean yielded a p-value of 0.0005 and a confidence interval of [1.852, 8.642]. Given that our confidence interval [1.852, 8.642] does not include zero and the p-value is very low, these results provide strong evidence against the null hypothesis, indicating a significant difference in the paired observations. This difference is estimated to be between 1.852 and 8.642.

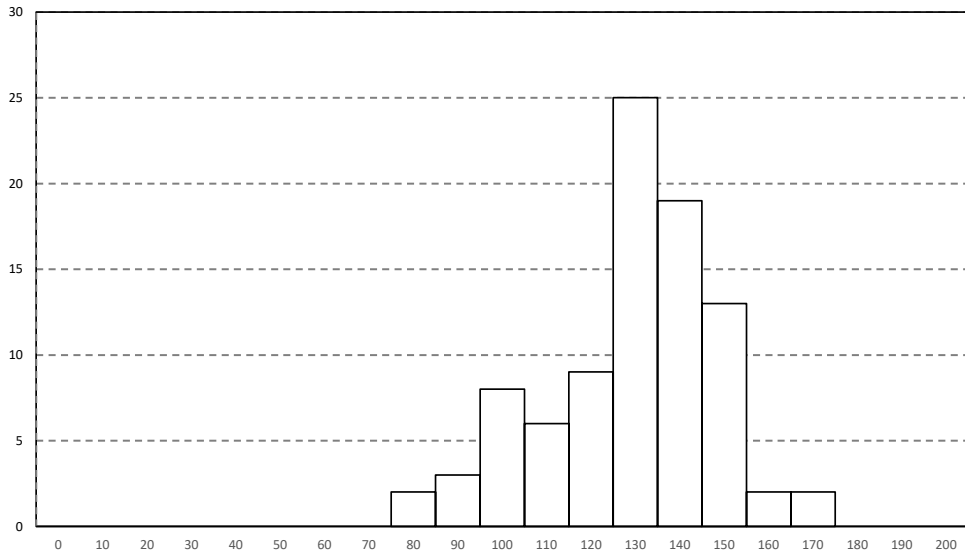
Figure 1 presents a combined box-and-whisker and beeswarm plot illustrating the distribution of scores for pre-tests and post-tests among participants. The box-and-whisker component outlines the interquartile range, with the top and bottom edges of the box marking the third (the lower limit of the top 25% of the data)

and first (the upper limit of the bottom 25% of the data) quartiles, respectively, and the median indicated by the central line. The pre-test scores have a first quartile at 120 points, a median of 130 points, and a third quartile at 140 points. The post-test scores show a first quartile at 120 points, a median of 140 points, and a third quartile at 150 points, indicating a slight increase in median scores and a wider interquartile range. The beeswarm plot adds a layer of detail by displaying each individual score as a point, indicating the density of data around the median. The subtle increase in the post-test median implies a potential instructional effect, and this conclusion is reinforced by statistical analysis already conducted. Integrating the statistical analysis with the visual insights from the beeswarm plot, which reveals the presence of two outliers alongside a wide distribution of scores, is crucial for a thorough interpretation of the instructional impact on participant outcomes.

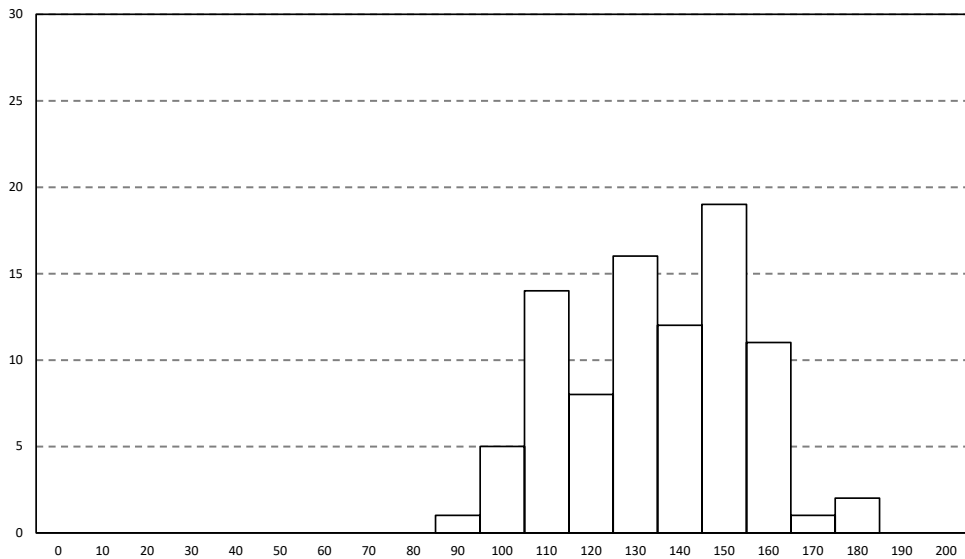


**FIGURE 1. Box-and-Whisker Plot for the Results of the Pre- and Post-Tests (N = 89)**

Figure 2 presents a histogram illustrating the score distribution for the writing pre-test conducted in October. The mode was 130 points, attained by 25 individuals, constituting 28.09% of the sample. Further breakdowns include 2 students (2.25%) scoring 80 points, 3 students (3.37%) scoring 90 points, 8 students (8.99%) scoring 100 points, 6 students (6.74%) scoring 110 points, 9 students (10.11%) scoring 120 points, 19 students (21.35%) scoring 140 points, 13 students (14.61%) scoring 150 points, 2 students (2.25%) scoring 160 points, and 2 students (2.25%) scoring 170 points. From this distribution, it can be observed that 64.04% of students are concentrated within the score range of 130 to 150 points, indicating that they were already achieving relatively high scores prior to receiving writing instruction.



**FIGURE 2. Score Distribution for the 2022 Writing Test in October**



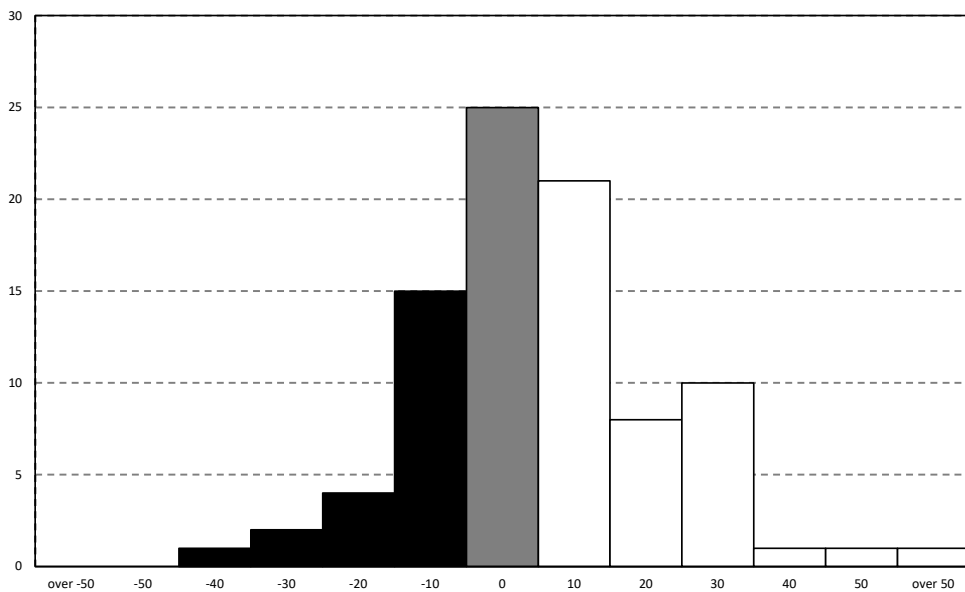
**FIGURE 3. Score Distribution for the 2022 Writing Test in January**

Figure 3 depicts a histogram showing the score distribution for the post-test conducted in January. The mode was 150 points, achieved by 19 individuals, which represents 21.35% of the sample. The remaining scores were as follows: 1 student (1.12%) scored 90 points, 8 students (8.99%) scored 100 points, 14 students (15.73%) scored 110 points, 8 students (8.99%) scored 120 points, 16 students (17.98%) scored 130 points, 12 students (13.48%) scored 140 points, 19 students (21.35%) scored 150 points, 11 students (12.36%)



scored 160 points, 1 student (1.12%) scored 170 points, and 2 students (2.25%) scored 180 points. The histogram reveals positive outcomes from the instruction, as evidenced by an increase in the mode to a higher score compared to the pre-test, an increase from 2 to 11 individuals scoring 160 points, and a decrease from 5 to 1 individual scoring below 100 points.

Figure 4 presents a histogram depicting the distribution of score differences between the writing tests taken in October 2022 and January 2023. The white bars represent students with a positive score difference, the gray bars denote students with no change in scores, and the black bars indicate students with a negative score difference. The number of students with unchanged scores was 25 (28.09%), while 22 students (24.72%) experienced a drop in scores. Conversely, 42 students (47.19%) saw an improvement in scores. Compared to the 2019 academic year (Amano et al., 2021), when the difference between the pre-test and post-test scores was roughly equivalent with about 5 points for this year, there was a significant increase in the percentage of students with unchanged scores. Meanwhile, the proportion of students with decreased scores substantially fell, and there was a slight decline in the percentage of students with improved scores. This suggests that despite the score differences being almost equivalent, there was less variance in student test scores in the 2022 academic year than in 2019. In particular, the decline in the proportion of students with lower scores might hint that the teaching approach in 2022 was potentially more effective than in the previous year. On the other hand, the marginal decrease in students with improved scores may suggest that some advanced students could benefit from additional support or instruction.



**FIGURE 4. Distribution of the Difference in Writing Test Scores between October and January**

### Survey Results

An analysis of the Can-Do Statements survey conducted before and after writing instruction again suggested that the instruction had a positive impact on students' writing skills. As shown in Table 4,

comparing the average response scores from October and January, the scores were approximately 3.14 in October and 3.26 in January, indicating a consistent improvement. Overall, among the 15 statements, 14 showed an increase in response score, while only one statement exhibited a decrease, suggesting the instruction was generally effective. The statement with the most notable improvement was “I can write the invitation for a party addressed to my friends.” On the other hand, the only statement that saw a slight decline was “I can write a letter of introduction for myself to address a stranger.” However, this decline was minimal at -0.01, suggesting that it is effectively negligible. Moreover, upon examining the score distribution for October and January, the standard deviations were approximately 0.60 for October and 0.59 for January. Both pre-instruction and post-instruction surveys showed similar values, implying a uniform distribution of the instructional effects.

**TABLE 4. Student Responses to Can-Do Statements for Writing Skills**

Can-do statements	October	January	Difference
W-1 I can write the list of things to take with me on a trip.	3.08	3.28	0.20
W-2 I can write simple cards such as birthday cards, Christmas cards, New Year’s cards, etc., to my friends	3.17	3.33	0.16
W-3 I can write the directions to my house.	3.03	3.16	0.13
W-4 I can write postcards to friends to let them know about my vacation.	3.25	3.33	0.08
W-5 I can write the reason for being late.	2.80	2.94	0.14
W-6 I can write both upper- and lower-case letters of the alphabet.	4.93	4.97	0.03
W-7 I can write the invitation for a party addressed to my friends.	3.30	3.54	0.24
W-8 I can write my home address.	3.06	3.25	0.20
W-9 I can prepare a report on my school research project.	2.20	2.31	0.11
W-10 I can write my resume (a document showing my background).	2.37	2.43	0.06
W-11 I can write a book report.	2.85	2.99	0.14
W-12 I can write a letter of introduction for myself to address a stranger.	3.37	3.36	-0.01
W-13 I can write a note to my friends letting them know where and when to meet me.	3.34	3.55	0.21
W-14 I can write an essay about my future dreams (e.g., what career I want to pursue).	3.14	3.20	0.06
W-15 I can keep a diary.	3.24	3.29	0.05

This survey, which used the Can-Do Statements list, was conducted with the dual intention of providing students with an opportunity to self-assess their writing skills and capturing subjective shifts in their language skills in everyday life. The results of the survey showed an improvement in students’ writing skills on the Can-Do Statements scale, which aligns with the rise in standardized test scores. This suggests that students’ skills are either stable or improving across both objective and subjective evaluations.

### Future Directions of Writing Courses

The liberal arts English course implemented at our university’s FLARE aims to enhance not only the students’ receptive skills (reading and listening) but also their productive skills (speaking and writing). This study attempts to evaluate the effectiveness of writing instruction by using the TOEIC® Writing Test IP as an indicator. Test scores from 2018 showed significant improvements, rivaling those achieved in the HiSPEC program. In contrast, the 2019 test displayed smaller gains in scores, along with larger individual variances. Meanwhile, the 2022 test exhibited improvements roughly equivalent to those in 2019 but with less score variability and higher initial proficiency levels. Several factors may contribute to these changes, such as

changes in teaching staff, variations in students' majors, and external elements like the COVID-19 pandemic.

From the results, we can identify two critical issues. First, the importance of continually administering standardized tests is evident; the TOEIC® Writing Test IP has consistently played a significant role in measuring the effectiveness of writing instruction over multiple years. Ideally, the test would be administered across all academic departments within the same academic year, but given budgetary constraints, ongoing staggered implementation remains the next best option. Despite the challenges brought on by the pandemic, this continuous administration has enabled school-to-school comparative analysis within the university. Second, there is a need to focus on tailored approaches to relatively advanced students. While average scores have historically plateaued between 130 and 140 points since the HiSPEC program, targeting specific advanced students is necessary to exceed this ceiling. Currently, although proficiency-based classes are utilized, addressing the range of proficiency levels in each class is left primarily to the individual efforts of individual instructors. It remains unclear whether this approach meets the needs of advanced students aiming for high scores. Although realizing these tailored approaches is extremely difficult, a review of customized curricula and class-structuring methods may consequently be necessary.

## **CONCLUSION**

This paper emphasizes the evaluation of the liberal arts English writing course at the Institute for Foreign Language Research and Education at Hiroshima University. Specifically, we focused on the role of the TOEIC® Writing Test IP and insights derived from its results. Amidst the academic disruptions brought about by the COVID-19 pandemic, this test retains its reliability in assessing students' English writing skills. The analysis, which also includes data from 2018 and 2019, provides insight into the current state of students' English language writing skills.

From the 2022 scores, it became evident that students from the School of Education had an average pre-test score in October that was notably higher compared to students from other schools. After receiving approximately ten writing lessons, scores further improved in the January test. Our findings suggest a meaningful impact of the instruction in general writing courses on students' TOEIC® Writing scores, as evidenced by a statistically significant difference in pretest and posttest scores. This is supported by our confidence interval and very low p-value, underscoring the practical significance of our teaching and research. Observing the histogram, we noticed that a majority of the students achieved relatively high scores in the October test, with further improvement evident in the January test. However, given the inherently higher starting scores, the progress seemed slightly underwhelming.

The survey results from Can-Do Statements also suggested that the writing instruction positively impacted students' writing skills. By providing students with an opportunity to self-assess their language abilities, it was possible to capture changes in language skills relevant to daily life. The findings aligned with the rise in standardized test scores. This indicates that students' skills improved on both objective and subjective assessment scales.

However, a challenge identified was that students struggled to surpass the average score barrier that ranges from 130 to 140 points. To breach this ceiling, targeted improvements for advanced students may be necessary. Although skill-based classes are currently in use, it remains unclear if this approach sufficiently caters to students aiming for higher scores. A review of customized curricula and class structuring tailored

to the specific needs of advanced students might potentially be needed.

Furthermore, the significance of administering standardized tests like the TOEIC® Writing Test IP was strongly recognized. Despite budgetary constraints on implementation which led to a rotation across three groups, the challenges posed by the pandemic did not prevent the resumption of collecting comprehensive data on students' current writing skills. The findings serve as a helpful foundation for us in refining material selection, innovating with instructional methods, and re-evaluating class assignment strategies. As we contemplate future educational reforms and curriculum redesign, the continued use of such standardized tests is indispensable in consistently enhancing educational quality and realizing instruction tailored to student needs. Lastly, it is our hope that the insights offered by this paper serve as a valuable foundation for researchers and educators in discussions aimed at designing and refining liberal arts English courses.

## ACKNOWLEDGEMENT

This manuscript was initially proofread for English language clarity using OpenAI's ChatGPT. Subsequently, it underwent further proofreading by all the authors.

## NOTES

- 1) Initiated by the Ministry of Education, Culture, Sports, Science and Technology, the “Top Global University Project” (2014–2023) seeks to enhance the global standing of Japanese higher education through university system reforms and active partnerships with foreign institutions.
- 2) The details of the TOEIC® Writing Test IP can be found on the website (<https://www.iibc-global.org/english/toEIC/test/sw/about/format.html>), and they are described by Amano et al. (2021).
- 3) Although we had hoped to match the descriptive statistics presented in Table 1 with those from later years, the raw data is no longer available, so we rely on the descriptions in the paper and present the data as best we can.
- 4) Ideally, we would have preferred to conduct the writing test in October, consistent with other years, rather than in April. However, owing to budgetary and scheduling constraints, we had to forgo this plan.
- 5) The fifth author worked at the Institute for Foreign Language Research and Education, Hiroshima University, until August 2023 and participated in this research project.
- 6) The Can-Do Statements for writing skills used in this study are consistent with those presented in Kida et al. (2017) and Amano et al. (2023). For the survey, the Japanese version was employed. The English version provided in this paper is not an official translation but is offered for the convenience and better understanding of the reader.

## REFERENCES

- Amano, S., Davies, W., Fraser, S., Lauer, J., Selwood, J., & Takita, F. (2021). Integration of the TOEIC® Speaking and Writing tests into English language courses for liberal arts education. *Hiroshima Studies in Language and Language Education*, 24, 1–14. <http://doi.org/10.15027/50443>
- Amano, S., Yamauchi, Y., Sakaue, T., Takahashi, Y., Nakagawa, A., & Yamauchi, K. (2023). Expansion of the Hiroshima University English Can-do List to support English learning outside the classroom. *Hiroshima Studies in Language and Language Education*, 26, 1–15. <http://doi.org/10.15027/53516>

- Educational Testing Service (n.d.). Content and format of the TOEIC® Speaking and Writing tests. <https://www.ets.org/toEIC/test-takers/speaking-writing/about/content-format>
- Institute for International Business Communication (n.d.). Test format and content. <https://www.iibc-global.org/english/toEIC/test/sw/about/format.html>
- Kida, S., Enokida, K., Uenishi, K., Kusanagi, K., Sakaue, T., Takita, F., Tatsukawa, K., Morita, M., Yamamoto, G., & Yoshikawa, L. (2017). Hiroshima daigaku eigo can-do risuto no kaihatsu [The development of a Hiroshima University English Can-Do List]. *Hiroshima Studies in Language and Language Education*, 20, 185–200. <http://doi.org/10.15027/42622>
- Morita, M. (2020). The impact on English education of adopting the quarter system at Hiroshima University. *Hiroshima Studies in Language and Language Education*, 23, 121–136. <http://doi.org/10.15027/48752>
- Uenishi, K., Sakaue, T., Lauer, J., Davies, W., Fraser, S., Howell, P., Selwood, J., Song, K., Morita, M., & Kida, S. (2017). Integrating the TOEIC® speaking test with small group classes. *Hiroshima Studies in Language and Language Education*, 20, 1–15. <http://doi.org/10.15027/42610>
- Uenishi, K., Sakaue, T., Fraser, S., Davies, W., Lauer, J., Selwood, J., Song, K., Morita, M., & Kida, S. (2018). Integrating the TOEIC® writing test with small-group classes. *Hiroshima Studies in Language and Language Education*, 21, 1–12. <http://doi.org/10.15027/45290>
- Uenishi, K., Sakaue, T., Fraser, S., Davies, W., Lauer, J., Selwood, J., Song, K., Morita, M., & Kida, S. (2019). Evaluating the second year of a program that integrates TOEIC® speaking and writing with small group classes. *Hiroshima Studies in Language and Language Education*, 22, 45–57. <http://doi.org/10.15027/47082>
- Uenishi, K., Fraser, S., Takita, F., Sakaue, T., Selwood, J., & Kida, S. (2020). Incorporating the TOEIC® speaking and writing test into regular first-year English writing classes. *Hiroshima Studies in Language and Language Education*, 23, 137–148. <https://doi.org/10.15027/48753>

## APPENDIX. R code Used in Statistical Analysis

R code

### **Robust t-test with 5% means trimming and bootstrapping**

```
# Step 1: Load the required packages
```

```
library(WRS)
```

```
library(readr)
```

```
# Step 2: Load the data and calculate the differences
```

```
data <- read_csv("Dataset(n=89).csv")
```

```
data$diff_scores <- data$Posttest - data$Pretest
```

```
# Step 3: Perform the robust paired samples t-test using the trimpb() function with the entire dataset of 89 participants
```

```
trimpb_result <- trimpb(data$diff_scores, tr = .05, alpha = .05, nboot = 2000)
```

```
print(trimpb_result)
```

## ABSTRACT

### **Exploring Directions for Improving General English Writing Courses through the TOEIC® Writing Test**

Shuichi AMANO

Daniel HOUGHAM

Marshall HIGA

Atsushi NAKAGAWA

Institute for Foreign Language Research and Education

Hiroshima University

Yuka TAKAHASHI

Global Education Center

Tokyo Kasei University

Simon FRASER

Institute for Foreign Language Research and Education

Hiroshima University

In this paper, we evaluate the efficacy of liberal arts English courses using the TOEIC® Writing Tests IP. The paper analyzes the outcomes of the TOEIC® Writing Test IP in 2022, contrasting these with the data from 2018 and 2019. The results in 2022, while having average scores close to those of 2019, exhibited less variability, indicating that instruction adapted necessary to a post-pandemic world has been achieved. This study is important in evaluating the current status and future challenges of writing instruction within these courses. The study lends support to the notion that proficiency tests can be valuable tools in liberal arts English courses, facilitating objective assessment and strategic enhancements in instructional methodologies, and thereby maintaining or enhancing students' English proficiency. For example, in our situation, our analysis has led us to consider the needs of students with the highest levels of English, and whether there may be better ways of organizing courses to help such students improve their writing proficiency.

## 要 旨

### TOEIC® ライティングテストを活用した 教養英語ライティング科目の改善の方向性の探求

天 野 修 一

ダニエル・ホフム

マーシャル・ヒガ

中 川 篤

広島大学外国語教育研究センター

高 橋 有 加

東京家政大学グローバル教育センター

サイモン・フレイザー

広島大学外国語教育研究センター

本稿では、TOEIC® Writing Tests IP を用いて、リベラルアーツ英語コースの有効性を精査した。2022年における TOEIC® Writing Tests IP の結果を、2018年および2019年のデータと対比させながら丹念に描き出している。特に、2022年は COVID-19の予防措置のため、Writing テストのみが実施された。2022年の結果は、2019年の結果に近い平均点ながら、変動が抑制されており、パンデミック後の指導に必要な適応の達成を示している。本研究は、これらのコースにおけるライティング指導の現状と今後の課題を評価する上で基礎となるものである。本研究は、習熟度テストがリベラルアーツの英語コースにおいて実に価値あるツールとして機能し、客観的な評価と指導方法の戦略的な強化を促進し、それによって学生の英語習熟度を維持または向上させる可能性があるという考え方を支持するものである。たとえば、私たちは分析から、最もレベルが高い学生たちのニーズを検討し、彼らがライティング能力を向上させるのに役立つ、よりよいコース編成についての検討の必要性を見出すことができた。