



Toeic Performance, Program-Level Disparities, and Career Readiness of Aviation Students: School Year 2025–2026

Dr. Marianne Shalimar G. Del Rosario

The Research Department, PATTS College of Aeronautics, Parañaque City, Philippines

DOI: <https://doi.org/10.68050/JAMS.2026.376>

Received: 01 Sept 2026; Accepted: 07 Sept 2026; Published: 30 September 2026

ABSTRACT

This descriptive-comparative study examined the Test of English for International Communication (TOEIC) Listening and Reading performance of 1,367 students at an aeronautical school in Parañaque for School Year 2025–2026, drawing on the institutional management report produced by Princeton Assessments and Training, Inc., the aeronautical school's TOEIC partner since 2014. The study compared results across year level, testing point, and academic program, and evaluated program averages against the score floors that Japan-based partner companies (ANA, JASCO, and KGS) and the United States-based partner NORDAM require for international internship placement. Results showed an institutional mean score of 762, placing the aeronautical school at CEFR B1, one band below the college's own B2 benchmark, with reading trailing listening by a wide margin. The academic program explained far more of the score variation than year level: BS Aeronautical Engineering averaged 890, while the Aircraft Technician Course averaged 593, a gap exceeding two standard deviations. Only two of five technical programs met the 800-point Japan-partner floor at the mean, and none met NORDAM's 900-point floor. Nearly half of all test-takers (48 percent) fell below the CEFR B2 benchmark. A comparison of the December 2025 and March 2026 testing sub-periods showed a 71-point institutional gap driven mainly by a shift in test-taker composition, since March was a smaller session concentrated in the two lowest-scoring technical programs, not a genuine decline in ability. The study recommends program-specific reading interventions, earlier placement into support programs, and closer alignment between English instruction and each program's workplace register.

Keywords: aviation English, career readiness, CEFR, English language proficiency, TOEIC performance

INTRODUCTION

English remains the working language of international aviation, business process outsourcing, and cross-border employment, and the Philippines has built much of its labor-export economy on the claim that its graduates can communicate in English at a level foreign employers trust. That claim is tested every time a graduate sits for an international assessment such as the Test of English for International Communication (TOEIC), which measures the listening and reading skills needed to function in an English-speaking workplace rather than in an English classroom. For a specialized aeronautical school such as the aeronautical school, the stakes are higher still: pilots, aircraft technicians, and airline personnel operate in an industry where a misread instruction or a missed radio call carries consequences well beyond a failing grade.

Concerns about the direction of Filipino English proficiency are not new. Asio and Quijano (2023) found a clear link between the English proficiency of Filipino college students and their academic performance, arguing that students who communicate confidently in English are better equipped to comprehend and interpret the academic material they are tested on. Working with senior high school learners, Manuel (2022) reported a similar pattern and traced part of the problem to how English is taught rather than to any lack of aptitude among students, since instruction that emphasizes grammar drills over actual communication can leave learners able to answer multiple-choice items yet unable to use the language confidently under pressure.

Aviation compounds this ordinary challenge with a specialized one. Zhao (2023) compared the linguistic structure of everyday conversational English against the English used in air traffic control communication and



found that aviation English behaves as its own register, built around short, formulaic exchanges that leave little room for the negotiation of meaning available in casual conversation. Estival and Molesworth (2020) demonstrated how costly that lack of room can be, recording pilots in flight and showing experimentally that even brief lapses in shared understanding between pilot and controller create the conditions for miscommunication. Taken together, these findings suggest that a general command of English is necessary but not sufficient for aviation-track students, who need practice with a variety of English that behaves differently from the English used in ordinary conversation.

High-stakes English tests also shape how students study, for better or worse. Sadeghi, Ballıdağ, and Mede (2021) examined this washback effect among university students preparing for TOEFL iBT and a locally developed proficiency exam and found that neither test reliably increased students' motivation or self-directed study habits, regardless of how much institutional weight was placed on the results. This complicates the common assumption that simply requiring a test will push students to improve; the test appears to need pairing with support that helps students act on their results, or a mandatory testing requirement risks becoming a hurdle rather than a lever for learning.

Part of the barrier is psychological rather than purely linguistic. Silaj, Schwartz, and Siegel (2021) reviewed how test anxiety interferes with the cognitive processes students rely on during an exam, showing that anxious test-takers often understand material well enough in low-stakes settings but underperform once the exam carries real consequences, because anxiety consumes cognitive resources that would otherwise go toward recalling and applying what they already know. For a group of students whose test outcomes are tied to an overseas job placement, this dynamic is worth taking seriously rather than assuming that any observed shortfall is purely a matter of language ability.

Confidence and motivation shape performance in the opposite direction. Wu, Yang, Liu, and Liu (2022) surveyed university students in China and found that students' belief in their own ability to use English predicted achievement more strongly than anxiety did, suggesting that programs built around strengthening students' sense of competence, rather than simply drilling test content, may do more to raise scores over time. Related work on English-medium instruction supports this view: Kamasak and Sahan (2023) found that students' language proficiency was one of the strongest predictors of how well they performed in courses taught in English, ahead of factors such as prior grades in other subjects.

Reading in particular deserves attention on its own terms. Liao, Zhu, and Cheong (2021) studied Chinese-speaking students with limited English proficiency and found that weaknesses in independent reading and grammar skills carried over directly into weaker integrated writing performance, indicating that reading comprehension functions as a foundation that other language skills draw on rather than as an isolated skill. This has a direct bearing on any institution, the aeronautical school included, where reading is consistently the weaker of the two TOEIC components and is treated as the target of a dedicated intervention program.

Programs that prepare students for a specific industry benefit from English instruction built around that industry's actual demands rather than a generic curriculum. Al-Malki, Javid, Farooq, Algethami, and Al-Harthi (2022) conducted a needs analysis of the English required of tourism workers in Saudi Arabia and found a persistent gap between the English taught in classrooms and the English graduates were actually expected to use on the job, a gap they used to argue for purpose-built ESP materials rather than repurposed general-English textbooks. The same logic applies to a school whose programs range from aeronautical engineering to airline business administration and tourism management, each of which sends graduates into a different register of workplace English.

Ultimately, institutions track English proficiency this closely because it affects whether graduates find and keep relevant work. Kamaruddin, Ahmad, Husain, and Abd Hamid (2021) studied graduate employability at a Malaysian public university in the aftermath of the COVID-19 disruption and found that communication skills remained one of the attributes employers weighed most heavily, even as the labor market tightened. In the Philippine setting, Albina and Sumagaysay (2020) traced the employment outcomes of information technology graduates and found that a majority secured work related to their field, but a meaningful share did not, a gap



they linked partly to skills that employers expect but that formal coursework does not always emphasize strongly enough, communication among them.

Read together, this body of work points to a consistent picture: English proficiency among Filipino students is uneven and has drawn sustained scholarly concern; aviation and technical industries demand a specialized register of English that ordinary instruction does not automatically deliver; high-stakes testing changes student behavior only when paired with genuine support; anxiety and self-efficacy shape performance independently of raw ability; reading proficiency underlies performance in other skills; and, in the end, all of this converts into consequences for employability that students, parents, and institutions can measure in very concrete terms, such as whether a graduate qualifies for an overseas internship. The aeronautical school sits squarely inside this picture. Its students are drawn from technical and business programs that feed directly into internship pipelines with named international partners, its TOEIC results are already collected as an institutional requirement, and the gap between what the literature predicts and what its own data show is the gap this study sets out to measure.

Background of the Study

Since 2014, the aeronautical school in Paranaque has partnered with what is now Princeton Assessments and Training, Inc. (formerly Hopkins International Partners, Inc.), the exclusive Philippine administrator of the TOEIC, TOEFL Junior, and TOEFL ITP examinations under the ETS Preferred Network, to administer the TOEIC Listening & Reading test to its students as an embedded academic requirement rather than an optional credential. Students first take the examination, referred to internally as TOEIC 1, during their first year, where it fulfills a requirement of the Speech Laboratory sequence (SPLB 300 and SPLB 400). They sit for it again, as TOEIC 2, during their third year, where it fulfills a requirement of the English Plus subject (ENGL 400). TOEIC 2 carries the greater practical weight of the two administrations, because it is the score that partner companies use when screening students for the international internship placements that many aeronautical school programs are built around.

Those placements come with explicit score floors. Japan-based partner companies — ANA, JASCO, and KGS among them — require a minimum TOEIC score of 800 from students in the college's technical courses (BS Aeronautical Engineering, BS Aircraft Maintenance Technology, BS Air Transportation, BS Avionics Technology, and the Aircraft Technician Course). NORDAM, the college's United States-based partner, sets a higher floor of 900. These are not aspirational targets; they are the literal cutoffs that separate a student who qualifies for a placement from one who does not, which is what makes an institution-wide accounting of TOEIC performance more than an academic exercise.

For School Year 2025–2026, the aeronautical school administered the TOEIC Listening & Reading test to 1,367 students across nine academic programs — BS Aeronautical Engineering, the Aircraft Technician Course, BS Aircraft Maintenance Technology, BS Air Transportation, BS Avionics Technology, BS Airline Business Administration, BS Tourism Management, BS Industrial Engineering, and Pre-Engineering — with testing concentrated among first-year students taking TOEIC 1 and third-year students taking TOEIC 2. The resulting institutional management report, produced by Princeton Assessments and Training, Inc., forms the entire data basis of this study and is examined here against the eleven-year trend the same partnership has produced since AY 2014–2015.

Theoretical Framework

This study is anchored on Canale and Swain's (1980) model of communicative competence, later elaborated by Canale (1983), which describes the ability to use a language effectively as the integration of four component competencies: grammatical competence (control of vocabulary, syntax, and the sound system), sociolinguistic competence (the ability to use language appropriately for a given social situation and register), discourse competence (the ability to connect utterances into a cohesive and coherent whole), and strategic competence (the ability to compensate for breakdowns in communication, whether caused by limited proficiency or by the situation itself). The model treats these four competencies as jointly constituting communicative competence rather than ranking one above the others.

Theoretical Framework: Canale and Swain's Model of Communicative Competence

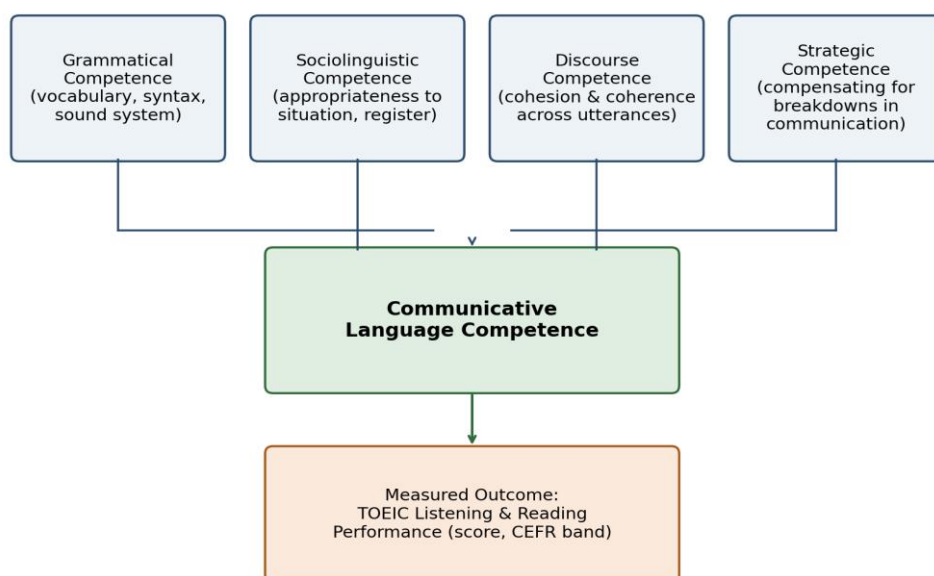
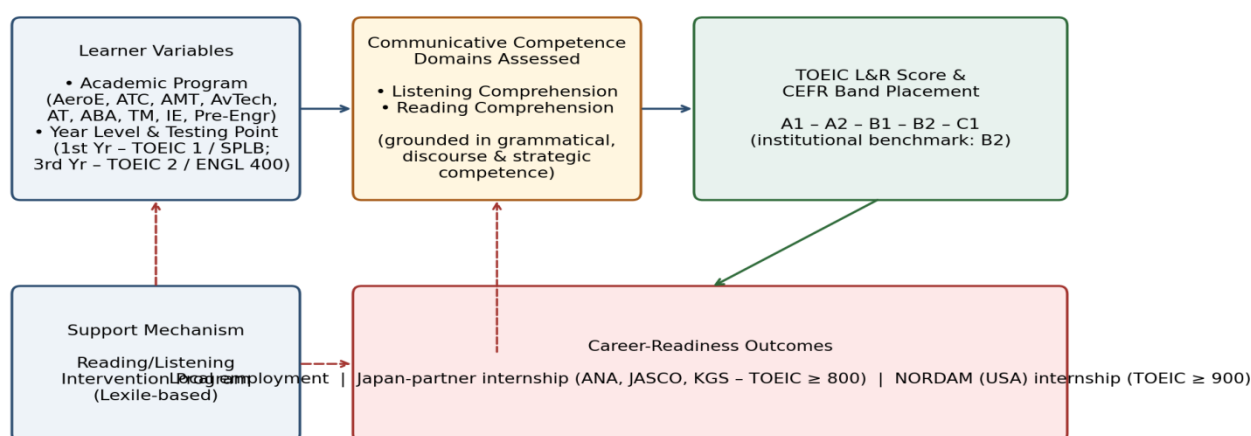


Figure 1. Theoretical Framework: Canale and Swain's Model of Communicative Competence

The TOEIC Listening & Reading test operationalizes this model within a narrower, receptive scope. Its listening section draws primarily on grammatical and discourse competence to test whether a student can follow the structure and cohesion of a spoken exchange, while its reading section draws on grammatical and discourse competence to test comprehension of written workplace texts. Neither section directly tests sociolinguistic or strategic competence, which is precisely why the test result — a total score placed on a CEFR band — is treated in this study as a partial but standardized proxy for overall communicative competence, not a complete measure of it. That distinction matters when the findings are interpreted: a student's TOEIC score speaks to a portion of communicative competence, and the recommendations proposed later follow from that limitation rather than around it.

Conceptual Framework

Conceptual Framework: PATTS TOEIC Proficiency-to-Career Pathway Model



Dashed arrows indicate the feedback loop: below-benchmark performance routes students into the intervention program, which is intended to raise subsequent competence and outcomes.

Figure 2. Conceptual Framework: Aeronautical School TOEIC Proficiency-to-Career Pathway Model



The conceptual framework adapts the theoretical model above to the specific variables this study examines. It begins with the learner variables that structure the aeronautical school's TOEIC administration — academic program and year level — since the same test is given at two different points in a student's academic life for two different institutional purposes. These variables feed into the communicative competence domains the TOEIC actually samples, listening and reading comprehension, which the test converts into a total score and a corresponding CEFR band against an institutional benchmark of B2. That placement, in turn, feeds into career-readiness outcomes with real consequences: general local employment prospects on one branch, and on the other, eligibility for the Japan-partner and NORDAM international internship pipelines, each gated by its own score floor. A feedback loop connects the outcome stage back to a support mechanism — the aeronautical school's Lexile-based reading and listening intervention program — so that a below-benchmark result is treated as an entry point into a structured second attempt at the same competence domains, rather than a dead end. This framework, rather than a generic input-process-output chain, is what the results and analysis section is organized around: each stage in the diagram corresponds to a specific research question below.

Statement of the Problem

This study examined the TOEIC Listening & Reading performance of aeronautical school students for School Year 2025–2026 in relation to program-level career-readiness benchmarks. Specifically, it sought to answer the following questions:

1. What is the overall level of TOEIC Listening & Reading performance of the aeronautical school students for SY 2025–2026, in terms of (a) mean total score, (b) mean listening score, (c) mean reading score, and (d) CEFR level?
2. What is the eleven-year trend of the aeronautical school's institutional TOEIC performance from AY 2014–2015 to AY 2025–2026?
3. What is the level of TOEIC performance when respondents are grouped according to (a) year level and testing point (TOEIC 1 among first-year students; TOEIC 2 among third-year students) and (b) academic program?
4. Is there a significant difference in TOEIC Listening & Reading performance across year level and across academic program?
5. To what extent do program-level TOEIC averages meet the minimum score requirements set by partner companies for international internship placement — a TOEIC score of 800 or above for the Japan-based partners (ANA, JASCO, and KGS) among the college's technical courses, and 900 or above for the United States-based partner, NORDAM?
6. What proportion of students fall below the institution's CEFR B2 benchmark and may be considered candidates for the reading and listening intervention program?
7. How did TOEIC performance and test-taker composition differ between the December 2025 and March 2026 testing sub-periods that together make up SY 2025–2026, both institution-wide and by academic program?

Hypothesis

H0(1): There is no statistically significant difference in TOEIC Listening & Reading mean scores among the aeronautical school students when grouped according to year level.

H0(2): There is no statistically significant difference in TOEIC Listening & Reading mean scores among the aeronautical school students when grouped according to academic program.

H0(3): There is no difference in institutional mean TOEIC Listening & Reading performance between the December 2025 and March 2026 testing sub-periods.

Significance of the Study

The findings of this study are expected to benefit the following:



- Students, particularly those enrolled in technical programs tied to international internship placements, who stand to gain a clearer picture of where their program stands relative to the score floors that determine their eligibility, and who may use the study's recommendations to seek intervention support earlier in their academic program rather than after a TOEIC 2 result has already closed off an option.
- Faculty handling Speech Laboratory and English Plus subjects, who can use the program-level and skill-level findings to adjust instructional emphasis, particularly toward reading comprehension, which this study identifies as the weaker of the two tested skills.
- The Guidance and Counseling Office, which may use the findings on test anxiety and motivation discussed in the review of literature to design support that addresses the psychological, and not only the linguistic, dimension of test performance.
- The Research Department and college administration, who can use the study as a baseline for evaluating the reading and listening intervention program and for deciding whether program-specific, rather than institution-wide, interventions are warranted, given the disparities the data reveal across programs.
- Industry partners, including the Japan-based and United States-based companies named in this study, who may find in these results a more precise account of which the aeronautical school programs are, and are not, currently producing graduates who clear their respective score requirements.
- Future researchers, who may extend this study by tracking the same cohort longitudinally from TOEIC 1 to TOEIC 2, by examining the qualitative reasons behind program-level disparities, or by evaluating the actual effectiveness of the intervention program once it has been implemented for a full cycle.

METHODOLOGY

Design

This study used a quantitative, descriptive-comparative research design. It is descriptive because it characterizes the current state of the aeronautical school students' TOEIC Listening & Reading performance in terms of central tendency, dispersion, and proficiency banding, and comparative because it examines whether that performance differs meaningfully across year level and academic program. The study relies entirely on secondary institutional data: the SY 2025–2026 TOEIC management report produced by Princeton Assessments and Training, Inc. for the aeronautical school, together with the eleven-year historical summary contained in the same report. No new instrument was administered for purposes of this study, and no additional data were collected from students beyond the institution's regular testing cycle.

Respondents

The respondents were the 1,367 the aeronautical school students who sat for the TOEIC Listening & Reading test during SY 2025–2026 as part of their regular academic requirements, drawn from nine academic programs: BS Aeronautical Engineering, BS Aircraft Maintenance Technology, BS Air Transportation, BS Avionics Technology, the Aircraft Technician Course, BS Airline Business Administration, BS Tourism Management, BS Industrial Engineering, and Pre-Engineering. Because the TOEIC is embedded in course requirements rather than administered to a drawn sample, this figure represents a total enumeration of test-takers for the period; no separate sampling procedure was applied.

Setting

The study was conducted at an aeronautical education institution located in Parañaque City, Metro Manila, Philippines. The TOEIC test itself was administered under standardized conditions set by Princeton Assessments and Training, Inc., the exclusive Philippine licensee for TOEIC administration.

Instrumentation

The primary instrument was the TOEIC Listening & Reading test, developed by Educational Testing Service (ETS) and administered locally by Princeton Assessments and Training, Inc. The test yields a Listening score



and a Reading score, each on a scale of 5 to 495, which combine into a total score on a scale of 10 to 990. Scores are further mapped onto Common European Framework of Reference for Languages (CEFR) bands, from A1 (beginner) through C1 (advanced), and onto Lexile reading measures, which the institution uses to match students to appropriately leveled reading material for its intervention program. This study did not construct or validate a new instrument; it analyzed the results of an existing, externally administered, standardized test as reported in the institutional management report.

Data Analysis

Data were analyzed using descriptive statistics — means, standard deviations, and percentage distributions — computed by Princeton Assessments and Training, Inc. and reported at the institutional level, by year level, and by academic program. Comparative analysis relied on the z-scores already included in the institutional report, which express each subgroup's mean score relative to the overall institutional mean and standard deviation; a z-score whose magnitude departs meaningfully from zero was treated as indicating a subgroup that performs differently from the institutional average. Program-level means were then directly compared against the two career-readiness benchmarks specified in the significance of the study — a TOEIC score of 800 for Japan-partner eligibility among technical courses, and 900 for NORDAM eligibility — to determine how many, and which, programs currently meet each threshold.

Ethical Considerations

This study used institutional, aggregate-level TOEIC data rather than data collected directly from individual students, and no student was contacted, surveyed, or tested for purposes of this research. Permission to use the SY 2025–2026 management report and the accompanying eleven-year historical summary was secured from the aeronautical school's Research Department prior to analysis. All data presented in this study are reported at the level of year level and academic program rather than at the level of the individual student; no student name, student number, or other individually identifying information appears anywhere in this study, in keeping with the Data Privacy Act of 2012 (Republic Act No. 10173) and its implementing rules on the processing of personal and academic information held by educational institutions. Where the source report identifies individual perfect scorers by name for institutional recognition purposes, this study reports only the aggregate count of perfect scorers for the period and does not reproduce any student's name. The researcher affirms that no funding source or personal interest influenced the interpretation of the findings, and that every figure reported in the results and analysis section is traceable to the institutional management report cited throughout this study.

RESULTS AND ANALYSIS

Results are presented in the order of the statement of the problem so that each finding answers a specific research question directly.

Overall TOEIC Listening & Reading Performance, SY 2025–2026 (SOP 1)

Indicator	Result
Total test takers	1,367
Mean total score (out of 990)	762
Standard deviation	136
Mean listening score (out of 495)	420
Mean reading score (out of 495)	342
Overall CEFR level	B1 (Intermediate)

Table 1. Institutional TOEIC Listening & Reading Performance, SY 2025–2026

The aeronautical school students posted an institutional mean of 762 out of a possible 990, placing the college as a whole at CEFR B1 (Intermediate) — one full band below the institution's own B2 (Upper Intermediate)

benchmark. The gap between the two tested skills is the more striking finding: the mean listening score of 420 out of 495 (85 percent of the maximum) considerably outpaces the mean reading score of 342 out of 495 (69 percent of the maximum). This asymmetry is consistent with Liao et al.'s (2021) observation that reading comprehension functions as a foundation for other language skills; when reading lags this far behind listening within the same population, it points to reading, specifically, as the more urgent target for remediation, a pattern that recurs throughout the program-level results below.

Eleven-Year Institutional Trend, AY 2014–2015 to AY 2025–2026 (SOP 2)

Academic Year	Test Takers	Mean Total Score	CEFR
2014–2015	1,733	630	B1
2015–2016	2,702	659	B1
2016–2017	1,177	691	B1
2017–2018	1,231	690	B1
2018–2019	1,720	726	B1
2019–2020	1,757	722	B1
2020–2021	1,092	806	B2
2021–2022	1,025	816	B2
2022–2023	2,174	778	B1
2023–2024	2,255	785	B2
2024–2025	1,780	743	B1
2025–2026	1,367	762	B1

Table 2. Institutional TOEIC Trend, AY 2014–2015 to AY 2025–2026

The eleven-year record shows a long, mostly gradual climb from 630 in AY 2014–2015 to a peak of 816 in AY 2021–2022, coinciding with the pandemic-era shift to remote instruction, followed by a decline across the four most recent cycles to the current 762. The institution has not returned to its 2021–2022 peak in any of the four years since, and the current mean sits closer to the levels recorded in the late 2010s than to the 2020–2022 high point. Read on its own, this decline could plausibly be attributed to a rebound in enrollment volume after the pandemic dip — the population more than doubled between AY 2021–2022 (1,025 test takers) and AY 2022–2023 (2,174) — diluting a smaller, possibly more self-selected pandemic-era cohort with a larger and more representative one. The trend nonetheless establishes that the SY 2025–2026 mean of 762 is not an isolated low point but part of a four-year downward drift that predates the specific cohort examined in this study.

Performance by Year Level and Testing Point (SOP 3a)

Year Level	Examinees	Mean	Listening	Reading	z-score	% ≥ 785
3rd Year (TOEIC 2)	521	792	433	359	0.221	62%
4th Year	117	769	426	343	0.044	55%
The aeronautical school (institutional)	1,367	762	420	342	0	52%
5th Year	9	754	422	322	-0.059	56%
1st Year (TOEIC 1)	671	744	411	332	-0.14	45%
2nd Year	49	689	393	296	-0.544	37%

Table 3. TOEIC Performance by Year Level, SY 2025–2026

Third-year students taking TOEIC 2 — the administration that internship-screening partner companies actually use — outperformed first-year students taking TOEIC 1 by 48 points on average (792 versus 744), a



standardized difference of roughly 0.36 standard deviations. This is the expected direction of movement: students gain nearly a full year of Speech Laboratory and general coursework between the two administrations. Second-year students, a small and largely irregular or repeating group of 49, posted the weakest result at any year level (689, $z = -0.544$), well below both first-year and third-year cohorts; because this group is small and non-representative of a normal cohort, it should not be read as evidence that proficiency declines between first and third year.

Performance by Academic Program (SOP 3b)

Program	Examinees	Mean	Listening	Reading	z-score	% ≥ 785
BS Aeronautical Engineering (AeroE)	94	890	465	425	0.934	93%
BS Industrial Engineering (IE)	12	866	458	408	0.758	92%
Pre-Engineering	104	847	455	391	0.618	79%
BS Avionics Technology (AvTech)	107	802	436	366	0.287	66%
BS Air Transportation (AT)	360	777	426	351	0.103	55%
The aeronautical school (institutional)	1,367	762	420	342	0	52%
BS Aircraft Maintenance Technology (AMT)	488	750	419	331	-0.088	47%
BS Airline Business Administration (ABA)	35	708	386	322	-0.397	43%
BS Tourism Management (TM)	62	672	378	294	-0.669	18%
Aircraft Technician Course (ATC)	105	593	350	243	-1.243	9%

Table 4. TOEIC Performance by Academic Program, SY 2025–2026

Program explains far more of the variation in TOEIC performance than year level does. The spread between the strongest program (BS Aeronautical Engineering, 890) and the weakest (Aircraft Technician Course, 593) is 297 points — more than two standard deviations — compared with a 103-point spread across year levels. Engineering and pre-engineering programs cluster at the top, all posting z-scores above +0.6; the Aircraft Technician Course sits more than a full standard deviation below the institutional mean ($z = -1.243$), the single largest departure recorded anywhere in the SY 2025–2026 data.

The institutional report also assigns each program its own internal target score, set higher for the engineering-track programs (850 for BS Aeronautical Engineering, Pre-Engineering, and BS Industrial Engineering) than for the other technical and business programs (750 for BS Avionics Technology, BS Air Transportation, BS Airline Business Administration, and BS Tourism Management; 725 for the Aircraft Technician Course and BS Aircraft Maintenance Technology). Measured against these internal targets rather than the external partner-company floors, the Aircraft Technician Course shows the widest shortfall by far, 132 points under its own 725 target, followed by BS Tourism Management (78 points under a 750 target) and BS Airline Business Administration (42 points under). BS Aeronautical Engineering, BS Avionics Technology, BS Aircraft Maintenance Technology, BS Air Transportation, and BS Industrial Engineering all meet or exceed their respective internal targets, and Pre-Engineering sits within 3 points of its target. The pattern matches the partner-company comparison in the next section: the same three programs — the Aircraft Technician Course, BS Tourism Management, and BS Airline Business Administration — fall short under both the institution's own internal benchmark and, where applicable, the external internship floor.

Significance of Differences Across Year Level and Program (SOP 4)

The institutional report's z-scores provide a standardized measure of how far each subgroup departs from the population mean, though they are descriptive rather than the product of a formal inferential test such as ANOVA. Judged against that standard, the differences across academic program are large enough to reject $H_0(2)$: a spread of more than two standard deviations between the highest- and lowest-performing programs, anchored by the Aircraft Technician Course's z-score of -1.243, is not consistent with a population in which



program membership makes no difference to TOEIC outcomes. The differences across year level are smaller and more mixed. The gap between first-year and third-year students (the two testing points that matter most operationally) is real and in the expected direction, but it is modest next to the program-level spread, and the second-year group's outlier result reflects a small, non-representative subgroup rather than a stable trend. H0(1) is therefore only partially supported: year level shows a genuine but comparatively minor difference, while H0(2) is not supported by the data.

Program-Level Readiness Against Internship Score Floors (SOP 5)

Technical Program	Mean Score	Meets Japan-Partner Floor (≥ 800)	Meets NORDAM Floor (≥ 900)
BS Aeronautical Engineering	890	Yes	No (10 points short)
BS Avionics Technology	802	Yes (marginal)	No
BS Air Transportation	777	No (23 points short)	No
BS Aircraft Maintenance Technology	750	No (50 points short)	No
Aircraft Technician Course	593	No (207 points short)	No

Table 5. Mean TOEIC Scores of Technical Programs Against Partner-Company Score Floors

Among the five technical programs whose graduates are eligible for the Japan-partner internship pipeline, only two — BS Aeronautical Engineering (890) and BS Avionics Technology (802, clearing the floor by just two points at the mean) — currently post program averages above the 800-point requirement. The remaining three (BS Air Transportation, BS Aircraft Maintenance Technology, and the Aircraft Technician Course) fall short at the mean by margins ranging from 23 to 207 points, which means that even an average-performing student in those three programs would not currently qualify for the Japan-partner pipeline without additional score gains. Against NORDAM's higher 900-point floor, no program clears the requirement at the mean; BS Aeronautical Engineering comes closest, 10 points under.

Proportion of Students Below the CEFR B2 Benchmark (SOP 6)

CEFR Band	Number of Students	Percentage
C1 (Proficient)	57	4%
B2 (Upper Intermediate) — institutional benchmark	656	48%
B1 (Intermediate)	544	40%
A2 (Elementary)	109	8%
A1 (Beginner)	1	0%

Table 6. CEFR Band Distribution, SY 2025–2026 ($n = 1,367$)

Combining the A1, A2, and B1 bands shows that 654 students — 48 percent of the SY 2025–2026 testing population — fall below the institution's CEFR B2 benchmark. On the reading component specifically, the institutional report flags 459 students, or 34 percent of test takers, as candidates for a reading intervention program, a figure the report arrives at independently of the CEFR banding and one that reinforces the pattern already visible in Table 1: reading, more than listening, is where the institution's proficiency gap concentrates.

December 2025 versus March 2026 Testing Sub-Periods (SOP 7)

Program	Dec 2025 n	Dec 2025 Mean	Mar 2026 n	Mar 2026 Mean
BS Aeronautical Engineering	94	890	0	n/a
BS Industrial Engineering	12	866	0	n/a
Pre-Engineering	102	848	2	770
BS Avionics Technology	103	808	4	635



BS Air Transportation	340	782	20	695
The aeronautical school (institutional)	1,152	774	215	703
BS Aircraft Maintenance Technology	339	759	149	732
BS Tourism Management	39	669	23	677
BS Airline Business Administration	34	715	1	475
Aircraft Technician Course	89	609	16	503

Table 7. TOEIC Performance by Testing Sub-Period, SY 2025–2026

December 2025 was the primary testing window, covering 1,152 of the 1,367 SY 2025–2026 test-takers (84 percent) across all nine programs, including the two strongest-performing programs, BS Aeronautical Engineering and BS Industrial Engineering, both of which had zero test-takers in March. March 2026 was a much smaller supplementary session of 215 test-takers (16 percent), and its composition was heavily skewed toward a single program: BS Aircraft Maintenance Technology alone accounted for 149 of the 215 March test-takers, 69 percent of that session.

The institutional mean fell from 774 in December to 703 in March, a 71-point gap that looks like a decline but is better read as a composition effect than a drop in ability over the school year. December's population included the entire Aeronautical Engineering and Industrial Engineering cohorts, the two highest-scoring programs; March's population was concentrated in the Aircraft Technician Course and BS Aircraft Maintenance Technology, two of the three lowest-scoring programs institution-wide. Within the programs that had a meaningful population in both sub-periods, scores were consistently lower in March than in December — the Aircraft Technician Course fell from 609 to 503, BS Avionics Technology from 808 to 635, and BS Airline Business Administration from 715 to 475 on a single test-taker — a pattern consistent with March functioning as a retake or completion window for students who had not yet tested, or had not met a required benchmark, in December, rather than as a second independent sample of the same population. This composition difference means $H_0(3)$ cannot be rejected on the institutional figures alone; the December-to-March gap is confounded with which programs are represented in each session, and a claim of an actual within-student decline would require matching individual retake records, which fall outside the scope of this study.

DISCUSSION

CONCLUSIONS

The following conclusions are drawn directly from the results reported above, in the order of the statement of the problem.

1. The aeronautical school students' SY 2025–2026 TOEIC performance sits at an institutional mean of 762, placing the college at CEFR B1 — one band below its own B2 benchmark — with reading (342/495) trailing listening (420/495) by a wide margin, confirming reading as the specific skill in greatest need of support.
2. The institution's TOEIC performance has declined across the four most recent academic years, from a peak of 816 in AY 2021–2022 to 762 in SY 2025–2026, indicating that the current shortfall is part of a multi-year drift rather than a single off-year result.
3. Academic program, far more than year level, determines where a student is likely to land on the proficiency scale: engineering and pre-engineering programs cluster well above the institutional mean, while the Aircraft Technician Course and BS Tourism Management cluster well below it.
4. The magnitude of the program-level spread — more than two standard deviations between the highest- and lowest-performing programs — does not support the null hypothesis that program membership

makes no difference to TOEIC outcomes, while the year-level difference, though real, is comparatively modest.

5. Only two of the five technical programs eligible for the Japan-partner internship pipeline currently meet the 800-point score floor at the mean, and no program meets NORDAM's 900-point floor at the mean, meaning that for three of five technical programs, the average graduate is not yet internship-ready by these specific external standards.
6. Close to half of the SY 2025–2026 testing population (48 percent) falls below the institutional CEFR B2 benchmark, and about a third (34 percent) has been identified through the reading component specifically as candidates for intervention, confirming that the gap this study documents is not confined to a handful of underperforming programs but is a substantial, institution-wide concern.
7. December 2025 and March 2026 are not comparable samples of the same population: December covered 84 percent of test-takers across all nine programs, including the two strongest-performing ones, while March was a smaller, supplementary session in which BS Aircraft Maintenance Technology alone made up 69 percent of test-takers. The 71-point institutional gap between the two sub-periods reflects this shift in composition rather than a within-student decline, and $H_0(3)$ cannot be rejected on the institutional data alone.

RECOMMENDATIONS

The following recommendations correspond to the stakeholder groups identified in the significance of the study.

1. For the Aircraft Technician Course, BS Aircraft Maintenance Technology, and BS Air Transportation — the three technical programs currently below the Japan-partner floor — the Research Department and program heads should consider a program-specific reading intervention rather than a single institution-wide track, following Al-Malki et al.'s (2022) argument that English instruction is more effective when built around the specific register learners will actually use, in this case technical maintenance and ground-operations English rather than generic business English.
2. The Guidance and Counseling Office should be brought into TOEIC 2 preparation directly, drawing on Silaj et al.'s (2021) findings on test anxiety and Wu et al.'s (2022) findings on self-efficacy, so that third-year students in low-performing programs receive support for test-taking confidence and not only additional content review.
3. Because reading, not listening, accounts for most of the institutional shortfall, the Lexile-based reading intervention currently offered later in the program should be made available to flagged students immediately after TOEIC 1 in the first year, rather than waiting until the TOEIC 2 result confirms the gap in the third year, when the internship-application window is already narrowing.
4. Faculty teaching the Speech Laboratory sequence and English Plus should incorporate aviation- and industry-register listening and reading materials specific to each technical program, in line with Zhao (2023) and Estival and Molesworth's (2020) findings that aviation English functions as a distinct register that general English instruction does not automatically cover.
5. Given Sadeghi et al.'s (2021) finding that a mandatory test does not by itself change student motivation, the college should pair each TOEIC administration with explicit, program-specific goal-setting — showing a BS Aircraft Maintenance Technology student, for example, exactly how many points separate their program average from the 800-point Japan-partner floor — rather than treating the test purely as a graduation requirement to be cleared.
6. Industry partners and the college administration should periodically review whether the number of instructional hours devoted to English proficiency in the Aircraft Technician Course and BS Tourism Management curricula is proportionate to the size of the gap those two programs currently show, since both post the lowest scores and, in the case of the Aircraft Technician Course, the largest shortfall against its relevant score floor.
7. Because the March 2026 session was disproportionately made up of Aircraft Technician Course and BS Aircraft Maintenance Technology retake or late test-takers, the Research Department should treat that session as a distinct, higher-risk population for intervention purposes and route resources accordingly,



rather than reading the December-to-March score drop as evidence that performance declines over the school year.

8. Future researchers should track individual students longitudinally from TOEIC 1 to TOEIC 2 and, within that, match individual records across the December and March sub-periods rather than comparing the two sessions as if they were independent samples; this would let a future study isolate genuine retake improvement from the composition effect identified here, and should also evaluate the reading intervention program's actual effect once a full cohort has completed it and investigate the qualitative reasons behind the Aircraft Technician Course's persistently low performance across multiple testing cycles.

REFERENCES

1. Al-Malki, E. A. A., Javid, C. Z., Farooq, M. U., Algethami, G. F., & Al-Harthi, A. A. (2022). Analysis of the English language needs of the Saudi tourism workforce: A first step into designing ESP teaching materials. *International Journal of Learning, Teaching and Educational Research*, 21(2), 72–88. <https://doi.org/10.26803/ijlter.21.2.5>
2. Albina, A. C., & Sumagaysay, L. P. (2020). Employability tracer study of Information Technology Education graduates from a state university in the Philippines. *Social Sciences & Humanities Open*, 2(1), 100055. <https://doi.org/10.1016/j.ssaho.2020.100055>
3. Asio, H. T., & Quijano, M. G. (2023). English proficiency and academic performance of college students. *International Journal of Multidisciplinary Research and Analysis*, 6(3), 1301–1317. <https://doi.org/10.47191/ijmra/v6-i3-50>
4. Canale, M. (1983). From communicative competence to communicative language pedagogy. In J. C. Richards & R. W. Schmidt (Eds.), *Language and communication* (pp. 2–27). Longman.
5. Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1(1), 1–47.
6. Estival, D., & Molesworth, B. R. C. (2020). Errors in air-ground pilot communication: An experimental study. *The ESPecialist*, 41(3). <https://doi.org/10.23925/2318-7115.2020v41i3a5>
7. Kamaruddin, M. I. H., Ahmad, A., Husain, M. A., & Abd Hamid, S. N. (2021). Graduate employability post-COVID-19: The case of a Malaysian public university. *Higher Education, Skills and Work-Based Learning*, 11(3), 710–724. <https://doi.org/10.1108/HESWBL-05-2020-0114>
8. Kamasak, R., & Sahan, K. (2023). Academic success in English medium courses: Exploring student challenges, opinions, language proficiency, and L2 use. *RELC Journal*, 55(3), 705–720. <https://doi.org/10.1177/00336882231167611>
9. Liao, X., Zhu, X., & Cheong, C. M. (2021). Direct and indirect effects of independent language skills on the integrated writing performance of Chinese-speaking students with low proficiency. *Reading and Writing*, 34(10), 2529–2557. <https://doi.org/10.1007/s11145-021-10152-7>
10. Manuel, J. B. (2022). English language proficiency of senior high school students. *Multidisciplinary Journal for Education, Social and Technological Sciences*, 9(1), 71–86. <https://doi.org/10.4995/muse.2022.16638>
11. Princeton Assessments and Training, Inc. (2026). TOEIC management report: PATTS College of Aeronautics, SY 2025–2026 [Institutional report].
12. Sadeghi, K., Ballıdağ, A., & Mede, E. (2021). The washback effect of TOEFL iBT and a local English Proficiency Exam on students' motivation, autonomy and language learning strategies. *Heliyon*, 7(10), e08135. <https://doi.org/10.1016/j.heliyon.2021.e08135>
13. Silaj, K. M., Schwartz, S. T., & Siegel, A. L. M. (2021). Test anxiety and metacognitive performance in the classroom. *Educational Psychology Review*, 33(4), 1809–1834. <https://doi.org/10.1007/s10648-021-09598-6>
14. Wu, X., Yang, H., Liu, J., & Liu, Z. (2022). English use anxiety, motivation, self-efficacy, and their predictive effects on Chinese top university students' English achievements. *Frontiers in Psychology*, 13, 953600. <https://doi.org/10.3389/fpsyg.2022.953600>
15. Zhao, W. (2023). A corpus-based study on aviation English from the perspective of systemic functional linguistics. *Discourse & Communication*, 17(5), 630–661. <https://doi.org/10.1177/17504813231171644>