Jurnal Pendidikan Bahasa
Volume 16. Nomor 4, Desember 2026 | ISSN: 2088-0316 | e-ISSN: 2685-0133
https://doi.org/10.37630/jpb.v16i4.5209

Hospitality Students' Communicative English Competence Gaps against Indonesian National Occupational Standards (SKKNI): A Mixed-Methods Importance–Performance Analysis

Ni Made Ernila Junipisa[1,*], Ni Nyoman Nidya Trianingrum[1]

[1]Universitas Triatma Mulya

*Corespondence: ernila.junipisa@triatmamulya.ac.id

Abstract

Indonesian vocational hospitality programmes must produce graduates who communicate in English at the level required by the Indonesian National Occupational Competency Standards (SKKNI), yet English for Specific Purposes (ESP) courses are seldom designed from a gap analysis against them. This study mapped the importance and performance of SKKNI-derived communicative tasks in front office, food and beverage service, and housekeeping, and explored the causes of the gaps. An explanatory sequential mixed-methods design was used. Eight experts judged the content validity of 24 items using Aiken's V; 40 students and 28 hotel practitioners then completed the questionnaire. Data were analyzed with Importance–Performance Analysis (IPA) and non-parametric tests, and interviews with 12 informants were analyzed thematically. Two items failed content validation. Across the remaining 22 items, students' performance (M = 3.48) fell below importance (M = 4.20), with significant gaps on 16 tasks. The priority quadrant contained complaint handling, menu explanation, and allergy enquiries. Industry rated graduates' performance lower than students rated themselves in all three departments, although both groups ranked the tasks almost identically (ρ = .88). Internship length was positively associated with performance. Interviews attributed the gaps to anxiety in unpredictable interactions, limited service vocabulary, and script-based practice.

Keywords: Communicative Competence; English for Specific Purposes; Hospitality SKKNI; Importance–Performance Analysis; Needs Analysis

Received: 15 Okt 2026; **Revised:** 29 Okt 2026; **Accepted:** 30 Okt 2026; **Available Online:** 9 Nov 2026

This is an open access article under the CC - BY license.

INTRODUCTION

For hotel employees, English is not a subject they once studied but a working tool used every hour of a shift. Check-in, food orders, and complaint handling all take place in spoken exchanges with guests whose accents, speech rates, and expectations vary widely. Fluency in general English does not guarantee success in these exchanges. When hotel staff and EFL students in Thailand responded to the same guest complaints, both used a similar overall move structure, but the students were markedly less accurate in professional steps such as thanking the guest, apologizing, and proposing a solution (Soratcha & Gampper, 2026). Hotel practitioners now regard staff English as a form of professional capital that combines linguistic agility, intercultural empathy, and emotional intelligence (Anwari et al., 2026).

In Indonesia, these workplace demands are codified in the Indonesian National Occupational Competency Standards (*Standar Kompetensi Kerja Nasional Indonesia*, SKKNI). The standardization system is regulated by Minister of Manpower Regulation No. 2/2016 (Permenaker RI, 2016), and the standard currently in force for the hotel sector was issued on 20 June 2024 through Minister of Manpower Decree No. 125/2024, which revoked Decree No. 145/2018 (Kepmenaker RI, 2018). Each SKKNI unit of competency is broken down into elements and performance criteria (*kriteria unjuk kerja*), and many performance criteria in guest-service units can only be met through spoken communication with guests. The same criteria are used by professional certification bodies (*Lembaga Sertifikasi Profesi*, LSP) when they assess prospective workers. SKKNI therefore already provides a measurable, industry-endorsed description of the target situation. What is rarely done is to

ISSN: 2088-0316 | e-ISSN: 2685-0133
https://doi.org/10.37630/jpb.v16i4.5209

use it as the starting point for designing language instruction, even though national policy on vocational revitalization calls for curricula aligned with labor-market needs (Peraturan Presiden (Perpres), 2022).

In English for Specific Purposes (ESP), such alignment is pursued through needs analysis. Hutchinson and Waters (1987) distinguish *necessities* (what the target situation demands), *lacks* (the distance between those demands and learners' current ability), and *wants* (what learners themselves wish to learn). Needs analysis underpins every subsequent design decision, from materials selection to assessment (Anthony, 2018; Woodrow, 2018). Long (2005) adds that information on needs should be triangulated across sources, because learners' perceptions alone are often inaccurate. Canale & Swain (1980) framework helps explain why tasks that require negotiation of meaning and choices of politeness are harder: they depend on strategic and sociolinguistic competence, not grammar alone.

A number of studies have documented English needs in the hospitality sector, but they define and measure those needs in ways that limit their use for course design. Among vocational high school hospitality students, speaking and hotel-specific vocabulary were the most needed, yet instruction did not match their field (Jubaedah & Wirza, 2021). Front office staff at a hotel in Bali acknowledged a real gap between their ability and job demands, especially in speaking and writing (De Carvalho & Waruwu, 2026), and more than 80% of 86 hospitality students still struggled with spontaneous interaction (Anwari et al., 2026) Students in Thailand likewise identified listening and speaking as the most problematic skills (Lertchalermtipakoon et al., 2021), while in Vietnam telephone reservations and complaint handling were among the most demanding tasks for front-desk staff (Nguyen, 2025). Reviews of English for tourism and hospitality purposes summarize the same pattern (Zahedpisheh et al., 2017). Across these studies, needs are described by skill (listening, speaking, writing) rather than by the workplace tasks against which graduates are certified, and the results are reported as percentages or means that do not indicate which tasks should be taught first.

These studies leave three gaps. First, the target situation is usually defined from respondents' perceptions or from the literature rather than from the official occupational standards used in certification. Nguyen (2025) is the only study above that starts from a national occupational standard, the Vietnam Tourism Occupational Standards, but its analysis remains descriptive, and no comparable study has derived hospitality English tasks from the SKKNI performance criteria. Second, most studies report percentages or means per skill (Anwari et al., 2026; Jubaedah & Wirza, 2021; Lertchalermtipakoon et al., 2021) and therefore cannot show which tasks are both important and weakest, which is the information a syllabus designer needs to allocate limited class hours. Third, students' perceptions are rarely compared statistically with industry judgements. This matters because self-assessment is only a moderate indicator of language performance: a meta-analysis of 67 studies found a pooled correlation of .47 between self-assessment and measured performance (Li & Zhang, 2021), and teachers tend to rate students' ability lower than students rate themselves (Mao & Zhou, 2024). A needs analysis based on students alone may therefore misjudge the size of the gap. In this journal, studies of English for tourism students have so far addressed the digital transformation of instruction (Trianingrum & Junipisa, 2026) and the basic proficiency of prospective students (Jokosaharjo et al., 2026), not competence gaps against occupational standards.

To address these gaps, this study combines three procedures. First, the target situation is defined by the hotel-sector SKKNI: every questionnaire item is derived from a performance criterion in Decree No. 125/2024. Second, priorities are determined with Importance–Performance Analysis (IPA), borrowed from service management (Martilla & James, 1977). IPA plots each attribute on a Cartesian grid according to its importance and performance, so that attributes which matter but perform poorly become immediately visible. The technique continues to be refined (Phadermrod et al., 2019) and has recently been used to evaluate English language instruction (Chiou et al., 2024), but it has not been used to rank ESP tasks against an occupational standard. Here, the importance axis represents industry-defined *necessities* and the performance axis represents students' current state, so the distance between them can be read as *lacks*. Third, students' and industry's ratings are compared with rank-based tests (Mann-Whitney U and Spearman's ρ), which show both whether the two groups differ in level and whether they agree on which tasks are hardest. The novelty of the study therefore lies in an SKKNI-anchored, quadrant-based priority map for hospitality ESP that is checked against industry judgement and explained with qualitative data.

ISSN: 2088-0316 | e-ISSN: 2685-0133
https://doi.org/10.37630/jpb.v16i4.5209

Accordingly, the study addresses four questions: (1) How important are SKKNI-based English communicative tasks, and how well are they performed, according to students and industry practitioners? (2) Which tasks are priorities for improvement according to the IPA map? (3) Do students' and industry's performance ratings differ, and does performance differ by internship experience? (4) What explains the gaps, and what are the implications for hospitality ESP syllabus design?

METHOD

This study used an explanatory sequential mixed-methods design (Creswell & Plano Clark, 2018), in which a quantitative phase measured the gaps and mapped priorities and a qualitative phase explained the results. The two phases were connected by selecting interview foci from items in the priority quadrant and items on which students and industry disagreed, and they were merged in a joint display to derive meta-inferences.

The quantitative phase involved two groups selected purposively (Etikan et al., 2016). The first comprised 40 hospitality students in semesters 4, 6, and 8. Students were included if they had completed the hospitality English course, which covers communicative tasks in all three departments; for this reason they rated all 24 tasks, even though all were registered in the front office department. The second comprised 28 hotel practitioners (managers, supervisors, and staff) who were included if they worked daily with interns or recent graduates. Practitioners rated only the tasks of their own department, because a rater can judge only the performance they observe directly; this gave 10 industry raters per item for front office, 10 for food and beverage service, and 8 for housekeeping. Sample size was determined by the availability of eligible respondents rather than by an a priori power analysis. A sensitivity analysis showed that the department-level comparisons (40 students vs 8–10 practitioners, $\alpha = .05$, power = .80) could detect only large differences ($d \geq 1.04$ for front office and food and beverage service; $d \geq 1.14$ for housekeeping), so item-level comparisons were interpreted with caution. To keep the two groups comparable, students' ratings were compared with industry ratings only on the items of the same department, with identical item wording and five-point scales, and with rank-based tests that do not assume equal group sizes. Table 1 presents the respondent profile.

Table 1. Profile of Quantitative-Phase Respondents

Group	Characteristic	Category	n (%)
Students (n = 40)	Gender	Female / Male	25 (62.5) / 15 (37.5)
	Age (years)	18–20 / 21–23 / ≥ 24	13 (32.5) / 23 (57.5) / 4 (10.0)
	Semester	4 / 6 / 8	13 (32.5) / 13 (32.5) / 14 (35.0)
	Internship experience	None / 3–6 months / > 6 months	13 (32.5) / 18 (45.0) / 9 (22.5)
Industry (n = 28)	Gender	Female / Male	17 (60.7) / 11 (39.3)
	Age (years)	25–34 / 35–44 / ≥ 45	18 (64.3) / 9 (32.1) / 1 (3.6)
	Position	Manager / Supervisor / Staff	10 (35.7) / 10 (35.7) / 8 (28.6)
	Department	Front office / Food and beverage / Housekeeping	10 (35.7) / 10 (35.7) / 8 (28.6)
	Work experience	< 3 / 3–5 / 6–10 / > 10 years	3 (10.7) / 9 (32.1) / 13 (46.4) / 3 (10.7)

The qualitative phase involved 12 informants: five students, two lecturers, three hotel supervisors, and two competency assessors, chosen so that every role involved in teaching, supervising, and assessing hotel English was represented.

The main instrument was a questionnaire of 24 English communicative tasks derived from the hotel-sector SKKNI (Kepmenaker RI, 2024). Each task was formulated from a performance criterion in a guest-service unit that requires spoken interaction with guests, covering front office (FO1–FO9), food and beverage service (FB1–FB9), and housekeeping (HK1–HK6). Each item was rated on two five-point Likert scales: importance for the job (1 = not at all important; 5 = very important) and performance (1 = not at all able; 5 = very able). Students rated their own performance, whereas practitioners rated the general performance of the interns and recent

ISSN: 2088-0316 | e-ISSN: 2685-0133
https://doi.org/10.37630/jpb.v16i4.5209

graduates they supervised. Content validity was judged by eight experts on a four-category relevance scale and computed with Aiken's V (Aiken, 1985; Retnawati, 2016); with eight experts and four categories, $V \geq .75$ ($p \leq .040$) was adopted as the validity threshold. Internal consistency of the performance scale was assessed with Cronbach's α and corrected item–total correlations. A semi-structured interview guide was developed after the quantitative phase.

Quantitative data were analyzed in four steps. Means and standard deviations of importance and performance were computed per item and per group. Students' importance and performance ratings were compared item by item with the Wilcoxon signed-rank test, with effect sizes $r = |Z|/\sqrt{N}$ and Holm-adjusted p values. Items were then plotted on an IPA grid with students' performance on the X axis and industry-rated importance on the Y axis, using the grand mean of each axis as the quadrant boundary; industry importance was used because industry defines and assesses the target situation, and the map was redrawn with students' importance ratings as a robustness check. Finally, students' and industry's performance ratings were compared with the Mann-Whitney U test per department and per item, agreement on task difficulty was assessed with Spearman's ρ between the two groups' item means, and differences by internship experience were tested with the Kruskal-Wallis test, Holm-adjusted pairwise comparisons, and partial Spearman correlations controlling for semester. Non-parametric tests were used because Likert data are ordinal, and all computations were performed in Python (SciPy).

Interviews were conducted between 14 and 24 September 2026, audio-recorded, transcribed, and analyzed thematically (Braun & Clarke, 2021). Codes were merged into themes and linked to the items they explained, and credibility was supported by source triangulation across informant roles. All participants gave consent, respondents and hotels were anonymized, and students' participation had no bearing on their grades.

RESULT AND DISCUSSION

Instrument Validity and Reliability

Twenty-two items obtained V values between .75 and 1.00 ($p \leq .040$) and were retained as valid. Two housekeeping items, HK4 (handling lost-and-found reports) and HK6 (explaining laundry procedures and charges), reached only $V = .71$ ($p = .078$). The experts' notes explain why: both tasks are more often handled through written procedures or forms and depend heavily on individual hotel policy. HK4 and HK6 were therefore excluded, and all subsequent inferential analyses and the IPA used 22 items.

The 22-item performance scale showed high reliability ($\alpha = .90$), with corrected item–total correlations between .27 and .71. The importance scale produced a very low α (.10). This reflects a ceiling effect: 85.9% of students' and 90.1% of practitioners' importance ratings were 4 or 5, leaving very little between-respondent variance (SD of respondent means = 0.14). Importance was therefore treated as an item-level judgement for positioning items on the grid rather than as a composite score.

Importance–Performance Gaps

Overall, students rated the SKKNI-based communicative tasks as highly important ($M = 4.20$, $SD = 0.14$) but rated their own performance considerably lower ($M = 3.48$, $SD = 0.45$). Thirty-seven of the 40 students reported a mean performance below their mean importance ($Z = -5.39$, $p < .001$, $r = .85$). Industry rated importance slightly higher ($M = 4.30$, $SD = 0.20$) and graduates' performance lower ($M = 2.91$, $SD = 0.44$). Compared against industry-rated importance, students' performance reached a conformity level of only 80.7%. Table 2 gives item-level results.

Table 2. Content Validity, Importance, Performance, and Gaps by Communicative Task

Code	Communicative task	V	I-S	P-S	Z	r	I-Ind	P-Ind	Quadrant
FO1	Greeting guests on arrival	0.88	4.55	4.00	−2.94*	0.46	4.60	4.00	II
FO2	Processing check-in	0.88	4.60	4.00	−3.59*	0.57	4.50	3.40	II
FO3	Taking reservations by phone	0.88	4.53	3.65	−4.39*	0.69	4.60	3.20	II
FO4	Explaining hotel facilities	0.79	4.28	3.90	−2.70*	0.43	4.20	3.30	IV
FO5	Giving information and directions	0.83	3.55	3.58	0.29	0.05	3.80	3.00	IV
FO6	Handling guest complaints	0.88	4.65	2.98	−5.36*	0.85	4.40	2.70	I

ISSN: 2088-0316 | e-ISSN: 2685-0133
https://doi.org/10.37630/jpb.v16i4.5209

Code	Communicative task	V	I-S	P-S	Z	r	I-Ind	P-Ind	Quadrant
FO7	Processing check-out and bills	0.88	4.48	3.55	−4.62*	0.73	4.30	3.20	IV
FO8	Taking and relaying messages	0.75	4.05	3.88	−1.16	0.18	4.20	3.30	IV
FO9	Offering room upgrades	0.79	3.93	2.78	−4.51*	0.71	4.00	2.80	III
FB1	Greeting and seating guests	0.75	4.20	4.05	−0.86	0.14	4.00	3.10	IV
FB2	Explaining the menu and ingredients	0.83	4.35	3.35	−4.70*	0.74	4.60	2.70	I
FB3	Recommending dishes	0.96	3.88	2.80	−4.84*	0.77	4.10	2.30	III
FB4	Taking and confirming orders	0.83	4.55	3.60	−4.72*	0.75	4.60	2.90	II
FB5	Asking about allergies and diets	0.88	4.45	3.28	−4.46*	0.71	5.00	2.60	I
FB6	Room-service orders by phone	0.92	4.13	3.18	−4.44*	0.70	4.20	2.70	III
FB7	Handling food/service complaints	0.83	4.53	2.78	−5.28*	0.83	4.80	1.90	I
FB8	Presenting the bill and payment	1.00	4.05	3.65	−2.12	0.34	4.30	2.90	IV
FB9	Small talk with guests	0.79	3.58	3.53	−0.36	0.06	3.70	3.00	IV
HK1	Greeting guests during room service	0.96	4.03	4.00	−0.22	0.04	4.50	3.38	II
HK2	Explaining the cleaning schedule	0.96	3.95	3.45	−2.85*	0.45	3.88	2.75	III
HK3	Handling special requests	0.88	3.93	3.30	−3.29*	0.52	4.00	2.75	III
HK4	Handling lost and found	0.71	—	—	—	—	—	—	Excluded
HK5	Responding to room complaints	0.96	4.15	3.20	−4.53*	0.72	4.50	2.38	I
HK6	Explaining laundry service	0.71	—	—	—	—	—	—	Excluded

Note. V = Aiken's index (8 experts); I-S and P-S = students' mean importance and performance (n = 40); Z and r = Wilcoxon signed-rank test of I-S vs P-S and its effect size; * = significant at α = .05 after Holm correction (22 tests); I-Ind and P-Ind = industry's mean importance and rating of graduates' performance (n = 10 for FO and FB, n = 8 for HK); Quadrant refers to Figure 1.

Significant gaps were found on 16 of the 22 tasks. The largest effects were for FO6 handling guest complaints (r = .85), FB7 handling food or service complaints (r = .83), FB3 recommending dishes (r = .77), FB4 taking and confirming orders (r = .75), and FB2 explaining the menu (r = .74). The six tasks without a meaningful gap (FO5, FO8, FB1, FB8, FB9, and HK1) share one feature: they are routine tasks with a fairly fixed conversational sequence, such as greeting guests, relaying messages, or making small talk.

Priority Mapping with Importance–Performance Analysis

Figure 1 shows the 22 tasks on the IPA grid, with boundaries at $\bar{X}$ = 3.48 (student performance) and $\bar{Y}$ = 4.31 (industry-rated importance). Table 3 summarizes the contents of each quadrant.

Quadrant I contains five tasks: three are complaint-handling tasks across all three departments (FO6, FB7, HK5), and two concern explaining the menu and allergies (FB2, FB5). FB5 deserves particular attention. All ten food and beverage raters gave it the maximum importance score of 5, the only item to reach a perfect score, yet students' performance was only 3.28 and industry's rating of graduates was 2.60. When the map was redrawn using students' importance ratings, four tasks remained in Quadrant I (FO6, FB2, FB5, FB7); only HK5 moved to Quadrant III because students considered it slightly less important than industry did. Two Quadrant IV items, FO7 and FB8, sit just below the importance boundary (4.30 vs 4.31), so their position should be read with caution.

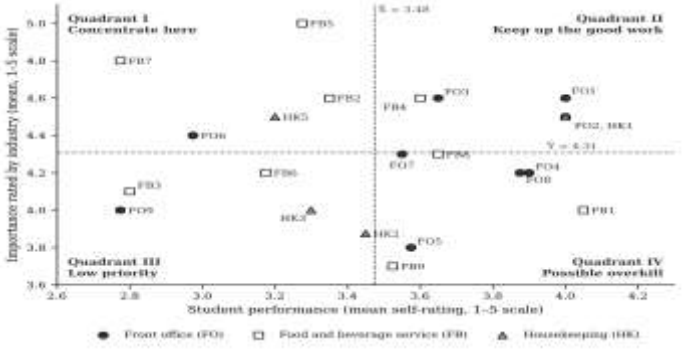


Figure 1. Importance–Performance Analysis Grid of 22 Communicative Tasks

ISSN: 2088-0316 | e-ISSN: 2685-0133
https://doi.org/10.37630/jpb.v16i4.5209

Table 3. Distribution of Communicative Tasks Across IPA Quadrants

Quadrant	Tasks	Shared characteristics
I. Concentrate here	FO6, FB2, FB5, FB7, HK5	Complaint handling; explaining the menu, ingredients, and allergies
II. Keep up the good work	FO1, FO2, FO3, FB4, HK1	Core, fixed-pattern tasks performed relatively well
III. Low priority	FO9, FB3, FB6, HK2, HK3	Persuasive selling, telephone work, and housekeeping requests
IV. Possible overkill	FO4, FO5, FO7, FO8, FB1, FB8, FB9	Routine tasks with above-average performance

Quadrant III holds an equally important finding. FO9 offering room upgrades and FB3 recommending dishes have the lowest student performance of all items (2.78 and 2.80), and industry also rated graduates' performance on FB3 very low (2.30). Both fall into Quadrant III only because their industry-rated importance (4.00 and 4.10) is slightly below the mean. These persuasive selling tasks are better described as hidden priorities. Moreover, industry-rated importance and student performance were virtually uncorrelated at item level (ρ = .03, p = .91): the tasks industry values most are not the tasks students perform best.

Students' versus Industry's Ratings

Table 4 compares mean performance by department as rated by students and by industry. In all three departments, students rated themselves as more capable than industry rated graduates, with the largest effect in food and beverage service.

Table 4. Performance as Rated by Students and by Industry, by Department (Mann-Whitney U Test)

Department	Students M (SD); Mdn	Industry M (SD); Mdn	U	p	r
Front office (9 items)	3.59 (0.45); 3.56	3.21 (0.30); 3.17	302.5	.013	.35
Food and beverage (9 items)	3.36 (0.54); 3.33	2.68 (0.22); 2.72	353.0	< .001	.52
Housekeeping (4 items)	3.49 (0.49); 3.50	2.81 (0.58); 2.75	258.5	.006	.39

Note. Students n = 40; industry n = 10 (front office), 10 (food and beverage), and 8 (housekeeping).

At item level, students' means exceeded industry's on 20 of the 22 tasks. After Holm correction, only FB1 greeting and seating guests differed significantly (U = 334.5, p = .009, r = .50), while FB7 and HK5 approached significance (p = .081). Because the number of industry raters per department was small, the department-level comparisons are more reliable. Notably, the two groups agreed closely on the ordering of tasks from easiest to hardest (ρ = .88, p < .001). The difference lies in level, not in rank order.

Performance by Internship Experience

Mean performance rose with internship length: 3.16 (SD = 0.38) for students with no internship, 3.48 (SD = 0.38) for 3–6 months, and 3.90 (SD = 0.27) for internships or work longer than six months. The difference was significant (H(2) = 15.17, p < .001, ε^2 = .39). Pairwise comparisons showed that the over-six-months group differed from the no-internship group (Holm p = .001, r = .78) and from the 3–6 months group (Holm p = .016, r = .51), whereas the difference between no internship and 3–6 months was not significant (p = .065). Because senior students were more likely to have completed an internship (ρ = .55), semester was controlled. The association between internship length and performance remained significant after controlling for semester (partial ρ = .48, p = .002), whereas semester was no longer significant after controlling for internship length (partial ρ = .21, p = .194). Male and female students did not differ (U = 186, p = .98).

Qualitative Findings

Thematic analysis of the interviews yielded four main themes. Table 5 lists the themes, informants, related items, and illustrative quotations.

ISSN: 2088-0316 | e-ISSN: 2685-0133
https://doi.org/10.37630/jpb.v16i4.5209

Table 5. Interview Themes and Illustrative Quotations

Theme	Informants	Related items	Illustrative quotation
Anxiety in unpredictable interactions	INF-01 (St), INF-08 (Su), INF-09 (Su)	FO6, FO9	"When a guest complains straight away in fast English, my mind goes blank and I just keep saying sorry." (INF-01)
Insufficient service vocabulary	INF-03 (St), INF-10 (Su)	FB2, FB3, FB5	"The names of ingredients and cooking methods in English are hard, especially when guests ask about allergies." (INF-03)
Script-based practice not aligned with the standards	INF-02 (St), INF-05 (St), INF-06 (L), INF-07 (L), INF-11 (A)	FO2, FO4, FO9, FB4, HK3	"We already use role-play, but not all scenarios follow the competency standards." (INF-07)
Internships build readiness	INF-04 (St), INF-12 (A)	FO1, FO2	"Students who have done an internship are much better prepared for the performance assessment." (INF-12)

Note. St = student; L = lecturer; Su = supervisor; A = competency assessor. Quotations translated from Indonesian.

The first theme, anxiety in unpredictable interactions, explains why the three complaint-handling tasks (FO6, FB7, HK5) combine Quadrant I positions with the largest gap effects (r = .72–.85). A student described how a complaint delivered in fast English left her repeating apologies: "When a guest complains straight away in fast English, my mind goes blank and I just keep saying sorry" (INF-01). Supervisors saw the same problem from the other side and named the ability to "calm an angry guest with polite sentences" as the most pressing need (INF-08). For upselling room upgrades (FO9; student performance 2.78), the barrier was not vocabulary but the fear of sounding pushy (INF-09). The second theme, insufficient service vocabulary, corresponds to the menu and allergy tasks. Students found "the names of ingredients and cooking methods in English" hard, "especially when guests ask about allergies" (INF-03), which matches the distance between industry importance (5.00) and student performance (3.28) on FB5. A supervisor added that interns tend to answer "only what is necessary" when asked to recommend dishes (INF-10), consistent with industry's very low rating of FB3 (2.30). The third theme, script-based practice not aligned with the standards, explains why the gaps persist and why students rate themselves higher than industry does. Memorized dialogues fail when guests ask questions outside the script (INF-02), housekeeping tasks are rarely practiced orally (INF-05), and large classes limit speaking practice (INF-06). In a lecturer's words, "we already use role-play, but not all scenarios follow the competency standards" (INF-07). Assessors see the consequence when candidates struggle to re-confirm orders (INF-11), the same task (FB4) on which industry rated graduates at 2.90 against students' self-rating of 3.60. The fourth theme came from a student and an assessor who both saw internships as building confidence; as the assessor put it, "students who have done an internship are much better prepared for the performance assessment" (INF-12), which mirrors the rise in mean performance from 3.16 to 3.90 across internship groups.

Integration of Quantitative and Qualitative Findings

Table 6 integrates the findings of both phases. Some meta-inferences confirm the numbers; others extend their meaning.

Table 6. Joint Display of Quantitative and Qualitative Findings

Quantitative finding	Qualitative explanation	Meta-inference
FO6, FB7, HK5 (complaints) in Quadrant I; largest gap effects (r = .85 and .83 for FO6 and FB7)	Anxiety and loss of words when guests complain (INF-01, INF-08)	Confirmation: gaps stem from demands on strategic competence, not vocabulary alone
FB2 and FB5 in Quadrant I; FB5 importance 5.00 but performance 3.28	Difficulty naming ingredients, cooking methods, and allergies (INF-03)	Confirmation and expansion: the vocabulary gap also concerns guest safety

ISSN: 2088-0316 | e-ISSN: 2685-0133
https://doi.org/10.37630/jpb.v16i4.5209

Quantitative finding	Qualitative explanation	Meta-inference
FO9 and FB3 in Quadrant III, yet with the lowest performance (2.78 and 2.80)	Fear of sounding pushy; staff answer only what is necessary (INF-09, INF-10)	Expansion: hidden priorities masked by the mean-importance boundary
Students rate performance higher than industry in all three departments	Memorized dialogues fail off-script; candidates struggle to re-confirm orders (INF-02, INF-11)	Expansion: students' self-ratings are anchored in controlled classroom situations
HK2 and HK3 show significant gaps despite lying in Quadrant III	Housekeeping tasks are rarely practiced orally (INF-05)	Expansion: housekeeping is under-represented in instruction
Performance rises with internship length, also after controlling for semester	Internships build confidence and assessment readiness (INF-04, INF-12)	Confirmation: authentic experience is associated with performance

Discussion

The clearest finding is that the largest gaps lie not in routine tasks but in tasks that cannot be completed with memorized conversational patterns. Routine tasks follow relatively fixed sequences and can be mastered through model dialogues. Complaint handling, allergy explanations, and menu recommendations, by contrast, require students to negotiate meaning, de-escalate emotion, and choose an appropriate level of politeness in real time, which is the domain of strategic and sociolinguistic competence in Canale & Swain (1980) framework. This is consistent with the pragmatic shortcomings of EFL students in complaint responses (Soratcha & Gampper, 2026), with Nguyen's (2025) finding on the difficulty of complaint handling, and with (Anwari et al., 2026) report of students' difficulty with spontaneous interaction.

The contribution of the present study lies in how these gaps are located and ranked. Earlier hospitality needs analyses describe needs by skill and in percentages (Jubaedah & Wirza, 2021; Lertchalermtipakoon et al., 2021) or derive the target situation from an occupational standard without prioritizing it (Nguyen, 2025). Anchoring every item to an SKKNI performance criterion and placing it on an IPA grid turns the needs analysis into a priority list expressed in the same units used in certification. Three results follow that a skill-based analysis would not show: complaint handling emerges as a priority in all three departments rather than in front office alone; persuasive selling tasks (FO9, FB3) are revealed as hidden priorities that the mean-importance boundary would otherwise mask; and routine tasks in Quadrant IV can be identified as candidates for reduced class time. Because the grid uses industry-rated importance and is checked against industry performance ratings, the priorities also carry the judgement of the people who assess graduates. IPA has been used to evaluate English instruction (Chiou et al., 2024), but to our knowledge this is its first use to prioritize ESP tasks against a national occupational standard.

The second finding is that students rated themselves more highly than industry did, particularly in food and beverage service. This mirrors the teacher–student rating divergence reported by Mao and Zhou (2024) and the only moderate link between self-assessment and measured language performance (Li & Zhang, 2021). The interviews suggest why: students judge themselves against familiar classroom scenarios, whereas industry judges performance in front of real guests. Because the rank order of task difficulty was almost identical (ρ = .88), the problem is not that students do not know which tasks are hard but that they miscalibrate their level. This supports Long's (2005) call for triangulated needs analysis, since an analysis relying on students alone would paint a rosier picture than reality.

The third finding concerns the misalignment between importance and performance. A near-zero correlation (ρ = .03) indicates that the tasks industry values most are not the ones students perform best. This pattern is consistent with the mismatch between needs and instruction reported by Jubaedah and Wirza (2021) and with lecturers' acknowledgement that role-play scenarios have not yet been mapped to the competency standards.

The fourth finding highlights the role of authentic experience. Students with internships longer than six months performed markedly better, and the association held after controlling for semester. This is in line with Yahya et al. (2025), who reported an effect of fieldwork on students' English communication in hospitality services. The cross-sectional design, however, does not permit causal interpretation.

ISSN: 2088-0316 | e-ISSN: 2685-0133
https://doi.org/10.37630/jpb.v16i4.5209

Drawing on the integrated findings, the hospitality ESP syllabus can be restructured in three stages, each tied to the way SKKNI is assessed. In certification, an LSP assessor judges a candidate against the performance criteria of each unit, using evidence such as observed demonstrations or simulations and oral questioning, and records the unit as competent or not yet competent. The same logic can be built into classroom assessment by turning each priority task into a short simulation and each related performance criterion into an observable checklist item, scored as met or not yet met and supplemented by a rubric for language quality. A foundation stage covers the routine tasks of Quadrants II and IV briefly through model dialogues; time spent on FO5, FO8, FB1, and FB9 can be reduced. A consolidation stage concentrates on Quadrant I and the hidden priorities (FO9, FB3) through graded simulations with unpredictable turns, telephone practice without visual cues, and a glossary of ingredients and allergens. Structured role-play has been shown to reduce speaking anxiety among tourism ESP students (Maldin & Sianipar, 2025), making it relevant to the anxiety theme found in the interviews. A transfer stage brings assessors and practitioners into mock competency assessments that use the same performance-criteria checklists, in line with the principle that ESP materials and assessment should be derived from real workplace tasks (Anthony, 2018; Woodrow, 2018). Table 7 shows how this applies to the priority tasks.

Table 7. Linking Priority Tasks to SKKNI-Based Assessment and Classroom Simulation

Priority tasks (IPA position)	Observable actions derived from the related SKKNI performance criteria	Classroom simulation and scoring
Complaint handling: FO6, FB7, HK5 (Quadrant I)	Listens to and acknowledges the complaint, apologizes, clarifies the problem, proposes and confirms a solution, and closes politely	Unscripted simulations with a practitioner as the guest; checklist of actions plus a rubric for politeness and fluency
Menu and allergies: FB2, FB5 (Quadrant I)	Describes dishes, ingredients, and cooking methods accurately; asks about allergies and dietary needs; confirms and relays them	Menu-explanation role-play with an allergen glossary; confirming and relaying the allergy is a critical item
Persuasive selling: FO9, FB3 (hidden priorities)	Recommends options matched to the guest's needs, explains benefits and price, and accepts a refusal politely	Upselling simulations with varied guest responses; scored on matching the offer to the need and on polite closing
Telephone tasks: FO3, FB6	Takes details, spells and reads them back, and confirms the request	Audio-only practice without visual cues; read-back accuracy checklist
Routine tasks (Quadrants II and IV)	Greets guests, gives information, and processes standard transactions	Brief model-dialogue practice; spot checks during mock competency assessments

Note. Each task is judged competent only when all of its observable actions are met, mirroring the competent / not yet competent decision used in SKKNI certification.

The study has several limitations. Performance was measured through perceptions rather than direct performance tests. Student respondents came from a single programme and were all registered in the front office department. The small number of industry raters per department (8–10) limited the power of item-level tests, and the importance scale showed a ceiling effect. Future research should use multi-institution samples, simulation-based performance tests, and quasi-experimental designs.

CONCLUSION

This study shows that students' performance on SKKNI-based communicative tasks remains below the importance of those tasks, with significant gaps on 16 of the 22 valid tasks. The main instructional priorities are complaint handling in all three departments and explaining menus and allergies, together with persuasive selling tasks whose performance is the lowest even though they lie outside Quadrant I. Industry rates graduates' performance lower than students rate themselves in every department, although both groups agree on the order of task difficulty, and performance increases with internship length. The gaps are explained by anxiety in unpredictable interactions, limited service vocabulary, and script-based practice that is not aligned with the competency standards. Hospitality programmes are therefore advised to build ESP syllabi on SKKNI performance criteria ordered by IPA priority, to assess priority tasks through simulations scored against those criteria, to involve assessors and practitioners in language assessment, and to consider longer, more structured

ISSN: 2088-0316 | e-ISSN: 2685-0133
https://doi.org/10.37630/jpb.v16i4.5209

internships; the validated 22-item instrument can serve as a periodic diagnostic tool before internships and competency assessments.

References

Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement*, *45*(1), 131–142. https://doi.org/10.1177/0013164485451012

Anthony, L. (2018). *Introducing English for specific purposes*. Routledge.

Anwari, H., Pratolo, B. W., & Sulisworo, D. (2026). Voices behind the front desk: A triangular lens on English competence in hospitality education. *English Language Teaching Educational Journal*, *9*(1), 17–30. https://doi.org/10.12928/eltej.v9i1.15079

Braun, V., & Clarke, V. (2021). *Thematic analysis: A practical guide*. SAGE.

Canale, M., & Swain, M. (1980). Theoretical Bases of Communicative Approaches to Second Language Teaching and Testing. *Applied Linguistics*, *I*(1), 1–47. https://doi.org/10.1093/applin/I.1.1

Chiou, K.-C., Huang, Y.-M., & Silalahi, A. D. K. (2024). Employing the performance evaluation method to examine the efficacy of English language instruction: Insights from primary educational institutions in Taiwan. *Education Sciences*, *14*(6), 672. https://doi.org/10.3390/educsci14060672

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE.

De Carvalho, A. H., & Waruwu, O. (2026). ESP needs analysis of front office staff in a hotel setting: A qualitative study. *Research on English Language Education*, *8*(1), 30–42. https://doi.org/10.57094/relation.v8i1.4883

Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics*, *5*(1), 1–4. https://doi.org/10.11648/j.ajtas.20160501.11

Hutchinson, T., & Waters, A. (1987). *English for specific purposes: A learning-centred approach*. Cambridge University Press.

Jokosaharjo, S., Sari, N. L. K. J. P., & Trianingrum, N. N. N. (2026). Analysis of the basic English proficiency of prospective university students based on diagnostic test results. *Jurnal Pendidikan Bahasa*, *16*(2), 194–198. https://doi.org/10.37630/jpb.v16i2.4391

Jubaedah, & Wirza, Y. (2021). An investigation of the needs analysis and English learning process for the hospitality students at a vocational high school. *Proceedings of the Fifth International Conference on Language, Literature, Culture, and Education (ICOLLITE 2021)*. https://www.atlantis-press.com/proceedings/icollite-21/125963477

Kepmenaker RI. (2018). *Keputusan Menteri Ketenagakerjaan Nomor 125 Tahun 2024 tentang Penetapan* Standar *Kompetensi Kerja Nasional Indonesia Kategori Penyediaan Akomodasi dan Penyediaan Makan Minum Golongan Pokok Penyediaan Akomodasi Bidang Hotel*. Kementerian Ketenagakerjaan Indonesia. https://jdih.kemnaker.go.id/peraturan/detail/2868/keputusan-menteri-ketenagakerjaan-nomor-125-tahun-2024

Kepmenaker RI. (2024). *Keputusan Menteri Ketenagakerjaan Nomor 125 Tahun 2024 tentang Penetapan Standar Kompetensi Kerja Nasional Indonesia Kategori Penyediaan Akomodasi dan Penyediaan Makan Minum Golongan Pokok Penyediaan Akomodasi Bidang Hotel*. Kementerian Ketenagakerjaan Indonesia. https://jdih.kemnaker.go.id/peraturan/detail/2868/keputusan-menteri-ketenagakerjaan-nomor-125-tahun-2024

Lertchalermtipakoon, P., Wongsubun, U., & Kawinkoonlasate, P. (2021). Need analysis: English language use by students in the tourism and hospitality and industry. *English Language Teaching*, *14*(3), 59–71. https://doi.org/10.5539/elt.v14n3p59

Li, M., & Zhang, X. (2021). A meta-analysis of self-assessment and language performance in language testing and assessment. *Language Testing*, *38*(2), 189–218. https://doi.org/10.1177/0265532220932481

ISSN: 2088-0316 | e-ISSN: 2685-0133
https://doi.org/10.37630/jpb.v16i4.5209

Long, M. H. (Ed.). (2005). *Second language needs analysis*. Cambridge University Press.

Maldin, S. A., & Sianipar, B. (2025). Role play to alleviate speaking anxiety: Action research in ESP class of Batam Tourism Polytechnic. *Cahaya Pendidikan*, *11*(1), 43–55. https://doi.org/10.33373/chypend.v11i1.7746

Mao, F., & Zhou, J. (2024). A needs analysis of ESP courses in colleges of art and design: Consensus and divergence. *PLoS ONE*, *19*(6), e0305210. https://doi.org/10.1371/journal.pone.0305210

Martilla, J. A., & James, J. C. (1977). Importance-performance analysis. *Journal of Marketing*, *41*(1), 77–79. https://doi.org/10.1177/002224297704100112

Nguyen, D. T. (2025). English proficiency requirements for hotel front-desk staff by the Vietnam Tourism Occupational Standards: Implications for ESP curriculum in Vietnam. *Educational Process: International Journal*, *18*, e2025491. https://doi.org/10.22521/edupij.2025.18.491

Peraturan Presiden (Perpres). (2022). *Peraturan Presiden (Perpres) Nomor 68 Tahun 2022 tentang Revitalisasi Pendidikan Vokasi dan Pelatihan Vokasi*. Jaringan Dokumentasi Dan Informasi Hukum Badan Pemeriksa Keuangan. https://peraturan.bpk.go.id/Details/207855/perpres-no-68-tahun-2022

Permenaker RI. (2016). *Peraturan Menteri Ketenagakerjaan Nomor 2 Tahun 2016 tentang Sistem Standardisasi Kompetensi Kerja Nasional*. Kementerian Ketenagakerjaan Indonesia. https://jdih.kemnaker.go.id/peraturan/detail/1203/peraturan-menteri-ketenagakerjaan-nomor-2-tahun-2016

Phadermrod, B., Crowder, R. M., & Wills, G. B. (2019). Importance-performance analysis based SWOT analysis. *International Journal of Information Management*, *44*, 194–203. https://doi.org/10.1016/j.ijinfomgt.2016.03.009

Retnawati, H. (2016). Proving content validity of self-regulated learning scale (The comparison of Aiken index and expanded Gregory index). *REID (Research and Evaluation in* Education*)*, *2*(2), 155–164. https://doi.org/10.21831/reid.v2i2.11029

Soratcha, T., & Gampper, C. (2026). A comparative move analysis of customer complaint responses between Thai hotel staff and EFL students. *An Online Journal of Education*, *21*(1), 1–18. https://doi.org/10.14456/ojed.2026.16

Trianingrum, N. N. N., & Junipisa, N. M. E. (2026). Transformasi digital dalam pendidikan bahasa Inggris bagi mahasiswa pariwisata. *Jurnal Pendidikan Bahasa*, *16*(2), 219–228. https://doi.org/10.37630/jpb.v16i2.4499

Woodrow, L. (2018). *Introducing course design in English for specific purposes*. Routledge.

Yahya, M., Rahmayanti, F., Widianto, A. T., & Ridwan, M. (2025). Impact of fieldwork on students' English communication in hospitality services. *LETS: Journal of Linguistics and English* Teaching *Studies*, *6*(2), 141–153. https://doi.org/10.46870/lets.v6i2.1556

Zahedpisheh, N., Abu Bakar, Z. B., & Saffari, N. (2017). English for tourism and hospitality purposes (ETP). *English Language Teaching*, *10*(9), 86–94. https://doi.org/10.5539/elt.v10n9p86