FOREIGN LANGUAGES AND VIETNAMESE

ENGLISH FOR SPECIFIC PURPOSES EDUCATION: ASSOCIATED FACTORS AND STRATEGIES FOR QUALITY IMPROVEMENT AT THE ENGLISH FOR SPECIFIC PURPOSES DEPARTMENT, HANOI UNIVERSITY

BUI LE MINH* - NGUYEN VU MINH ANH - NGUYEN THI QUYNH HOA*** - NGO THI THU TRANG******

ABSTRACT: This study examines the factors associated with the perceived quality of English for Specific Purposes (ESP) education at Hanoi University, drawing on survey responses from 849 first-year students across ten majors. The data were analyzed using descriptive statistics, multiple regression and thematic analysis. Students evaluated the overall quality of ESP education positively, with a mean score of 4.09 on a five-point scale. Learner factors (M = 3.85) and learning materials (M = 3.97) received the lowest ratings. The regression model comprising seven factors explained 84.3% of the variance in perceived quality. Assessment and feedback had the largest standardized regression coefficient ($\beta = 0.614$, $p < 0.001$), while learner factors had a smaller but statistically significant coefficient ($\beta = 0.067$, $p = 0.002$). Students' open-ended responses also highlighted errors in learning materials, delays in score reporting and limited opportunities for practical application. Based on these findings, the study recommends strengthening feedback, tailoring learning support to students' proficiency levels, reviewing learning materials and expanding opportunities for discipline-specific practice.

KEY WORDS: English for Specific Purposes; quality of training; students' perception; assessment and feedback; English-medium instruction.

RECEIVED: September, 29th, 2026. EDITED: November, 18th, 2026

1. Introduction

The fast growth of English-medium instruction (EMI) in Vietnamese higher education has turned English for Specific Purposes (ESP) from a marginal service course to an essential component of degree programs. Students' English proficiency is repeatedly reported to be the most serious constraint on EMI program effectiveness in Vietnam, with students completing a preparatory year of English reported to experience considerable shock when moving to content courses delivered in English (Dang, 2023). The ESP preparatory year is therefore a high-stakes intervention, and as such requires rigorous examination.

The English for Specific Purposes Department (ESPD) at Hanoi University is responsible for providing the first year of English education to all students enrolled in 10 English-medium majors. During this one acadenic year, students progress from the B1 to the B2 level of the Common European Framework of Reference (CEFR) through a combination of intensive skill development courses and a discipline-specific Language of the Specialism module. The department therefore straddles the interface of general English, English for Academic Purposes and subject-specific ESP, serving a large and diverse cohort within a limited timeframe.

Research on ESP, both international and in Vietnam, has shown many quality issues: lack of

*-****; *Hanoi University, Hanoi, Vietnam*

***Hanoi University, Hanoi, Vietnam; Email: anhnvm@hanu.edu.vn*

needs analysis, ESP teachers' weak knowledge of subject matter, generic or outdated materials, large classes, heterogeneity of proficiency, theory-heavy content, and assessment practices mismatched with workplace communication (Hoa & Mai, 2016; Bui, 2022; Chi & Vu, 2024; Vo et al., 2026). However, few studies have quantified the relative contribution of different factors to students' overall judgment of quality. Very few studies have also explored ESP programs functioning as a preparatory year for EMI, where the ESP course is the main academic activity in the first year, rather than an add-on to a Vietnamese-medium degree.

The present study seeks to address these gaps with three research questions:

1. How do first-year ESPD students assess the seven elements of ESP provision (teaching staff, curriculum, learning materials, teaching methods, learner factors, facilities and assessment and feedback) and the overall quality of their ESP education?
2. What are the major predictors of students' perceived quality of ESP education and the relative importance of these predictors?
3. What are the solutions suggested by the quantitative and qualitative data for the improvement of the quality of ESP education at ESPD?

2. Literature review

2.1. English for Specific Purposes and the concept of quality

ESP has been defined as an approach to language teaching where all decisions as to content and method are based on the learner's reason for learning (Hutchinson & Waters, 1987, as cited in Tran et al., 2019). Dudley-Evans and St John (1998, as cited in Hoa & Mai, 2016) specified this definition by distinguishing between absolute characteristics (ESP meets specific learner needs, uses the methodology and activities of the discipline it serves, and focuses on the language, skills, discourse and genres appropriate to those activities) and variable characteristics (ESP may be discipline-specific, may use a different methodology from general English, and is generally designed for intermediate or advanced adult learners). According to Basturkmen (2010) the design of ESP courses is cyclical, involving needs analysis, specification of outcomes, selection of content and methods and evaluation. The quality of an ESP course is by its very nature relational in this tradition: it is good to the extent that it identifies and responds to learners' target and learning needs.

Dou et al. (2023) identified learner-specific needs, course goals, properly designed materials, learner motivation, learning strategies, technology integration and intercultural awareness as factors shaping ESP outcomes. Bui (2022) surveyed 216 students and 32 teachers from five Vietnamese universities and revealed divergent perceptions among students and teachers on effective ESP teaching in four domains (task-based teaching, corrective feedback and assessment, discipline specifications, and culture integration).

2.2. Quality of training in higher education: a multi-dimensional construct

The SERVQUAL model (as discussed in Abdullah, 2006) conceptualizes quality as the difference between expectations and perceptions. Abdullah (2006) used this same logic for higher education when developing the HEdPERF instrument. Studies on Vietnamese applications of HEdPERF confirm that academic aspects (teaching staff and program design) are the strongest indicator of student satisfaction (Le et al., 2021). On the other hand, a cross-national study demonstrated that program features predicted satisfaction whereas teacher attributes and facilities did not, illustrating that the importance of different quality factors need to be determined throgh empirical research within specific contexts (Bahadur et al., 2024).

Biggs (1996) argued that quality is generated when there is constructive alignment between intended learning outcomes, teaching and learning activities, and assessment tasks. Assessment tells students what is important and shapes their learning behavior. Hattie and Timperley (2007) rated feedback as one of the most powerful influences on achievement. A subsequent meta-analysis of 435 studies found an average effect of d = 0.48, increasing to d = 0.99 for high-information feedback that

tells learners how to improve (Wisniewski et al., 2020). Assessment and feedback may therefore be a way in which students experience and judge the effectiveness of teaching, materials and curriculum.

2.3. Factors influencing ESP quality in Vietnam

Structural constraints on ESP have been repeatedly highlighted in Vietnamese studies. Hoa and Mai (2016) surveyed 362 students and teachers at eleven universities in Hanoi. The most reported difficulties were large class sizes, the quality of lectures and textbooks, students' limited vocabulary and general proficiency, and the linguistic distance between Vietnamese and English. Tran et al. (2019) evaluated ESP courses through questionnaires with 657 students, interviews, observations and document analysis and found deficiencies in syllabus coherence, relevance of materials, methodology and testing. Dong (2025) identified the lack of systematic needs analysis, limited subject-matter knowledge, insufficient localization of materials, and heterogeneous proficiency. Chi and Vu (2024) found that graduates perceived ESP courses as too theoretical, insufficiently specific and outdated. A mixed-methods evaluation with 1,463 students found positive perceptions of teaching methods and feedback, but low confidence in applying English professionally, generic content, insufficient productive-skills practice and variable assessment practices (Vo et al., 2026).

Learner-side variables are also important: low proficiency students in EMI preparatory contexts experience heightened anxiety and reduced engagement (Dang, 2023), and ESP researchers have long recognized that learners' present situation (existing proficiency, learning habits and attitudes) constrains what any course can achieve (Basturkmen, 2010; Dou et al., 2023).

2.4. Hypotheses and conceptual framework

Drawing on the ESP literature, higher-education quality literature and empirical findings from Vietnam reviewed above, the quality of ESP education is conceptualized as a student-perceived outcome influenced by seven input factors:

Teaching Staff (GV): ESP language knowledge of the lecturers, disciplinary knowledge, pedagogical knowledge, clarity, feedback, motivation, ability to connect language to professional situations.

Curriculum (CT): Course content that relates to the student's major, preparation for future EMI studies, logical sequencing of topics, reasonable workload, and development of discipline-specific communication skills.

Learning Materials (HL) - Materials' capacity to promote discipline-specific competencies and vocabulary acquisition, authentic and current examples from professionals within the field, attractiveness of the material, supplemental digital resources, and accessibility.

Pedagogical Methods (PP) - Active learning techniques, integration of skills practice, interaction among students, opportunities to practice English in a classroom setting, task-based assignments that reflect real-world disciplinary scenarios, integration of technology into the learning process, variety of instructional strategies employed, and cultivation of critical thinking abilities.

Learner Characteristics (NH) - Students' existing language proficiency levels, motivation to learn, amount of time allocated to independent studying, perceived value of the course, self-confidence in using English, and willingness to independently seek out additional resources.

Facilities(CS) - Classroom equipment, availability of e-learning platforms, library resources, information and communication technology (ICT) infrastructure, and class size.

Assessment and Feedback (DG) - Coverage of all four language skills (listening, speaking, reading, writing), variety of assessment instruments, balanced representation of continuous versus summative assessments, validity of assessment tools, authenticity of assessment tasks, clarity of assessment criteria, timeliness of feedback, and usefulness of feedback for improvement.

The perceived quality of ESP education (CL) is operationalized by satisfaction, comprehensive skill development, confidence for future educational and professional use, preparedness for the disciplinary faculty and fulfillment of expectations. H1 to H7 are seven hypotheses that each factor is positively related to perceived quality. The study also examined whether perceived quality differed

by gender, major, self-assessed CEFR level, recent ESP grade, major-related work experience, and possession of an international English certificate.

3. Materials and methods

3.1. Research context

The ESPD program includes first-semester intensive skill courses (B1 level), second-semester intensive skill courses (B2 level) and a discipline-specific Language of the Specialism module. All instruction is in English and is taught by ESPD lecturers. Successful completion is a prerequisite for moving to content courses taught in English by the disciplinary faculties starting from the second year. The target population was the 2025 intake (K2025), the first cohort to have experienced the current version of the program.

3.2. Research design and instruments

The study used a cross-sectional, mainly quantitative design, supplemented with qualitative data from open-ended questions. The questionnaire was designed in Vietnamese based on the literature reviewed in Section 2 and the ESPD's internal quality-assurance criteria. Three senior ESP lecturers reviewed the questionnaire for content validity, and it was informally piloted with a small group of students prior to administration. The questionnaire was divided into four sections. The first section, referred to as Part A, has six questions regarding the respondent's gender; their major; their self-assessed CEFR level; their last ESP grade; any relevant work experience/part-time employment related to their major; and whether they hold an international English certificate. The second section, referred to as Part B, includes 50 questions that measure the seven factor scales using a five-point Likert scale (1 = strongly disagree, 5 = strongly agree). These factor scales include: teaching staff (GV, 10 items); curriculum (CT, 5 items); learning materials (HL, 7 items); teaching methods (PP, 8 items); learner factors (NH, 7 items, of which NH7 is negative); facilities (CS, 5 items); and assessment & feedback (DG, 8 items). The third section, Part C, covers 5 items investigating the perceived quality of ESP education (CL) measured with the same scale. The fianl section of the questionnaire - Part D consists of several open-ended questions on course strengths and weaknesses; selection of at least three of six improvement areas, with an open text box for each; and a forced-choice question identifying the most influential quality factor, with an option to explain.

3.3. Sampling and data gathering

The questionnaire was distributed online using Google Forms between 25 and 28 May 2026, when students had completed most of the first-year program. The K2025 cohort comprised approximately 1,200 students across the ten English-medium majors, all of whom were invited via class groups and lecturers. We collected 849 full responses, with no missing values for parts A to C because every closed item was compulsory. The sample exceeds ten observations per item for factor analysis. Respondents were informed that participation was voluntary and anonymous and the results would be used only for program improvement and research.

3.4. Analysis of data

Data were analyzed in Python. Reliability was assessed using Cronbach's alpha, corrected item-total correlations and alpha-if-item-deleted statistics, with 0.70 as the reference threshold for alpha (Abdullah, 2006). Dimensionality was explored using principal component analysis (PCA) with varimax rotation, after assessing the Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy and Bartlett's test of sphericity. PCA was treated as exploratory data reduction rather than confirmation of distinct latent constructs. Composite scores were calculated as the mean of the retained items of each scale. Items and composites were summarized using means, standard deviations and agreement percentages (item responses of 4 or 5, or composite scores of at least 4). H1 to H7 were examined using Pearson correlations and ordinary least squares multiple regression. Regression coefficients were interpreted as associations rather than causal effects. Multicollinearity was assessed using variance inflation factors (VIFs), and heteroscedasticity-robust (HC3) standard errors were estimated as a sensitivity check. Homogeneity of variance was assessed using median-centered Levene's test.

Group differences were examined using one-way ANOVA, with Tukey HSD for appropriate post-hoc comparisons. When Levene's test indicated unequal variances, Welch's ANOVA was used; Tukey comparisons were not applied to the Welch results. Categories with fewer than ten respondents were excluded from inferential comparisons. Open-ended responses were cleaned of non-substantive responses (e.g., "none", "I don't know") and coded thematically by the researcher through an inductive approach. Frequency counts of themes are reported to indicate salience, not statistical prevalence.

4. Results and Discussion

4.1. Respondent profile

The sample profile is shown in Table 1. Most respondents self-assessed at B2 (66.5%), but 27.9% remained at B1. Only 24.7% had major-related experience or part-time work and 23.6% had an international English certificate, almost exclusively IELTS with band scores of 5.5 to 8.5.

Table 1. *Profile of respondents (N = 849)*

Characteristic	Category	n	%
Gender	Female	586	69.0
	Male	252	29.7
	Other / prefer not to say	11	1.3
Major	Information Technology	184	21.7
	International Studies	101	11.9
	Tourism Management	91	10.7
	Finance and Banking	81	9.5
	Accounting	78	9.2
	Business Administration	75	8.8
	Financial Technology	64	7.5
	Multimedia Communication	64	7.5
	Marketing	55	6.5
	Development Studies	52	6.1
	Other	4	0.5
Self-assessed CEFR level	B1	237	27.9
	B2	565	66.5
	C1	41	4.8
	C2	6	0.7
Grade in most recent ESP course	Below 6.0	69	8.1
	6.0 to below 8.0	598	70.4
	8.0 to below 9.0	177	20.8
	9.0 and above	5	0.6
Major-related experience or part-time work	Yes	210	24.7
	No	639	75.3
International English certificate	Yes	200	23.6
	No	649	76.4

4.2. Reliability and construct validity

All eight subscales showed good internal consistency (Table 2). NH7 ("I feel that learning ESP is too difficult for my current ability") was intended as a reverse-worded indicator of learner readiness, but correlated positively with the other learner items (r = 0.27 to 0.56). Reverse-scoring reduced the learner-scale alpha from 0.905 to 0.754. NH7 was therefore excluded from the composite, leaving a

six-item scale with alpha of 0.919. Its raw distribution is reported descriptively because it carries substantive information about perceived difficulty.

Table 2. *Scale reliability*

Scale	Code	Items retained	Cronbach's alpha	Range of corrected item-total correlations
Teaching staff	GV	10	0.959	0.76 to 0.85
Curriculum	CT	5	0.934	0.78 to 0.86
Learning materials	HL	7	0.951	0.77 to 0.86
Teaching methods	PP	8	0.964	0.84 to 0.88
Learner factors	NH	6 (NH7 dropped)	0.919	0.71 to 0.82
Facilities	CS	5	0.913	0.70 to 0.84
Assessment and feedback	DG	8	0.963	0.83 to 0.88
Perceived quality of ESP education	CL	5	0.952	0.85 to 0.88

Exploratory factor analysis of the 49 independent items showed KMO of 0.986 and significant Bartlett's test (χ^2 = 47,654.9, df = 1176, $p < 0.001$). The first unrotated component explained 64.1% of total variance (eigenvalue 31.4); four components had eigenvalues above 1 (73.6% of variance). The theoretically expected seven-factor varimax solution accounted for 78.0% of variance. Teaching staff, teaching methods, learner, facilities and assessment and feedback dimensions were recovered, with no cross-loadings above 0.50. Curriculum and materials items loaded on a common factor. The seven a priori composites were retained given their high reliability and the practical importance of differentiating program design from resources. This overlap was treated as a caveat. The five outcome items loaded on one factor (KMO = .903; eigenvalue 4.20; 84.1% of variance; loadings .90 to .93).

4.3. Students' assessment of the seven factors and overall quality

All composite means exceeded 3.8 (Table 3). Teaching staff and facilities scored highest (both M = 4.16); learning materials (M = 3.97) and learner factors (M = 3.85) scored lowest. Overall perceived quality was high (M = 4.09, SD = 0.82); 65.7% had a composite score of at least 4.0 and only 3.8% averaged below 3.0.

Table 3. *Descriptive statistics of composite scores (N = 849)*

Factor	Mean	SD	Agreement (%)	Disagreement (%)	Rank
Teaching staff (GV)	4.16	0.80	67.1	4.8	1
Facilities (CS)	4.16	0.82	69.3	5.2	2
Assessment and feedback (DG)	4.09	0.81	64.0	4.6	3
Teaching methods (PP)	4.08	0.84	63.7	5.8	4
Curriculum (CT)	4.06	0.83	61.8	5.7	5
Learning materials (HL)	3.97	0.86	56.7	8.2	6
Learner factors (NH)	3.85	0.83	47.9	8.4	7
Perceived quality of ESP education (CL)	4.09	0.82	65.7	3.8	–

Note: Agreement is the percentage of respondents whose composite score is 4.0 or higher; disagreement is the percentage below 3.0.

The highest-rated items were dominated by teaching staff and facilities (Table 4). The eight lowest-rated items came from learner and materials scales. Workload appropriateness was the

weakest curriculum item (CT4, M = 3.95). On the excluded item NH7, 48.9% agreed that ESP was too difficult for their current ability and 18.1% disagreed, consistent with low self-assessed readiness.

Table 4. *Highest-and lowest-rated items*

Item	Statement (abridged)	Mean	SD	Agreement (%)
Highest-rated				
GV1	Lecturers have solid knowledge of ESP	4.35	0.89	84.1
CS1	Classrooms are fully equipped	4.28	0.92	81.2
GV2	Lecturers understand the professional field	4.24	0.91	80.1
GV4	Lecturers create a positive, communicative climate	4.21	0.90	78.9
CS5	Class size allows practice and interaction	4.21	0.91	78.4
PP4	Students regularly practise English in class	4.18	0.89	78.2
CS2	E-learning resources are accessible	4.17	0.93	76.0
GV6	Lecturers give timely, useful feedback	4.16	0.92	77.0
Lowest-rated				
NH5	I am confident using English for disciplinary content	3.72	1.04	56.1
NH3	I spend time on ESP self-study outside class	3.73	1.01	56.8
NH6	I actively seek ESP materials beyond the textbook	3.74	1.04	57.7
HL6	Supplementary materials (video, podcast, online) are varied and useful	3.87	1.04	64.2
HL5	Materials are lively and attractive	3.88	1.02	64.1
HL3	Materials contain many professional examples	3.89	1.02	64.8
NH2	I have clear motivation for ESP	3.90	0.93	65.0
NH1	My English foundation is sufficient for ESP	3.91	0.96	65.5
CT4	Course workload is reasonable	3.95	0.99	67.3

4.4. Relationships between factors and perceived quality

All seven factors correlated positively with perceived quality (Table 5; all $p < 0.001$), providing bivariate support for H1 to H7. Assessment and feedback had the strongest association ($r = 0.906$). Predictors also correlated highly with one another ($r = 0.65$ to 0.88), especially curriculum with materials and materials with methods (both $r = 0.881$), confirming the overlap identified in the factor analysis.

Table 5. *Pearson correlation matrix of composite scores*

	GV	CT	HL	PP	NH	CS	DG	CL
GV	1							
CT	0.870	1						
HL	0.835	0.881	1					
PP	0.871	0.869	0.881	1				
NH	0.693	0.721	0.724	0.732	1			
CS	0.786	0.744	0.716	0.798	0.653	1		
DG	0.826	0.828	0.827	0.868	0.715	0.816	1	
CL	0.811	0.814	0.812	0.843	0.713	0.778	0.906	1

Note: All coefficients significant at $p < 0.001$ (two-tailed).

The multiple regression model (Table 6) explained 84.3% of the variance in perceived quality (adjusted $R^2 = .841$; $F(7, 841) = 642.68$, $p < .001$). VIFs ranged from 2.42 to 7.68, below 10, although several values suggest considerable shared variance. Assessment and feedback was the strongest predictor ($\beta = 0.614$, $p < 0.001$); learner factors made a small independent contribution ($\beta =$

0.067, p = 0.002). The other coefficients were positive but non-significant at the 5% level. Heteroscedasticity-consistent standard errors did not alter these conclusions. Thus, H7 and H5 were supported in the adjusted model; H1, H2, H3, H4 and H6 were supported only at the bivariate level.

Table 6. *Multiple regression of perceived quality of ESP education on seven factors*

Predictor	B	SE	β	t	p	VIF
Constant	0.134	0.063		2.12	0.034	
Teaching staff (GV)	0.063	0.034	0.062	1.87	0.062	5.78
Curriculum (CT)	0.056	0.034	0.057	1.63	0.104	6.55
Learning materials (HL)	0.063	0.032	0.067	1.95	0.052	6.23
Teaching methods (PP)	0.064	0.037	0.065	1.72	0.085	7.68
Learner factors (NH)	0.066	0.021	0.067	3.14	0.002	2.42
Facilities (CS)	0.043	0.026	0.043	1.67	0.096	3.56
Assessment and feedback (DG)	0.617	0.032	0.614	19.17	< 0.001	5.47

Note: $R^2 = 0.843$; adjusted $R^2 = 0.841$; $F(7, 841) = 642.68$, $p < 0.001$; Durbin-Watson = 1.95.

Assessment and feedback alone accounted for 82.1% of the variance in perceived quality. Adding learner factors increased this to 83.0%, while the other five predictors added 1.3 percentage points. When assessment and feedback was removed, the six remaining factors explained 77.4% of the variance and all became significant ($p < 0.01$). This pattern indicates substantial shared variance; mediation by assessment and feedback remains a possibility requiring formal testing.

4.5. Differences between student groups

Perceived quality did not significantly differ by gender ($F(2, 846) = 0.93$, $p = 0.397$), major ($F(9, 835) = 1.59$, $p = 0.112$), self-assessed CEFR level ($F(2, 840) = 1.46$, $p = 0.233$), grade band ($F(2, 841) = 0.50$, $p = 0.609$), major-related experience (t-test $p = 0.958$) or certificate status ($p = 0.650$).

Learner factors varied by self-assessed CEFR level ($F(2, 840) = 9.98$, $p < 0.001$, $\eta^2 = 0.023$). B1 students reported lower readiness, motivation and autonomy (M = 3.66) than B2 students (M = 3.91; Tukey $p < 0.001$) and C1 students (M = 4.09; $p = 0.005$). Learning material ratings varied by major ($F(9, 835) = 2.07$, $p = 0.030$, $\eta^2 = 0.022$), with Multimedia Communication rating M = 3.63, Marketing M = 3.85, and International Studies M = 4.15. These small effects point to specific Language of the Specialism modules for priority review.

4.6. Students' own ranking of quality determinants and improvement priorities

When asked to identify the single factor with the greatest impact on ESP quality, students selected teaching methods (32.6%), teaching staff (23.8%), learner factors (18.5%), curriculum (14.1%), materials (5.9%), assessment and feedback (2.8%) and facilities (2.2%). The highest share of attribution to the learner was among B1 students (24%), the group with the lowest readiness.

Curriculum (61.0%), teaching methods (60.8%) and learning materials (58.7%) were the improvement areas most frequently selected, followed by teaching staff (36.9%), facilities (30.2%) and assessment and feedback (27.8%).

4.7. Qualitative results

Of the 256 responses on greatest strength, 240 were substantive. Teaching staff were the most common theme (94 mentions), described as knowledgeable, enthusiastic, approachable and experienced in the professional field. Other themes were the practical, discipline-oriented nature of courses (49), comprehensive four-skills coverage and logical B1-to-B2 progression (48), and the full-English environment (31). Materials (12) and facilities (5) were rarely nominated.

Of 224 responses on greatest weakness, 60 (27%) indicated they had none. The remaining 164 were grouped into eight different themes. First, materials (40 mentions) were criticized for typographical and answer-key errors, audio-text mismatches, outdated content, AI-generated illustrations and dense layouts. Second, assessment and score reporting (36) concerned delayed scores, unexplained formats, task-course mismatches and perceived subjective grading. Third, uneven

teaching quality (35) included monotonous delivery, lateness and inattention to weaker students. Other themes were lack of practical application (34), workload and pace (31), difficulty relative to ability (28), skill-specific difficulties (27), and facilities (6), primarily Wi-Fi.

4.8. Discussion

The results indicate a program perceived positively by students, alongside weaknesses in materials and learner readiness. Mixed perceptions of materials relevance and course design have also been reported in Vietnamese ESP assessments (Tran et al., 2019; Chi & Vu, 2024). One persistent problem in the literature is limited subject-matter knowledge among ESP teachers (Hoa & Mai, 2016; Dong, 2025). In contrast, lecturers' ESP knowledge and understanding of the professional field were strengths at ESPD, supported by the qualitative data: teaching staff were identified as the greatest strength more often than any other theme.

Almost half of the students agreed that ESP was too difficult for their current ability and only 56% felt confident using English for disciplinary content, despite two-thirds self-assessing at B2. The combination of a demanding, fast-paced program and imperfect materials may burden the least prepared students, a dynamic also observed in Vietnamese EMI-preparatory contexts (Dang, 2023).

Assessment and feedback as the key determinant:

The predominance of assessment and feedback is consistent with constructive alignment theory (Biggs, 1996), in which assessment defines what students perceive the course to be about and how they judge their learning. The effectiveness of the course as a whole can be perceived by students when there is variation in task types, authenticity of task type, transparency of task type and informative feedback on each task. Additionally, this supports the literature on the importance of high-information feedback (Hattie & Timperley, 2007; Wisniewski et al., 2020), as well as Bui's (2022) findings regarding the differences between how students and teachers view assessment. Although criticisms such as delayed score release, lack of explanation of format, and inadequate feedback exist, there are also high ratings for overall assessment quality which indicate specific areas where improvements can be made.

On the other hand, the weakness of the other factors as independent predictors should not be misinterpreted as unimportance. All six showed strong correlations with perceived quality and became significant when assessment was removed. Good teaching, cohesive curriculum design, and sound materials can enhance students' perceptions of assessment experiences as fair, relevant and formative. Due to the high level of inter-correlation among the factors studied, as well as the fact that all variables were measured using self-report methods, the researcher cannot determine whether any of the factors mediated or caused the relationship between student perception of the quality of assessment and the factors studied.

Learner readiness:

Learner factors were the only other significant independent predictor and the lowest-rated dimension. B1 students reported weaker readiness, motivation and autonomy than B2 and C1 students. This supports the ESP principle that courses should address learners' present situation as well as the target situation (Basturkmen, 2010). Interventions that build students' foundation, autonomy and confidence should be central to program quality.

Localized weaknesses and differential attributions:

The absence of any large disparities among student perceptions of module quality indicate similar levels of satisfaction with their courses. However, differences in materials ratings by major and textbook complaints concerning Language of the Specialism modules support targeted improvements.

Students considered teaching methods as the most important factor when evaluating module quality, whereas assessment and feedback were found to be more predictive than other variables in the regression analysis. These measures capture different judgments: explicit attribution and

conditional statistical association with overall ratings. The solutions below address both students' stated priorities and the associations identified in the model.

4.9. Quality improvement solutions

Formative assessment and feedback:

Because assessment and feedback is the dominant predictor of perceived quality and the area with the most specific complaints, the department should treat it as a managed system: (i) publish, at the start of each module, the assessment plan, formats, weightings and marking rubrics, and walk students through them in class; (ii) set a binding standard for the release of mid-term and continuous assessment scores (for example, within ten working days) with itemized feedback, and make compliance a criterion in lecturers' performance review; (iii) require at least one piece of written, criterion-referenced feedback per student per module in the productive skills courses, using shared rubrics referenced to CEFR descriptors; (iv) audit each assessment task for alignment with the module's stated outcomes, so that a Critical Reading module is not assessed primarily through an unrelated presentation; (v) diversify assessment towards authentic disciplinary tasks (professional emails, reports, case presentations, simulated meetings); and (vi) moderate grading across parallel classes through calibration sessions and sample double-marking. These measures are a direct response to the evidence that high-information, timely feedback supports learning (Wisniewski et al., 2020) and that alignment of outcomes, activities and assessment is a marker of quality (Biggs, 1996).

Tiered learner support:

Learner factors were the least-rated dimension, an independent predictor of quality and significantly weaker among B1 entrants. The department should therefore: (i) administer a diagnostic placement test at entry and use it to stream students into differentiated pathways or to assign B1 entrants to smaller support sections in the first semester; (ii) offer a structured self-study programme (weekly vocabulary and listening quotas, online practice with answer keys, learning logs) with formal recognition, since self-study time and autonomy were among the weakest learner items; (iii) establish a peer-tutoring scheme pairing C1/B2 students with B1 students, drawing on the 23.6% of students who hold international English certificates; (iv) train lecturers in learning-strategy instruction and scaffolding for weaker students; and (v) make the rationale of ESP visible from week one by connecting each module to second-year EMI courses and professional tasks, building on the perceived usefulness that already scores comparatively well (NH4, M = 4.09).

Learning materials:

Materials was the second lowest rated factor, the most common theme in weaknesses, and the dimension that varied most across majors. The department should: (i) commission an immediate errata review of the most frequently named Language of the Specialism textbooks, correcting audio-text mismatches, answer-key errors and typographical mistakes before the next print run; (ii) adopt a materials quality-assurance protocol with peer review, student piloting and an annual update cycle; (iii) enrich textbooks with glossaries of discipline-specific terminology, worked examples and case studies drawn from the Vietnamese and regional professional context, responding to the low ratings for professional illustration (HL3) and attractiveness (HL5); (iv) expand supplementary digital resources on the learning management system, the weakest materials item (HL6), and provide every textbook in both print and digital form; and (v) involve lecturers from the disciplinary faculties in reviewing content accuracy, a form of interdisciplinary collaboration recommended across the Vietnamese ESP literature (Dong, 2025; Vo et al., 2026).

Practical discipline-based teaching:

Teaching methods were most often named by students as the decisive factor and insufficient practical application as a leading weakness. The department should: (i) increase the percentage of class time devoted to simulations, role-plays, case discussions and project work based on professional situations in each specialism, following ESP's principle of using the methodology of the discipline served (Dudley-Evans & St John, 1998, as cited in Hoa & Mai, 2016); (ii) include mini-projects

requiring authentic disciplinary outputs (a marketing pitch, a financial summary, a tourism itinerary, a software specification) which can also be used for authentic assessment; (iii) rebalance the workload and coordinate deadlines across parallel modules using a common semester calendar, addressing the workload item (CT4) and the many comments on pace; (iv) review the B1-to-B2 progression in Reading and Listening, where students perceived little difference in challenge between semesters, and the design of the Debate module; and (v) systematize the use of learning technology so that digital resources are used consistently across classes.

Consistency of teaching quality:

The highest-rated factor was the teaching staff, although students contrasted the majority of excellent lecturers with a minority who were described as monotonous, late or inattentive. Thus, quality improvement is a matter of consistency and not of average improvement. The department should: (i) institutionalize peer observation and mentoring where highly rated lecturers support colleagues, with observation criteria derived from the GV items of this study; (ii) provide professional development in interactive methodology, differentiated instruction, feedback literacy and educational technology; (iii) deepen lecturers' disciplinary knowledge through short placements, guest lectures and co-teaching with the disciplinary faculties; (iv) adopt an English-only classroom policy with agreed, principled use of Vietnamese for terminology clarification; and (v) use student evaluation data at module and lecturer level, following up promptly where ratings are consistently low.

Targeted facility upgrades:

Based on the highly rated facilities (Wi-Fi access, electrical outlets, functioning projectors), there is a need for the university to make targeted facility improvements. These include: providing reliable Wi-Fi, making sure all sockets work, having working projectors, purchasing good quality audio equipment for students to use when listening and speaking, and continuing to focus on keeping class sizes small.

5. Conclusion

Based upon849 first- year students enrolled in10 English-medium majors, this study indicated that ESP education at Hanoi University was positively evaluated by students. While teaching staff and facilities were viewed as the two most effective aspects, student readiness and learning materials were considered the least effective. Seven factors accounted for 84.3 percent of the variance in perceived quality of ESP education. Assessment and feedback were found to be the most influential factor predicting perceived quality, while learner factors were second. Differences in rating levels for learning materials existed among majors; therefore, targeted reviews of materials used in specializations are needed. The study contributes quantitative evidence on assessment and feedback and highlights the learner as an explicit quality factor in preparatory ESP. It also translates the findings into recommendations for formative assessment, tiered learner support, materials revision, discipline-embedded teaching, lecturer development and targeted facility upgrades. A comprehensive plan for improving ESP education should include who will take responsibility for implementing changes, when those changes will occur, and how progress will be measured. Annual surveys can be used to track progress.

The data in this study were collected cross-sectionally and self-reported; thus, the data represent students' perceptions of ESP education, not actual learning outcomes. Future studies should examine the relationship between students' perceptions of ESP education and their performance in second-year EMI courses. Due to the high correlation between variables and because curriculum and materials were combined during factor analysis, caution must be exercised when interpreting the coefficients and further validation of constructs is required. Additionally, structural equation modeling may allow researchers to test whether assessment serves as a mediator in students' perceptions of ESP education. The negatively worded learner variable needs to be revised and pretested. Finally, perceptions of

lecturers, faculty members in disciplines, and former graduates may provide a more complete picture of what constitutes quality in ESP education.

REFERENCES

1. Abdullah, F. (2006), The development of HEdPERF: A new measuring instrument of service quality for the higher education sector. *International Journal of Consumer Studies, 30*(6), 569–581. https://doi.org/10.1111/j.1470-6431.2005.00480.x.

2. Bahadur, G. C. S., Gurung, S. K., Poudel, R. L., Yadav, U. S., Bhattacharjee, A., & Dhungana, B. R. (2024), The effect of higher education service quality on satisfaction among business students. *Cogent Education, 11*(1), 2393521. https://doi.org/10.1080/2331186X.2024.2393521.

3. Basturkmen, H. (2010), *Developing courses in English for specific purposes*. Palgrave Macmillan. https://doi.org/10.1057/9780230290518.

4. Biggs, J. (1996), Enhancing teaching through constructive alignment. *Higher Education,* 32(3), 347–364. https://doi.org/10.1007/BF00138871.

5. Bui, H. P. (2022), Students' and teachers' perceptions of effective ESP teaching. *Heliyon, 8*(9), e10628. https://doi.org/10.1016/j.heliyon.2022.e10628.

6. Chi, D. N., & Vu, N. T. (2024), English for specific purposes courses and Vietnamese graduates' employability. In T. L. H. Nghia, L. T. Tran, & M. T. Ngo (Eds.), *English language education for graduate employability in Vietnam*. Springer. https://doi.org/10.1007/978-981-99-4338-8_12.

7. Dang, N. A. (2023), English medium instruction practices in Vietnamese universities: Institutional, practitioner and student perspectives. *VNU Journal of Foreign Studies, 39*(6), 175-180. https://jfs.ulis.vnu.edu.vn/index.php/fs/article/download/5155/4487/12206.

8. Dong, T. K. D. (2025), Challenges in teaching English for specific purposes at universities in Ho Chi Minh City. *Tạp chí Thiết bị Giáo dục* (Journal of Educational Equipment), 338(2), 69-71. https://vjol.info.vn/index.php/tctbgd/article/view/124187.

9. Dou, A. Q., Chan, S. H., & Win, M. T. (2023), Changing visions in ESP development and teaching: Past, present, and future vistas. *Frontiers in Psychology, 14*, 1140659. https://doi.org/10.3389/fpsyg.2023.1140659.

10. Hattie, J., & Timperley, H. (2007), The power of feedback. *Review of Educational Research, 77*(1), 81–112. https://doi.org/10.3102/003465430298487.

11. Hoa, N. T. T., & Mai, P. T. T. (2016), Difficulties in teaching English for specific purposes: Empirical study at Vietnam universities. *Higher Education Studies, 6*(2), 154–161. https://doi.org/10.5539/hes.v6n2p154.

12. Le, D. M. T., Tran, T. H. L., & Nguyen, Q. T. (2021), Service quality in higher education: Applying HEdPERF scale in Vietnamese universities. *Ho Chi Minh City Open University Journal of Science – Social Sciences, 11*(1), 101-115. https://doi.org/10.46223/HCMCOUJS.soci.en.11.1.1905.2021.

13. Tran, T. D., Hoang, N. T., Le, T. H. G., Vu, T. N., Bui, G., & Cao, T. H. H. (2019), An evaluation of English for specific purposes courses at a Vietnamese university. *VNU Journal of Foreign Studies, 35*(5), 41–53. https://js.vnu.edu.vn/FS/article/view/4419.

14. Vo, T., Bui, N., Duong, U., Ha, L., Kieu, T., Nguyen, N., & Nguyen, T. (2026), Evaluating ESP programs for non-English majors: A case study from a Vietnamese higher education institution. *Educational Process: International Journal,* 23, e2026076. https://doi.org/10.22521/edupij.2026.23.76.

15. Wisniewski, B., Zierer, K., & Hattie, J. (2020), The power of feedback revisited: A meta-analysis of educational feedback research. *Frontiers in Psychology,* 10, 3087. https://doi.org/10.3389/fpsyg.2019.03087.

Đào tạo tiếng Anh chuyên ngành: các yếu tố liên quan và giải pháp nâng cao chất lượng tại Khoa tiếng Anh chuyên ngành, Trường Đại học Hà Nội

Tóm tắt: Nghiên cứu tìm hiểu các yếu tố liên quan đến chất lượng đào tạo tiếng Anh chuyên ngành thông qua khảo sát 849 sinh viên năm nhất thuộc 10 ngành tại Trường Đại học Hà Nội. Dữ liệu được phân tích bằng thống kê mô tả, hồi quy bội và phân tích chủ đề. Điểm đánh giá chất lượng đào tạo trung bình đạt 4,09/5; yếu tố người học (3,85) và học liệu (3,97) có điểm thấp nhất. Mô hình hồi quy gồm bảy yếu tố giải thích được 84,3% phương sai của chất lượng cảm nhận. Đánh giá và phản hồi có hệ số chuẩn hóa lớn nhất ($\beta = 0,614$; $p < 0,001$); yếu tố người học có hệ số thấp hơn nhưng vẫn có ý nghĩa thống kê ($\beta = 0,067$; $p = 0,002$). Các ý kiến phản hồi cho thấy học liệu còn sai sót, việc công bố điểm chậm và cơ hội thực hành còn hạn chế. Từ đó, nghiên cứu đề xuất tăng cường phản hồi, hỗ trợ học tập phù hợp với trình độ, rà soát học liệu và mở rộng thực hành gắn với chuyên ngành.

Từ khóa: Tiếng Anh chuyên ngành; chất lượng đào tạo; cảm nhận của sinh viên; kiểm tra, đánh giá và phản hồi; giảng dạy bằng tiếng Anh.