ACADEMIC JOURNAL PERSPECTIVE: Education, Language, and Literature
https://jurnal.ugj.ac.id/index.php/Perspective/index
p-ISSN 2354-7340
e-ISSN 2621-6981

PERCEPTION AND USABILITY EVALUATION OF AN ENGLISH FOR ENGINEERING E-BOOK IN THE ROAD TRANSPORTATION SYSTEM ENGINEERING STUDY PROGRAM

Dani F. Brilianti[1*], Nurul Fitriani[2], Aat Eska Fahmadi[3], Abdul Rokhim[4]
Politeknik Keselamatan Transportasi Jalan, Tegal, Indonesia[1,2,3]
Politeknik Transportasi Darat, Bali, Indonesia[4]
d_fiabrilianti@pktj.ac.id*

Abstract
English for Engineering (EFE) materials must be both contextually relevant and easy to use. This study evaluated cadets' perceptions of the usability of an EFE e-book (version 1.0) that had been developed from a needs analysis, after it was used in eight learning sessions in the Road Transportation System Engineering (RSTJ) study program of Politeknik Keselamatan Transportasi Jalan, Tegal, Indonesia. A quantitative descriptive design was applied. Data were collected from 105 third-semester cadets (88.2% of the population of 119) using the ten-item System Usability Scale (SUS) and a short questionnaire on respondent characteristics. The data were analyzed with descriptive statistics, a one-sample *t*-test against the benchmark of 68, item-level analysis, and the Learnable and Usable subscales of the SUS. The mean SUS score was 82.4 (SD = 6.9; 95% CI [81.1, 83.7]), which falls in the "Excellent" band of the working categories adopted in this study (80–100) and is above the benchmark of 68, $t(104) = 21.26$, $p < .001$, $d = 2.08$ (the large d partly reflects the narrow spread of scores). All 105 cadets scored 68 or higher: 73.3% scored in the Excellent band and 26.7% in the Good band. The strongest items concerned low complexity, ease of use, and the absence of cumbersome operation, whereas the weakest items concerned the need to learn many things before starting (Q10) and the need for technical support (Q4). Descriptively and tentatively, the Learnable subscale (75.1) was lower than the Usable subscale (84.2), although this two-factor split is of uncertain stability and the Learnable subscale has only two items. Internal consistency of the SUS in this restricted-range sample was low ($\alpha = 0.41$) and all scores lay between 70.0 and 92.5. The survey was completed in class by cadets taught by the lecturers who developed the e-book, so social desirability may have inflated the ratings. In this cohort, as rated in a post-use survey, the e-book was perceived very favorably. The findings suggest that onboarding and embedded help are the relative priorities for version 2.0, a conclusion that rests on neutral rather than unfavorable answers and needs confirmation with qualitative evidence.
Keywords: Usability, e-book, English for Engineering, system usability scale, road transportation system engineering

DOI: http://dx.doi.org/10.33603/perspective.v14i2.13015

INTRODUCTION

English competence in engineering has become a fundamental prerequisite for graduates who want to compete in an increasingly integrated global labor market. For cadets of the Road Transportation System Engineering (RSTJ) study program, mastery of English for Specific Purposes (ESP), and especially English for Engineering (EFE), is not an optional extra but an essential need, because international standards for road design, road safety regulation, and Intelligent Transportation Systems (ITS) are largely documented and communicated in English (Sasabone et al., 2024). Cadets are expected to read manuals such as the Highway Capacity Manual, analyze crash investigation reports, understand geometric design standards, and communicate in traffic engineering and transportation planning contexts (Nateghian, 2024).

Although the importance of ESP is widely acknowledged in engineering curricula, the reality in many classrooms is a gap between the materials taught and the specific needs of the learners (Mohamad Saleh & Mohammed, 2024). Failures in ESP instruction are often attributed to a one-size-fits-all approach that overlooks the distinctive features of a discipline (Efendi et al., 2025; Nateghian, 2024; Ramadhani et al., 2023). ESP is defined by its focus on learners' needs, the use of the methodology and activities of the discipline, and language appropriate to those activities (Dudley-Evans & St John, 1998; Hutchinson & Waters, 1987), and needs analysis is therefore the foundation of material development. In the RSTJ program, this principle was applied in the first phase of a research roadmap, in which a needs analysis was conducted and an EFE e-book version 1.0 was developed (Brilianti et al., 2026).

Developing materials from a needs analysis, however, is not sufficient to ensure that the resulting product can be used easily and effectively by its intended users (Rubin & Chisnell, 2008). In the digital era, usability, which refers to the ease, efficiency, and satisfaction with which users interact with a learning product, is an important determinant of whether learners can use a learning product effectively and of their intention to adopt it (Davis, 1989; ISO, 2018). E-books offer accessibility, interactivity, and personalization of content (Wahyuni & Rosana, 2025), and interactive e-books have been reported to support critical thinking and to bridge out-of-class and in-class learning (Marisda et al., 2025). Navigation problems and unfamiliar interfaces have been reported as barriers to students' use of e-books (Hernon et al., 2007). Usability evaluation is therefore a necessary stage in the development cycle of digital learning products (Rubin & Chisnell, 2008; Zaharias & Poylymenakou, 2009).

Usability is commonly defined as the extent to which a product can be used to achieve specified goals effectively, efficiently, and satisfactorily in a specified context of use (ISO, 2018). The concept of usability has been decomposed into five components: learnability, efficiency, memorability, errors, and satisfaction (Alkadrie et al., 2025). A widely used standardized instrument, the System Usability Scale (Brooke, 1996), summarizes perceived usability in a single score from 0 to 100 that can be interpreted against established benchmarks (Lewis & Sauro, 2009, 2017) and has been used to evaluate online learning platforms (Alkadrie et al., 2025). The SUS can additionally be read through two subscales proposed by Lewis and Sauro (Lewis & Sauro, 2009): *Learnable* (items 4 and 10), which was interpreted as reflecting how much support and prior learning users need, and *Usable* (the remaining eight items), which was interpreted as reflecting the ease and coherence of operation. However, on the next research it has been re-analyzed more than 9,000 SUS questionnaires and found that this two-factor structure was not consistently replicated; a split by item tone (positively versus negatively worded items) fitted better, and they no longer routinely compute the subscales (Lewis & Sauro, 2017). Because the Learnable subscale

DOI: http://dx.doi.org/10.33603/perspective.v14i2.13015

contains only two items and the internal consistency of the SUS in the present sample was low, the comparison of the two subscales in this study is treated as descriptive and tentative. In this study, usability is operationalized through the overall SUS score as the the primary indicator and these two subscales as a secondary, exploratory indicator.

This study draws on two complementary frameworks. The Technology Acceptance Model (TAM) explains that perceived ease of use and perceived usefulness shape attitude, intention, and actual use of a technology (Davis, 1989). In this study, the usability of the e-book corresponds closely to perceived ease of use. User-centered design (UCD) places users at the center of every stage of design and treats usability testing and user feedback as the basis for iteration (Rubin & Chisnell, 2008). The present evaluation is the evaluation phase of this cycle: cadets' perceptions are not an end in themselves but input for the continuous improvement of the e-book. Empirical studies support this view. From the previous research students responded positively to the ease of use of an e-book but made limited use of its interactive features because of low familiarity (Lim & Hew, 2014), and another researcher identified learner satisfaction as a critical factor in successful e-learning (Sun et al., 2008).

Research on EFE has mostly focused on needs analysis and material development (Changpueng & Pattanapichet, 2023) or on learning outcomes (Buroidah et al., 2023) and relatively little attention has been paid to how learners perceive the usability of the resulting digital materials. Three gaps can be identified. First, studies of users' perceptions of e-book usability in technical ESP, particularly EFE, remain limited. Second, most studies stop at the development or outcome stage without a usability evaluation as part of the user-centered design cycle. Third, in the literature the authors were able to identify (no systematic literature search was conducted for this study), studies that evaluate an ESP e-book for road transportation cadets with a standardized instrument such as the SUS were not found.

This study is the second phase of a research roadmap that began with an exploratory study of engineering students' perceptions of ESP, continued with needs analysis and e-book development in 2025, and will continue with the revision and extension of the e-book in 2027. Its aims are: (1) to analyze cadets' overall perception of the usability of the EFE e-book using the System Usability Scale; (2) to identify the relatively strongest and weakest aspects of the e-book at the item and subscale level; and (3) to formulate recommendations for improving the usability of the e-book in version 2.0. The study measures perceived usability only and does not assess learning outcomes. Theoretically, it extends research on technology-enhanced language learning and human–computer interaction into technical ESP for road transportation. Practically, it provides evidence-based input for the revision of the e-book and a replicable evaluation procedure for other vocational programs.

METHOD

Research Design

This study employed a quantitative descriptive design in the form of a cross-sectional survey. It is an evaluation study in the second phase of the research roadmap, conducted in real learning conditions rather than in a laboratory. Perceived usability was measured once, after cadets had used the e-book in eight sessions, with the System Usability Scale as the main instrument.

DOI: http://dx.doi.org/10.33603/perspective.v14i2.13015

E-Book Under Evaluation

The evaluated product is version 1.0 of the *English for Engineering* e-book developed in the first phase from a needs analysis of RSTJ cadets. The e-book covers technical reading in highway engineering, professional correspondence, technical report writing, and terminology related to traffic safety and transportation planning, and is distributed as an interactive PDF with a hyperlinked table of contents, a clickable glossary, and embedded multimedia links. The e-book was used in eight sessions of the EFE course in the third semester of the 2026/2027 academic year, both in class and for independent study.

Participants and Setting

The study was conducted at Politeknik Keselamatan Transportasi Jalan, Tegal, Indonesia. The population consisted of all 119 active third-semester RSTJ cadets who used the e-book in the EFE course. A total of 105 cadets provided valid responses (a response rate of 88.2%); the remaining 14 did not participate because of illness, approved leave, or academic leave. Three response-rate targets were set: ideal 119 (100%), realistic 107 (90%), and minimum 101 (85%). The achieved rate exceeded the minimum target and the minimum sample size of 92 calculated with Cochran's formula, but fell slightly short of the realistic target. The study population consisted of 119 third-semester cadets enrolled in the Road Transportation System Engineering (RSTJ) study program. A total of 105 cadets provided valid responses, resulting in a response rate of 88.2%. This figure exceeded the minimum target of 101 respondents (85%) and the statistically required minimum sample size of 92 respondents, although it fell slightly below the realistic target of 107 respondents (90%). The remaining 14 cadets did not participate due to illness, officially approved leave, or academic leaves.

Respondents had to fulfil four inclusion criteria: (1) being an active third-semester RSTJ cadet in the 2026/2027 academic year; (2) having attended at least six of the eight class meetings in which the e-book was used (75% attendance based on the signs on the attendance list); (3) having active access to the e-book during the study period; and (4) willing to participate as the respondent. Cadets on academic leave, those with attendance below 75%, those who could not access the e-book because of continuing technical problems, and those who refused to participate were excluded. "Meeting" refers to class meeting attended in which the e-book was used, whereas "frequency of access" (Table 3) refers to self-reported access to the e-book for independent study outside class; the two are therefore not contradictory.

Research Instruments

The questionnaire was delivered online through Google Forms and consisted of two parts. The first part was the ten-item System Usability Scale (Brooke, 1996), in which respondents rated each statement on a five-point scale from 1 (strongly disagree) to 5 (strongly agree). Five items are positively worded (odd items) and five are negatively worded (even items); the items and their scoring rules are listed in Table 1. The second part collected respondent characteristics, they are gender, the main device used to open the e-book, the frequency of access, the number of sessions in which the e-book was used (6–8), and the connectivity used to access the e-book. These data were collected through the same Google Form as the SUS responses but were not linked to individual SUS scores using respondent codes. The English wording of the SUS items was used, as in the original instrument; using English considered relatively easy to understand; no validated Indonesian version was used. Items

DOI: http://dx.doi.org/10.33603/perspective.v14i2.13015

such as “cumbersome” and the reverse-worded statements may be misread by non-native readers, which is acknowledged as a limitation.

Table 1: Items of the System Usability Scale and scoring rules

No.	Item	Wording	Adjusted score	Subscale
Q1	I think that I would like to use this system frequently.	Positive	response − 1	Usable
Q2	I found the system unnecessarily complex.	Negative (R)	5 − response	Usable
Q3	I thought the system was easy to use.	Positive	response − 1	Usable
Q4	I think that I would need the support of a technical person to be able to use this system.	Negative (R)	5 − response	Learnable
Q5	I found the various functions in this system were well integrated.	Positive	response − 1	Usable
Q6	I thought there was too much inconsistency in this system.	Negative (R)	5 − response	Usable
Q7	I would imagine that most people would learn to use this system very quickly.	Positive	response − 1	Usable
Q8	I found the system very cumbersome to use.	Negative (R)	5 − response	Usable
Q9	I felt very confident using the system.	Positive	response − 1	Usable
Q10	I needed to learn a lot of things before I could get going with this system.	Negative (R)	5 − response	Learnable

Note: (R) = reverse-scored. Subscales follow (Lewis & Sauro, 2009); their stability was questioned by (Lewis & Sauro, 2017). Source: (Brooke, 1996).

Data Collection Procedure

The academic semester commenced on 6th July 2026. The first meeting was held on 6th July 2026 and the eighth on 15th September 2026, and the questionnaire was open from 16 September to 1 October 2026. After the eight meetings, the questionnaire link was distributed to the cadets in the population by the lecturer as the researcher. Participation was voluntary, informed consent was obtained, and responses were anonymized using respondent codes (R1–R105). To reduce social desirability, cadets were told that responses were anonymous, that participation or refusal would not affect grades or cadet standing, and that lecturers would only see aggregated results after grading. The authors developed the e-book and teach the cadets.

Data Analysis

The data were analyzed in several steps. First, the quality of the data was checked for incomplete responses, out-of-range values, duplicate identifiers, and straight-lining, and the internal consistency of the SUS was examined with Cronbach’s alpha, with $\alpha \geq .70$ considered acceptable (Nunnally, 1978); item-total correlations and alpha for the positively and negatively worded items separately were also examined. Second, the SUS score of each respondent was computed following (Brooke, 1996) that states the score of odd items is the response minus 1, the score of even items is 5 minus the response, the sum of the ten adjusted scores is multiplied by 2.5, and the result ranges from 0 to 100. Third, descriptive statistics (mean, standard deviation, median, range, 95% confidence interval) were calculated and the scores were classified with the categories in Table 2. Fourth, the mean score was tested

DOI: http://dx.doi.org/10.33603/perspective.v14i2.13015

against the benchmark of 68 with a one-sample t-test; this benchmark is consistent with the mean of about 70 reported across studies (Bangor et al., 2008; Lewis & Sauro, 2017), and the effect size was expressed as Cohen's d (Cohen, 1988). Because n = 105 and the benchmark is a general reference value, the confidence interval and the share of respondents above 68 are treated as the main indicators, and the *t*-test and *d* as supplementary; *d* is inflated when the standard deviation is small. Fifth, item-level results were summarized as raw means, adjusted means (0–4), and the percentage of favorable responses (4–5 on positive items and 1–2 on negative items), and the Learnable and Usable subscale scores were computed following (Lewis & Sauro, 2017): the Learnable score is the sum of the adjusted means of Q4 and Q10 multiplied by 12.5, and the Usable score is the sum of the adjusted means of the other eight items multiplied by 3.125. Finally, the respondent characteristics were summarized as frequencies and percentages.

Table 2: Interpretation criteria for SUS scores

Category	SUS score	Interpretation
Excellent	80–100	Very favorable perceived usability (Excellent band)
Good	68–79	Above the acceptability benchmark
OK	51–67	Below benchmark; improvement needed
Poor	0–50	Serious usability problems

Source: categories are the authors' working categories for this study, not a published standard. The benchmark of 68 follows (Lewis & Sauro, 2017); the 80 cut-off approximates the lower end of the A range of their curved grading scale (80.8), and the labels were informed by the adjective ratings of (Bangor et al., 2008)). Scoring follows (Brooke, 1996).

RESULTS AND DISCUSSIONS

This section reports the findings of the second-phase evaluation of the *English for Engineering* (EFE) e-book, version 1.0, after eight learning sessions in the Road Transportation System Engineering (RSTJ) program. The results follow the three research aims: the overall perceived usability, the strongest and weakest aspects at item and subscale level, and the resulting priorities for improvement.

Respondent Profile and Response Rate

The population consisted of 119 third-semester RSTJ cadets. A total of 105 cadets provided valid responses, a response rate of 88.2% of the population. All 105 responses were complete, within the 1–5 range, and free from duplicate identifiers. Table 3 summarizes the respondents' characteristics.

DOI: http://dx.doi.org/10.33603/perspective.v14i2.13015

Table 3: Characteristics of respondents (n = 105)

Characteristic	Category	n (%)
Gender		
	Male	59 (56.2)
	Female	46 (43.8)
Main device used to open the e-book		
	Smartphone	15 (14.3)
	Laptop	90 (85.7)
	Tablet	0 (0.0)
Frequency of access		
	Every session	85 (81.0)
	Once a month	15 (14.3)
	Rarely	5 (4.8)
Sessions in which the e-book was used (mean = 7.52 of 8)		
	8 sessions	65 (61.9)
	7 sessions	30 (28.6)
	6 sessions	10 (9.5)
Connectivity when accessing the e-book		
	Campus Wi-Fi	80 (76.2)
	Mobile data	10 (9.5)
	Offline (downloaded file)	15 (14.3)

Source: primary data, processed by the authors.

The cohort consisted of 59 male cadets (56.2%) and 46 female cadets (43.8%). Most cadets opened the e-book on a laptop (90 cadets, 85.7%), while 15 cadets (14.3%) used a smartphone and none used a tablet (Figure 1). This distribution is relevant for interpreting the results because the e-book is an interactive PDF, a format that is designed primarily for larger screens. Regarding the frequency of access (distinct from the number of class sessions attended; see Method), 85 cadets (81.0%) reported accessing the e-book in every session, 15 (14.3%) once a month, and 5 (4.8%) rarely. Cadets used the e-book in 7.52 of the eight meetings on average, and 65 cadets (61.9%) used it in all eight sessions. Most cadets accessed the e-book through the campus Wi-Fi (80 cadets, 76.2%), 15 cadets (14.3%) used a downloaded offline copy, and 10 cadets (9.5%) relied on mobile data (Figure 2).

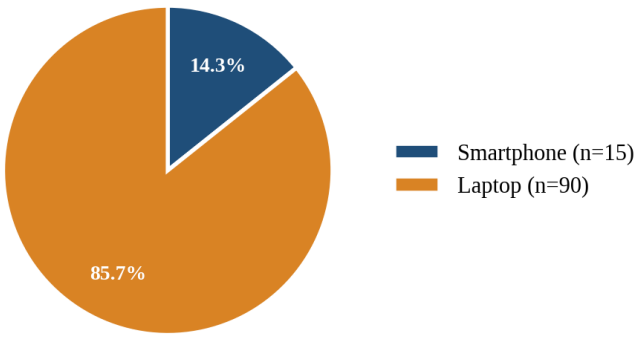


Figure 1: Distribution of the main device used by cadets to open the e-book (n = 105)

DOI: http://dx.doi.org/10.33603/perspective.v14i2.13015

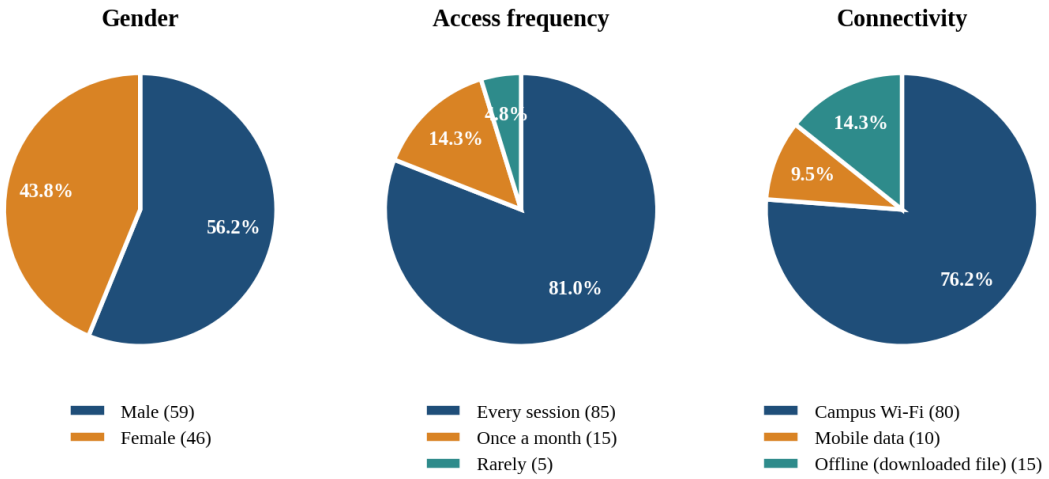


Figure 2: Gender, frequency of access, and connectivity of respondents (n = 105)

Data Quality and Reliability

The data quality checks showed no incomplete responses, no values outside the 1–5 range, no duplicate respondent identifiers, and no straight-lining (cadets giving the same answer to all ten items). The internal consistency of the SUS in this sample was low, Cronbach's α = 0.41, which is below the conventional threshold of .70 (Nunnally, 1978) and well below the high values usually reported for the SUS (Bangor et al., 2008). This result should be interpreted in light of the characteristics of the data: scores are concentrated in a narrow range (70.0–92.5), which restricts the variance on which alpha depends, the scale has only ten items, and positively and negatively worded items are mixed and the English wording of the items may not have been equally clear to all cadets. The SUS score is a composite index and was therefore still computed and interpreted following the standard procedure, but the low alpha is acknowledged as a limitation in the Discussion.

Overall System Usability Scale (SUS) Score

The SUS score was calculated for each cadet following (Brooke, 1996). The mean SUS score was **82.4** (SD = 6.9; median = 82.5), with individual scores ranging from 70.0 to 92.5 (Table 4). The 95% confidence interval of the mean was [81.1, 83.7], which lies entirely within the *Excellent* band (80–100) of Table 2. A one-sample *t*-test showed that the mean was significantly higher than the benchmark of 68, $t(104) = 21.26$, $p < .001$, with a standardized difference of Cohen's $d = 2.08$ (Cohen, 1988). This value is large by conventional benchmarks but is inflated by the small standard deviation of the scores (restricted range), so the confidence interval and the fact that every respondent scored above 68 are more informative than the effect size. On the Sauro–Lewis curved grading scale (Lewis & Sauro, 2017), a mean of 82.4 corresponds to grade A (80.8–84.0).

DOI: http://dx.doi.org/10.33603/perspective.v14i2.13015

Table 4: Overall SUS results (n = 105)

Statistic	Value	Note
Number of valid respondents	105	—
Mean SUS score	82.40	Excellent band, table 2.
Standard deviation	6.94	—
Median	82.5	—
Minimum – maximum	70.0 – 92.5	—
95% confidence interval	[81.06, 83.75]	—
Benchmark	68	(Lewis & Sauro, 2009)
One-sample t (df = 104)	21.26	p < .001
Cohen's d	2.08	Lage: inflated by small SD
Respondents ≥ 68	105 (100.0%)	—
Cronbach's α	0.41	Below .70
Learnable subscale (Q4, Q10)	75.1	Good band
Usable subscale (other eight items)	84.2	Excellent band

Source: primary data, processed by the authors.

Figure 3 shows the distribution of the individual scores. No cadet scored below 70, and the scores cluster between 80 and 95 (77 cadets), with the highest frequency in the 80–85 interval (34 cadets), followed by the 90–95 interval (23 cadets). Categorically (Figure 4 and Table 5), 77 cadets (73.3%) rated the e-book as *Excellent* and 28 cadets (26.7%) as *Good*, while no cadet fell in the *OK* or *Poor* categories. In total, all 105 cadets (100.0%) gave a score at or above the benchmark of 68. This pattern indicates a broadly shared, favorable experience of using the e-book rather than an average driven by a few very high scores.

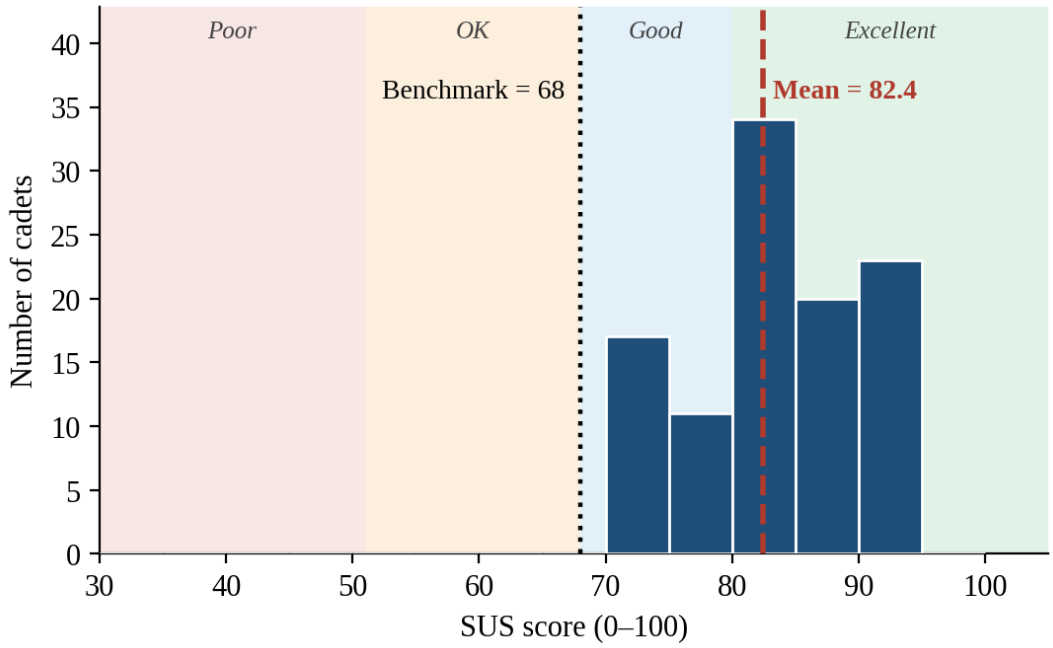


Figure 3: Distribution of individual SUS scores (n = 105); shaded bands show the interpretation categories

DOI: http://dx.doi.org/10.33603/perspective.v14i2.13015

Table 5: Distribution of SUS categories (n = 105)

Category	SUS range	n	Percentage (%)
Excellent	80–100	77	73.3
Good	68–79	28	26.7
OK (needs improvement)	51–67	0	0.0
Poor (serious problems)	0–50	0	0.0
Total		**105**	**100.0**

Source: primary data, processed by the authors.

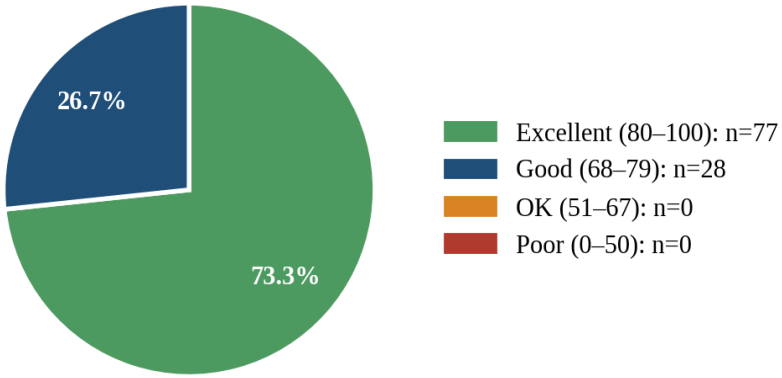


Figure 4: Proportion of cadets in each SUS category (n = 105)

Item-Level Results

Item-level results are presented in Table 6 and Figures 5 and 6. All ten items were rated favorably, with adjusted means between 2.99 and 3.48 on the 0–4 scale, well above the neutral midpoint of 2.0. The highest adjusted mean was obtained by Q2, the e-book was not unnecessarily complex (3.48; 91.4% favorable), followed by Q3, ease of use (3.43; 91.4%), and Q8, not cumbersome to use (3.43; 92.4%). Q1, the wish to use the e-book frequently, obtained a favorable response from every cadet (100.0%): 61 cadets agreed and 44 strongly agreed. Integration of functions (Q5; 3.41; 91.4%) and confidence in using the e-book (Q9; 3.30; 89.5%) were also rated favorably.

DOI: http://dx.doi.org/10.33603/perspective.v14i2.13015

Table 6: Item-level results of the System Usability Scale (n = 105)

No.	Item (short form)	Mean (raw)	SD	Adjusted (0–4)	Favorable (%)	Rank
Q1	Would like to use frequently	4.42	0.50	3.42	100.0	4
Q2	Not unnecessarily complex (R)	1.52	0.65	3.48	91.4	1
Q3	Easy to use	4.43	0.65	3.43	91.4	2
Q4	Need technical support (R)	1.98	0.72	3.02	75.2	9
Q5	Functions well integrated	4.41	0.65	3.41	91.4	5
Q6	Too much inconsistency (R)	1.75	0.90	3.25	77.1	7
Q7	Most people learn quickly	4.25	0.73	3.25	82.9	7
Q8	Very cumbersome (R)	1.57	0.63	3.43	92.4	2
Q9	Felt confident	4.30	0.65	3.30	89.5	6
Q10	Needed to learn a lot first (R)	2.01	0.81	2.99	66.7	10

Note: Favorable = responses of 4–5 on positive items and 1–2 on negative items. Rank 1 = highest adjusted mean (ties share a rank). Source: primary data, processed by the authors.

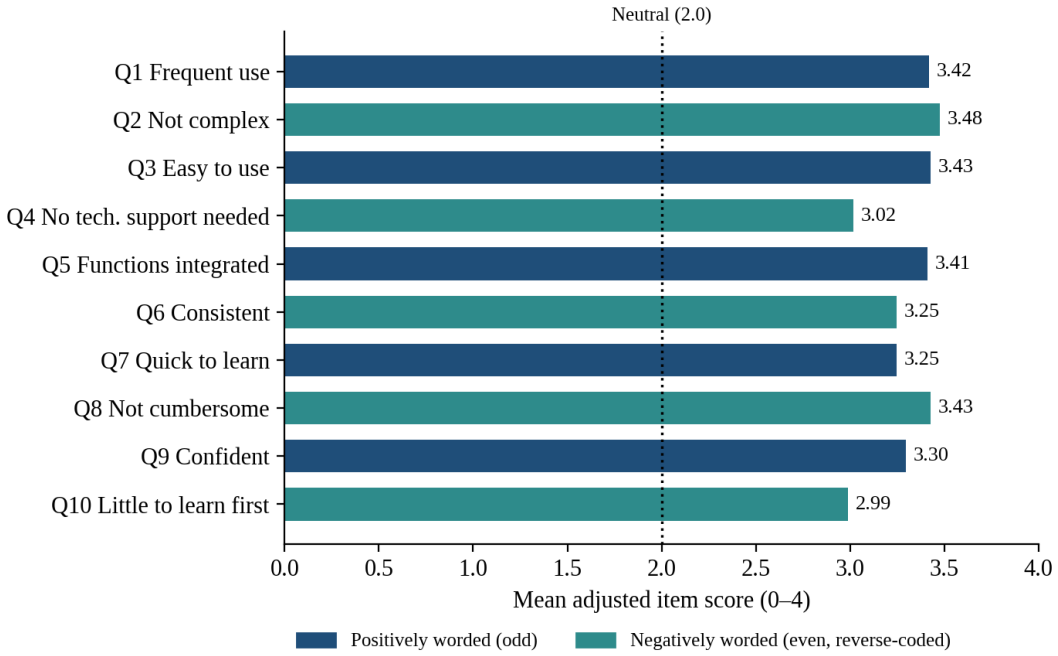


Figure 5: Mean adjusted score of each SUS item (0–4 scale)

The weakest results were found on two items that form the SUS Learnable subscale. Q10, needing to learn many things before getting started, had the lowest adjusted mean (2.99) and the lowest favorable share (66.7%); 35 cadets (33.3%) answered neutrally. Q4, needing the support of a technical person, followed with 3.02 and 75.2% favorable, and 26 cadets (24.8%) answered neutrally. Q6, too much inconsistency, and Q7, most people would learn quickly, shared an adjusted mean of 3.25; for Q6, 4 cadets (3.8%) agreed that there was too much inconsistency and 20 (19.0%) were neutral. As Figure 6 shows, these four items, together with Q9, are the only items for which the favorable share is below 90%, and Q10, Q4, and Q6 are the only items below 80%. For Q4 and Q10 no cadet gave an unfavorable response (4 or 5); the lower favorable shares are entirely due to neutral answers (3).

DOI: http://dx.doi.org/10.33603/perspective.v14i2.13015

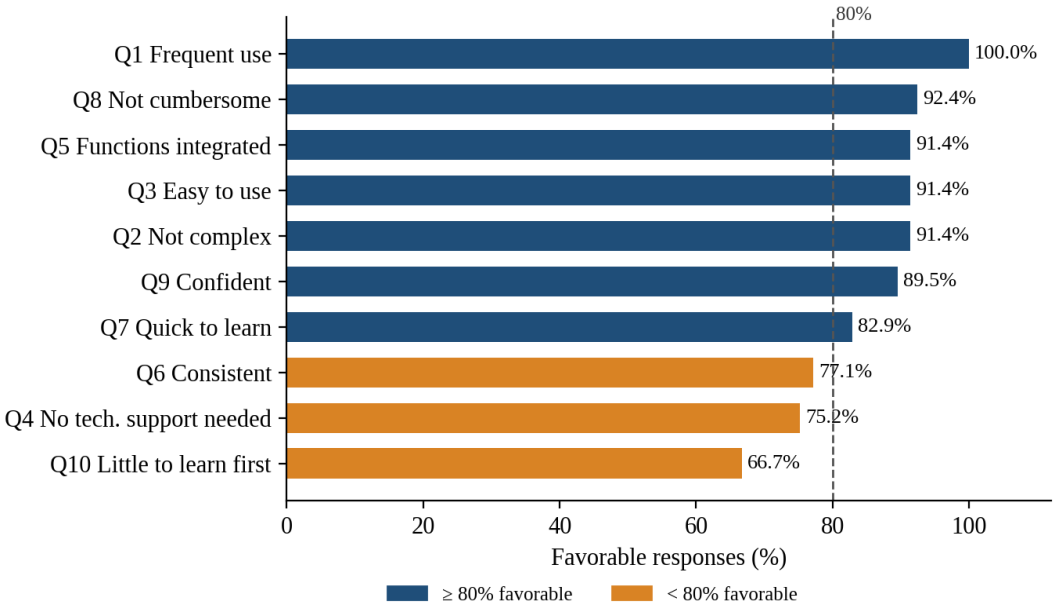


Figure 6: Percentage of favorable responses for each SUS item (sorted)

Learnable and Usable Subscales

Following (Lewis & Sauro, 2009), the SUS was also analyzed through two subscales (Figure 7), as a descriptive and tentative comparison only (see the caveats on the stability of this structure in the Introduction). The *Usable* subscale, comprising the eight items on complexity, ease of use, integration, consistency, learning speed, cumbersomeness, confidence, and frequency of use, reached 84.2, in the *Excellent* range. The *Learnable* subscale, comprising Q4 and Q10, reached 75.1, in the *Good* range but 9.1 points lower. Tentatively, cadets rated the e-book as easy and coherent to operate, while a share of them gave neutral rather than unfavorable answers about needing to learn something first or needing some support. Because the Learnable subscale has only two items and the two-factor structure is unstable (Lewis & Sauro, 2017), the 9.1-point gap may partly reflect item wording and should not be read as evidence of a distinct learnability problem.

DOI: http://dx.doi.org/10.33603/perspective.v14i2.13015

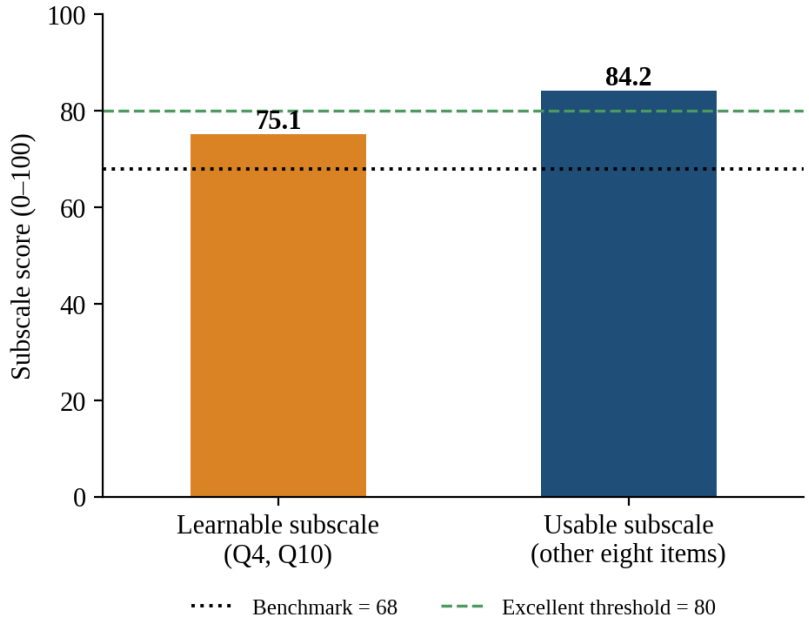


Figure 7: SUS Learnable and Usable subscale scores compared with the benchmark and the Excellent threshold

Priorities for Improving the E-Book

The item-level results were used to identify priorities for version 2.0 (Table 7). Relative priority was assigned according to the adjusted mean and the favorable share of the related item: the two Learnable items were given high priority, and the remaining items below 90% favorable were given medium or low priority. The cut-offs were set post hoc and are heuristic, without a published basis: high = adjusted mean below 3.10 (Q10, Q4); medium = 3.10–3.29 (Q6, Q7); low = 3.30 or above with a favorable share below 90% (Q9). Priority is relative within a very favorable profile: no cadet answered unfavorably on Q4 or Q10, so "high" priority does not indicate a usability problem. The suggested actions are derived from the content of the items and from the design features of the e-book; they are hypotheses to be confirmed with qualitative evidence from users, which was not part of the data analyzed here.

DOI: http://dx.doi.org/10.33603/perspective.v14i2.13015

Table 7: Prioritized improvement areas derived from the SUS item results

Item	Finding	Evidence	Suggested action	Relatively Priority
Q10	Cadets still need to learn several things before getting started	Adj. = 2.99; 66.7% favorable; 33.3% neutral	Add a one-page quick-start guide and a short orientation unit explaining navigation, glossary, and exercises	Relatively high
Q4	Some cadets would need technical support	Adj. = 3.02; 75.2% favorable; 24.8% neutral	Embed a help page or FAQ; demonstrate the main functions in the first session	Relatively high
Q6	A minority perceive inconsistency	Adj. = 3.25; 77.1% favorable; 3.8% agree	Audit layout, icons, and link behavior across units and standardize one template	Medium
Q7	Learning speed could be higher	Adj. = 3.25; 82.9% favorable	Shorten the path to core functions with a persistent navigation bar and a visible table of contents	Medium
Q9	Confidence can be strengthened	Adj. = 3.30; 89.5% favorable	Include short practice items with answer keys in the orientation unit	Low

Source: synthesis of the SUS item results.

Discussion: Overall Perceived Usability of the First Version

The mean SUS score of 82.4 suggests that, in this cohort and as rated in a post-use classroom survey, cadets perceived the EFE e-book very favorably (Excellent band in Table 2), far above the benchmark of 68 that is widely used to mark acceptable systems (Lewis & Sauro, 2017). Because the lower bound of the confidence interval (81.1) is itself above 80, sampling error alone is unlikely to explain the result, although the interval does not account for bias arising from the rating situation. All 105 cadets scored at or above the benchmark and the lowest score was 70.0, which means that the favorable evaluation was shared across the group. For a first-version educational product that was developed from a needs analysis and used within regular classes, this is a promising result. It is consistent with the premise of the research roadmap that materials grounded in learners' needs can also be accepted by learners, although a post-use score without a comparison group cannot show that needs-based design caused this acceptance. It is also consistent with (Lim & Hew, 2014), who reported positive student responses to the ease of use of an e-book.

Three cautions apply. First, scores were restricted to 70.0–92.5 and clustered at the upper end of the scale. This is a restricted range rather than strictly a ceiling effect, since no score approached 100, but it limits variance, the interpretability of alpha, and the discrimination among items. Second, the cadets rated a product developed by their own lecturers in a classroom setting, so social desirability and acquiescence (agreeing regardless of item content) may have raised the ratings; the survey administration described in the Method can reduce but not remove this risk. The very high favorable share on Q1 (100.0%) should therefore be read as a positive signal of acceptance rather than as proof of future use. Third, SUS measures perceived usability, not learning. A usable e-book is a necessary but not sufficient condition for improved technical English, which should be examined in later phases of the roadmap.

Strengths: Low Complexity, Ease of Use, and Integration

The highest-rated items concerned low complexity (Q2), ease of use (Q3), the absence of cumbersome operation (Q8), the wish to use the e-book frequently (Q1), and the integration

of functions (Q5). These items together describe a product that is simple to operate and coherent. This profile corresponds to the structural choices made in the development phase: modular units, a hyperlinked table of contents, a clickable glossary, and consistent navigation features. From a TAM perspective, perceived ease of use is a key antecedent of the intention to use a technology (Davis, 1989; Venkatesh et al., 2003) and the unanimous favorable response to Q1 is consistent with this logic. However, because perceived usefulness and actual use were not measured here, the link between usability and acceptance is suggested rather than tested.

The Learnability Gap

The clearest descriptive differentiation in the data is between the Usable subscale (84.2) and the Learnable subscale (75.1), a gap of 9.1 points, which should be read as tentative: the Learnable subscale has only two items, the Usability/Learnability structure was not consistently replicated in a large re-analysis (Lewis & Sauro, 2017) and the difference may partly reflect item tone rather than two real dimensions. Q10 and Q4 ranked lowest, and a quarter to a third of the cadets gave a neutral response to them. No cadet answered unfavorably on either item. Cadets therefore rated the e-book as easy to operate overall, and a quarter to a third of them were neutral, not negative, about needing to learn something or to receive support first. This is consistent with (Lim & Hew, 2014), who observed that students found an e-book easy to use but did not exploit its interactive features because they were not familiar with them, and with (Hernon et al., 2007), who identified navigation-related barriers as a factor hindering student use of e-books. An interactive PDF with hyperlinks, a clickable glossary, and embedded multimedia contains functions that are not self-evident, particularly for cadets whose first experience with this type of document is the EFE course. The recommendation to add a quick-start guide and embedded help (Table 7) is therefore low in cost and directly aligned with the weakest items. In user-centered design terms (Norman, 2013), it makes the available actions more discoverable and reduces the need for external support.

Q6, perceived inconsistency, was also relatively weak among the usable items, with a small group of cadets agreeing that the e-book was inconsistent. The present data cannot reveal the source of this perception, which may relate to differences in layout or link behavior between units or devices. Qualitative inquiry, such as short interviews or think-aloud sessions, is needed to locate it precisely before design changes are made.

Device and Context of Use

The respondent profile shows a context of use that is dominated by laptops (85.7%) and campus Wi-Fi (76.2%), with only 14.3% of cadets using a smartphone, no tablet users, and 9.5% relying on mobile data. The favorable SUS scores should therefore be understood as the experience of a mainly laptop-based, well-connected group. An interactive PDF is well suited to this context, but the results cannot be generalized to smartphone users, who are likely to face smaller text and tables, or to learners with limited connectivity. This point is consistent with the argument of (Sung et al., 2016) that the effectiveness of technology in education depends on the fit between media design, device, and learning context. Although device type and internet connectivity were collected alongside the SUS responses, these variables could not be reliably linked to individual SUS scores because the respondent identification codes were not consistently recorded across the relevant data fields. Consequently, subgroup comparisons based on device type and connectivity could not be

DOI: http://dx.doi.org/10.33603/perspective.v14i2.13015

performed. The analysis was therefore limited to descriptive summaries of the respondent profile, without examining differences in perceived usability across user subgroups.

Implications for ESP E-Book Development and Quality Assurance

The findings carry several implications. First, a usability evaluation should be a standard step after the development of ESP digital materials, because content that is accurate and contextual can still fail to deliver learning benefits if the user experience is poor (Fakhrurriana & Ningsih, 2023; Rubin & Chisnell, 2008). Second, the SUS provides a quick and comparable indicator that can be repeated for each version of the e-book, and its subscale reading may help to locate where intervention is needed, although the subscale structure is unstable. Third, the results suggest that an ESP e-book can be perceived very favorably even in a first version when its structure is kept simple and coherent, and that the remaining effort could be directed at onboarding rather than at changing the basic design, to be confirmed with qualitative evidence. Fourth, from an institutional perspective, converting the item-level evidence into a prioritized revision plan (Table 7) can serve as preliminary input for the third phase of the roadmap, namely the revision of the e-book into version 2.0; extending the e-book to other study programs would require separate evaluation in those programs.

Limitations

(1) *Single-institution context.* The study involved one cohort in one polytechnic, so the results may not generalize to other programs, devices, or connectivity conditions.

(2) *Perception-based, single-instrument measurement.* Only the SUS was analyzed. Task-based usability testing (for example, time on task and error counts), learning analytics, and supplementary scales for specific usability dimensions were not part of the dataset analyzed here.

(3) *No qualitative data.* Interviews or open-ended comments were not part of the data analyzed in this article, so the reasons behind the weaker items could not be explored and the improvement actions in Table 7 remain to be confirmed with users.

(4) *Restricted score range and low reliability.* Scores were concentrated between 70.0 and 92.5 (a restricted range at the upper end, not strictly a ceiling effect), and Cronbach's alpha was 0.41, which limits the precision of item-level comparisons and means that the total score cannot be treated as a clean single index.

(5) *Post-use design without a comparison group.* Cadets evaluated the e-book after eight sessions, with no baseline or comparison with printed material or other digital media, and the 88.2% response rate leaves room for non-response bias.

(6) *Evaluator–developer relationship and social desirability.* The cadets rated a product developed by their own lecturers, in a setting with a strong hierarchy, which raises the risk of social desirability and acquiescence even though participation was voluntary and anonymous; ratings may therefore be inflated.

(7) *English SUS not validated for this population.* The original English wording was used with Indonesian cadets without a documented validation or comprehension pilot, which may have affected the reverse-worded items and the low alpha.

(8) *Unstable subscale structure.* The Learnable/Usable split rests on two and eight items respectively, and a large re-analysis (Lewis & Sauro, 2017) did not replicate this structure, so the subscale comparison is descriptive and tentative.

DOI: http://dx.doi.org/10.33603/perspective.v14i2.13015

(9) *No systematic literature search.* The statement that studies of ESP e-book usability are scarce is based on the literature the authors were able to identify, not on a systematic search.

CONCLUSION

This study evaluated cadets' perceptions of the usability of the *English for Engineering* e-book for the RSTJ study program after real classroom use. Regarding the first research aim, the mean SUS score of 82.4 (SD = 6.9; 95% CI [81.1, 83.7]) falls in the *Excellent* band of the working categories used in this study and is significantly above the benchmark of 68, $t(104) = 21.26$, $p < .001$ ($d = 2.08$, inflated by the narrow spread of scores). All 105 respondents scored at or above 68, with 73.3% in the *Excellent* and 26.7% in the *Good* band. These ratings come from one cohort, one self-report instrument, and a post-use survey completed in class by cadets taught by the developers of the e-book.

Regarding the second aim, the strongest aspects were the low complexity, ease of use, absence of cumbersome operation, and integration of functions of the e-book, and every cadet responded favorably to the item on using the e-book frequently. The weakest aspects were the need to learn several things before getting started (Q10; adjusted mean 2.99) and the need for technical support (Q4; 3.02), so that the Learnable subscale (75.1) was 9.1 points lower than the Usable subscale (84.2), a tentative descriptive difference because the subscale structure is unstable and the two lowest items drew neutral rather than unfavorable answers. Perceived inconsistency (Q6) was a further, smaller concern.

Regarding the third aim, five relatively prioritized improvement areas (hypotheses to be confirmed with qualitative evidence) were identified for version 2.0: a quick-start guide and orientation unit, embedded help and a demonstration of the main functions, an audit and standardization of layout and links, a shorter path to core functions, and short practice items to build confidence. Overall, the study suggests that, in this cohort and in a post-use classroom survey, a needs-based ESP e-book was perceived very favorably in its first version, and that the remaining effort could concentrate on helping new users get started. These conclusions are limited by the single-institution sample, the reliance on a single self-report instrument administered to cadets taught by the e-book's developers, the unvalidated English SUS, the restricted range of scores, and the low internal consistency of the scale in this sample.

For developers and lecturers, it is recommended to: (1) add a quick-start guide, an orientation unit, and an embedded help page or FAQ to version 2.0, and demonstrate the main functions (navigation, glossary, bookmarks) in the first session; (2) audit the layout, icons, and link behavior across units and use one consistent template; and (3) consider providing an offline copy and a reflowable version for smartphones, and test them with cadets who use small screens or mobile data.

For the institution, it is recommended to: (1) consider using usability evaluation with the SUS as one routine step in the quality assurance of digital learning products, complemented by qualitative data and administered by someone other than the lecturers, and repeat it for each new version of the e-book; and (2) use the evidence from this study as preliminary input for the revision of the e-book into version 2.0, and evaluate the e-book separately before extending it to other study programs.

For future research, it is recommended to: (1) complement the SUS with interviews or open-ended questions to explain the weaker items and with task-based usability tests and learning analytics; (2) collect respondent characteristics at the individual level so that SUS scores can be compared across devices, connectivity, and intensity of use; (3) evaluate version 2.0 with

DOI: http://dx.doi.org/10.33603/perspective.v14i2.13015

a larger and more heterogeneous sample, including smartphone users, and re-examine the reliability of the scale; (4) use a pre-post or quasi-experimental design to assess the effect of the e-book on technical vocabulary, reading, and writing performance; and (5) use a validated or piloted Indonesian version of the SUS and an independent survey administrator, and report item-level response distributions.

ACKNOWLEDGEMENTS

The authors express gratitude to Politeknik Keselamatan Transportasi Jalan for the research funding support and to all cadets who participated in this research.

REFERENCES

Alkadrie, S. A., Fauzan Albahy, A., Handayani, Y., & Zein, A. F. (2025). Evaluation of the Usability of Online Learning Platforms Using the System Usability Scale Method and Nielsen Usability Attributes. *ITDM: Information Technology Development And Management*, *1*(1), 107–120. https://journal.uinjkt.ac.id/index.php/itdm/article/view/50374

Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An empirical evaluation of the system usability scale. *International Journal of Human-Computer Interaction*, *24*(6). https://doi.org/10.1080/10447310802205776

Brilianti, D. F., Fitriani, N., Fahmadi, A. E., Pradana, B., Rokhim, A., & Najwa, D. A. (2026). Integration of Needs Analysis in Developing E-Book Materials for English for Engineering in Road Transportation System Engineering Study Program. *Allure Journal*, *6*(2). https://doi.org/10.26877/allure.v6i2.1000

Brooke, J. (1996). SUS -A quick and dirty usability scale Usability and context. *Usability Evaluation in Industry*, *189*(194).

Buroidah, H., Zubaidah, S., & Mahanal, S. (2023). Effects of Project-Based Learning with Project Guide E-book on Critical Thinking and Metacognitive Skills: A Case from Undergraduate Biology Students in Genetic 1 Course. *Prisma Sains : Jurnal Pengkajian Ilmu Dan Pembelajaran Matematika Dan IPA IKIP Mataram*, *11*(2). https://doi.org/10.33394/j-ps.v11i2.6727

Changpueng, P., & Pattanapichet, F. (2023). A Needs Analysis of English for Meeting Lessons for Thai Undergraduate Engineering Students. *LEARN Journal: Language Education and Acquisition Research Network*, *16*(2), 603–623. https://so04.tci-thaijo.org/index.php/LEARN/article/view/266974

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly: Management Information Systems*, *13*(3). https://doi.org/10.2307/249008

Dudley-Evans, T., & St John, M. J. (1998). *Developments in English for specific purposes: A multi-disciplinary approach*. Cambridge University Press.

Efendi, M. Y., Mayasari, E., Iswanto, J., & Apriyani, L. (2025). Needs Analysis of English for Specific Purpose at Engineering Students in Vocational High School. *Issues in Applied Linguistics & Language Teaching*, *7*(1).

https://doi.org/10.37253/iallteach.v7i1.9333

Hernon, P., Hopper, R., Leach, M. R., Saunders, L. L., & Zhang, J. (2007). E-book Use by Students: Undergraduates in Economics, Literature, and Nursing. *Journal of Academic Librarianship*, *33*(1). https://doi.org/10.1016/j.acalib.2006.08.005

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

ISO. (2018). *Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts* (ISO 9241-11:2018). International Organization for Standardization. https://www.iso.org/standard/63500.html

Lewis, J. R., & Sauro, J. (2009). The factor structure of the system usability scale. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *5619 LNCS*. https://doi.org/10.1007/978-3-642-02806-9_12

Lewis, J. R., & Sauro, J. (2017). Revisiting the factor structure of the System Usability Scale. *Journal of Usability Studies*, *12*(4), 183–192. https://uxpajournal.org/wp-content/uploads/sites/7/pdf/JUS_Lewis_August2017.pdf

Lim, E. L., & Hew, K. F. (2014). Students' perceptions of the usefulness of an E-book with annotative and sharing capabilities as a tool for learning: a case study. *Innovations in Education and Teaching International*, *51*(1). https://doi.org/10.1080/14703297.2013.771969

Marisda, D. H., Tolla, I., & Arsyad, M. (2025). Developing and Validating an Interactive Flipbook E-Book to Foster Critical Thinking in Modern Physics Education. *TWIST*, *20*(1), 279–288. https://doi.org/10.5281/zenodo.10049652

Mohamad Saleh, N. N., & Mohammed, Y. (2024). Tourist Guides' Needs for Arabic for Specific Purposes According to The Hutchinson-Waters Model. *Al-Azkiyaa - International Journal of Language and Education*, *3*(1). https://doi.org/10.33102/alazkiyaa267

Nateghian, N. (2024). English language needs of Iranian students of civil engineering: Are the courses aligned with workplace needs? *English for Specific Purposes*, *76*. https://doi.org/10.1016/j.esp.2024.08.002

Norman, D. A. (2013). *The design of everyday things* (Rev. and expanded ed.). Basic Books.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). McGraw-Hill.

Ramadhani, R., Aulawi, H., Setiawan, R., & Rahayu, S. (2023). The Speaking English Material Needs of Engineering Students for the English Club Program: An ESP Analysis. *Scope : Journal of English Language Teaching*, *7*(2). https://doi.org/10.30998/scope.v7i2.16310

Rubin, J., & Chisnell, D. (2008). *Handbook of usability testing: How to plan, design, and conduct effective tests* (2nd ed.). Wiley.

Sasabone, L., Irman, Taufiq, A., & Hasbi, N. A. (2024). Needs Analysis of English for Specific Purposes (ESP) in Students' Speaking Materials. *EDULEC: Education, Language, and Culture Journal*, *4*(1). https://doi.org/10.56314/edulec.v4i1.218

Sun, P. C., Tsai, R. J., Finger, G., Chen, Y. Y., & Yeh, D. (2008). What drives a successful e-Learning? An empirical investigation of the critical factors influencing learner

satisfaction. *Computers and Education*, *50*(4). https://doi.org/10.1016/j.compedu.2006.11.007

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly: Management Information Systems*, *27*(3). https://doi.org/10.2307/30036540

Wahyuni, S., & Rosana, D. (2025). Using Interactive Science E-Books Using PBL Model on Critical Thinking Skills of Junior High School Students. *JPI (Jurnal Pendidikan Indonesia)*, *14*(1). https://doi.org/10.23887/jpiundiksha.v14i1.84159

Zaharias, P., & Poylymenakou, A. (2009). Developing a usability evaluation method for e-learning applications: Beyond functional usability. *International Journal of Human-Computer Interaction*, *25*(1). https://doi.org/10.1080/10447310802546716