Evaluating Course Performance and Student Perceptions through Direct and Indirect Assessment: A Study of the Information Technology and Information Systems I Course

Nyamkhuu Lkhagva[1,2]*, Khongorzul Baldorj[2], Nasanbayar Baavgai[2], Bayarmaa Gombo[1]

[1]School of Educational Studies, Mongolian National University of Education, Ulaanbaatar, Mongolia

[2]Etugen University, Ulaanbaatar, Mongolia

***Corresponding author**

DOI: https://doi.org/10.47772/IJRISS.2026.100900701

Received: 01 October 2026; Accepted: 06 October 2026; Published: 21 October 2026

ABSTRACT

This study examined course performance and student perceptions to identify improvement priorities for the Information Technology and Information Systems I course at Etugen University, Mongolia. A quantitative descriptive design used a course assessment summary representing 109 first-year undergraduate dentistry students in the fall semester of the 2025–2026 academic year. Direct evidence comprised reported scores for seminar tasks, typing performance, a course test, an independent assignment, and the end-of-course examination; attendance was treated as a participation indicator. Indirect evidence came from three separate questionnaires administered in Mongolian via Google Forms after course instruction. All 109 students were invited; the skills and teaching practices questionnaires each received 70 responses (64.2%), and the course assessment questionnaire received 66 (60.6%). Findings were analyzed descriptively. Mean scores represented 94.5% of the maximum for the independent assignment, 87.3% for seminar tasks, 70.8% for the course test, and 70.7% for the end-of-course examination. These percentages described performance within each component and did not establish comparable competence across tasks. Favorable skill self-ratings were lowest for Excel functions (72.9%), and 45.7% of respondents selected Excel for further practice. Group work and consultation outside class received favorable ratings of 52.9% and 65.7%, respectively. Favorable ratings for assessment aspects ranged from 86.4% to 90.9%. Questionnaire responses were not linked to performance records, and criterion-level rubric scores were unavailable. The findings identify Excel practice, group activities, consultation access, and communication of assessment weights as priorities for review. Direct and indirect evidence provided complementary perspectives, but did not establish learning gains, student-level relationships, or causal effects.

Keywords: direct assessment, indirect assessment, course evaluation, rubric-based assessment, student perceptions

INTRODUCTION

Higher education institutions need evidence of student learning to evaluate educational quality and identify areas for improvement. Outcome-based education (OBE) organizes curriculum design, teaching, and assessment around intended learning outcomes (Tam, 2014). Constructive alignment connects these outcomes with learning activities and assessment tasks (Biggs, 1996; Biggs & Tang, 2011). However, overall course grades may combine achievement, participation, and task completion. Examining assessment components separately can therefore provide a clearer account of students' performance.

Direct and indirect assessments offer complementary information for course evaluation. Examinations, practical tasks, and assignments provide direct evidence of performance under specified assessment conditions. Student questionnaires provide indirect evidence of perceived skill acquisition and experiences of teaching and assessment. Both sources can inform improvement, provided that their distinct meanings are

retained (Banta & Palomba, 2014; Suskie, 2018). Student evaluations should be interpreted alongside performance evidence rather than treated as equivalent measures of learning (Spooren et al., 2013; Constantinou & Wijnen-Meijer, 2022).

This distinction is relevant to introductory information technology courses, where students develop practical skills in document processing, spreadsheet use, and presentation preparation. Analytic rubrics can clarify expectations and structure the assessment of these tasks (Reddy & Andrade, 2010). However, evidence of improved assessed performance does not, by itself, establish gains in self-regulated learning or self-efficacy (Panadero et al., 2023). A related course-level study by Abd Latiff et al. (2025) illustrates how performance results, student feedback, and self-assessments can inform course evaluation.

The present study evaluates the Information Technology and Information Systems I course for first-year undergraduate dentistry students at Etugen University, Mongolia. It draws on a course assessment summary representing 109 students and aggregate questionnaire summaries containing 70 responses to the skills self-assessment and teaching practices sections and 66 responses to the course assessment section. The study describes performance across assessment components and examines student perceptions to identify priorities for practical instruction, learning support, and assessment communication. Because individual questionnaire responses could not be linked to performance records, the analysis is descriptive and does not estimate student-level associations or causal effects.

This study addresses an institutional need to interpret course assessment results alongside student feedback in an introductory digital skills course for dentistry students in Mongolia. Its contribution lies in showing how these sources can inform course improvement while preserving the distinction between assessed performance and perceived learning. Component scores describe performance within particular assessment conditions, whereas student feedback identifies perceived learning needs and experiences that scores alone cannot capture. Considering these sources together provides a basis for selecting priorities for further investigation without assuming that favorable perceptions demonstrate learning outcome attainment or that aggregate patterns establish student-level relationships.

Research Aim

To evaluate course performance and student perceptions in the Information Technology and Information Systems I course using complementary direct and indirect assessment evidence and identify areas for course improvement.

Research Objectives

- To describe students' performance across course assessment components, distinguishing participation indicators from measures of assessed knowledge and skills.
- To summarize students' self-assessments of acquired skills and their evaluations of teaching practices and course assessment.
- To consider direct and indirect assessment findings together and identify areas for improving teaching, student support, and assessment in the course.

LITERATURE REVIEW

Assessment serves several purposes in teaching and learning. Black and Wiliam (1998) examined how classroom assessment can support learning when the evidence collected informs subsequent instructional decisions. Schellekens et al. (2021) reviewed assessment of learning, assessment for learning, and assessment as learning, identifying overlapping characteristics and inconsistencies in their definitions and application. These perspectives suggest that assessment practices should be understood in terms of how evidence is used. Assessment can document achievement, guide instruction, or support students' evaluation of their work, depending on how it is implemented.

Feedback helps learners use assessment evidence to improve their work. Sadler (1989) argued that learners need to understand the expected standard, compare their current work with that standard, and identify actions to close the gap. Nicol and Macfarlane-Dick (2006) connected formative assessment and feedback with students' capacity to regulate their learning. Hattie and Timperley (2007) further emphasized that the effects of feedback depend on its content and focus. These frameworks draw attention to whether assessment offers clear expectations and actionable guidance as well as scores.

Students also need opportunities to interpret and use assessment information. Carless and Boud (2018) described feedback literacy as the capacities and dispositions needed to understand feedback and act on it. Tai et al. (2018) emphasized evaluative judgement, or the ability to make decisions about the quality of work. Boud and Soler (2016) argued that assessment should prepare students to evaluate their learning beyond the immediate course. These perspectives support examining students' experiences of assessment clarity, feedback, and learning support. Favorable questionnaire responses, however, do not establish that feedback literacy or evaluative judgement has developed.

Analytic rubrics specify criteria and describe levels of performance, providing a structure for assessing practical work. Reddy and Andrade (2010) reviewed their use in higher education and their potential to clarify expectations. Jonsson and Svingby (2007) examined rubric reliability, validity, and educational consequences, while Brookhart (2018) emphasized the importance of appropriate criteria. In information technology courses, rubric criteria can describe observable features of documents, spreadsheets, and presentations. The resulting scores should be interpreted in relation to the tasks and criteria assessed, rather than as comprehensive measures of digital competence.

Rubrics may support learning when students use them to review and improve their work. Panadero and Jonsson (2013) examined their formative use and the conditions that influence their usefulness. In a focus-group study involving 14 undergraduate students, Andrade and Du (2005) reported that participants used rubrics to plan assignments, check their work, and interpret feedback. Some participants nevertheless viewed rubrics primarily as a means of meeting an instructor's requirements. These findings illustrate possible learning uses but also show that providing a rubric does not ensure meaningful engagement with its criteria.

Evidence concerning rubric use also indicates that different outcomes require separate interpretation. Panadero and Romero (2014) compared rubric-supported and unsupported self-assessment in a concept-mapping task and reported differing findings for performance, assessment accuracy, learning strategies, and task-related difficulties. Panadero et al. (2023) found a positive pooled effect of rubric use on academic performance, whereas estimates for self-regulated learning and self-efficacy remained uncertain. Improved assessed performance therefore does not, by itself, establish improvement across all dimensions of learning. Course evaluation should distinguish the use of rubrics from evidence of their effects.

Self-assessment encompasses practices with different purposes. Andrade (2019) distinguished formative self-assessment intended to guide improvement from activities focused on assigning a score to one's own work. Boud and Falchikov (1989) critically examined comparisons between student-generated and teacher-generated marks. Falchikov and Boud's (1989) meta-analysis showed that agreement varied across study characteristics and educational contexts. Students' judgements can provide useful information, but their accuracy requires investigation. General reports of skill acquisition should also be distinguished from criterion-based judgements of a completed task.

Empirical research further supports separating perceived competence from demonstrated achievement. Sitzmann et al. (2010) examined self-assessed knowledge in relation to cognitive and affective outcomes, providing evidence against treating perceived knowledge as interchangeable with objectively assessed learning. In undergraduate physics, Deslauriers et al. (2019) found that students' feelings of learning did not consistently reflect measured achievement. Although the subject and instructional setting differed from those of the present study, the findings illustrate why perceived learning and assessed performance should remain distinct in course evaluation.

Student evaluations of teaching provide information about educational experiences, but their interpretation

requires care. Spooren et al. (2013) reviewed their validity and highlighted the complexity of interpreting teaching ratings. Uttl et al. (2017) questioned their use as indicators of student learning based on a meta-analysis. Constantinou and Wijnen-Meijer (2022) emphasized considering student evaluations alongside other evidence when assessing teaching effectiveness. These studies support using questionnaires to examine experiences of teaching, support, and assessment. Comparatively lower ratings can identify issues for further inquiry, while favorable ratings do not establish higher achievement.

Studies combining direct and indirect evidence offer examples at both program and course levels. Purewsuren et al. (2026) examined the relationship between direct and indirect assessments in a software engineering program. Their program-level analysis differs in scope from the evaluation of a single course. Abd Latiff et al. (2025) combined direct outcome attainment, student feedback, and entrance–exit self-assessments in a Data Communication and Network course involving 27 students. Their approach illustrates how multiple sources can inform evaluation while retaining the distinction between assessed performance and self-reported learning. Numerical findings from such studies should not be compared directly across courses with different tasks, scoring systems, and questionnaires.

The literature supports examining assessed performance and student perceptions as complementary sources of evidence for course improvement. It also highlights the need to interpret rubric scores in relation to specific criteria and to distinguish self-reported skills from demonstrated proficiency. These principles guide the present descriptive analysis of course assessment summaries and questionnaire findings. The two sources are examined together to identify improvement priorities, without assuming student-level agreement or attributing the findings to particular teaching or assessment practices.

RESEARCH METHODOLOGY

Research Design

This study used a quantitative descriptive design to evaluate course performance and student perceptions in the Information Technology and Information Systems I course within an outcome-based education framework. Course assessment summaries provided evidence of assessed performance, while aggregate questionnaire summaries described students' perceptions of their skills, teaching practices, and course assessment. The two sources were analyzed separately and considered together to identify priorities for course improvement. The analysis did not estimate baseline-to-final performance changes, student-level associations, or causal effects of teaching and assessment practices.

Participants and Setting

The study focused on first-year undergraduate dentistry students enrolled in the Information Technology and Information Systems I course at Etugen University, Mongolia, during the fall semester of the 2025–2026 academic year. The course assessment summary represented 109 students.

All 109 students were invited to complete questionnaires administered through separate Google Forms. The skills self-assessment and teaching practices questionnaires each received 70 responses (64.2% of invited students), while the course assessment questionnaire received 66 responses (60.6%). These totals reflect participation in separate forms rather than missing responses within a single questionnaire. Responses across forms could not be matched at the individual level. Findings were therefore reported separately for each questionnaire and interpreted as respondents' perceptions rather than as representative of the entire cohort.

Direct Assessment

Direct assessment evidence comprised reported scores for seminar and practical tasks, typing performance, a course test, an independent assignment, and the end-of-course examination. The grading scheme allocated 70 points to continuous assessment and 30 points to the end-of-course examination, yielding a maximum total course score of 100 points. Table 1 presents the assessment components and their maximum scores.

Table 1. Course Assessment Components and Maximum Scores

Assessment component	Maximum score
Attendance	8
Seminar and task performance	32
Continuous assessment 1: Typing performance	10
Continuous assessment 2: Test	10
Independent assignment	10
Continuous assessment subtotal	70
End-of-course examination	30
Total course score	100

Note. The continuous assessment subtotal includes attendance. Attendance was treated as an indicator of participation rather than a direct measure of acquired knowledge or skills.

The seminar and task performance component comprised eight seminars, each contributing a maximum of four points to the course grade, for a total of 32 points. Practical tasks were developed to align with the relevant assessment criteria, and students' completed work was evaluated using the corresponding analytic rubrics or task-specific scoring checklists.

Five-level analytic rubrics were used to assess practical tasks in Microsoft Word and Excel. The Word rubric included four criteria covering page setup; image insertion and formatting; table creation and formatting; and headers, footers, and page numbering. The Excel rubric included three criteria covering data entry, formatting, and calculations; worksheet management; and chart creation and formatting. Performance levels V, IV, III, II, and I corresponded to 4, 3, 2, 1, and 0 points, respectively, yielding maximum raw scores of 16 for the Word rubric and 12 for the Excel rubric.

The course test comprised 30 selected-response questions drawn from an item bank covering computer components and device functions, information units, data and information, networks, system software, cybersecurity concepts, and the functions and commands of Microsoft Word, Excel, and PowerPoint. The item bank included multiple-choice and true–false questions. Each correct response received one point, yielding a maximum raw score of 30. The raw score was divided by three to obtain the test's contribution to the course grade, with a maximum of 10 points. The item bank contained content relevant to computer fundamentals (K1), application use (K2), and cybersecurity (K3). However, the number of administered items corresponding to each outcome was not documented in the materials available for this analysis. The test assessed knowledge rather than demonstrated performance on practical tasks.

The independent assignment required students to prepare both a Microsoft Word document and a PowerPoint presentation. These were assessed using a scoring checklist, with a maximum of 12 points allocated to the Word document and 8 points to the PowerPoint presentation. The combined raw score of 20 was divided by two to obtain the assignment's contribution to the course grade, with a maximum of 10 points.

The end-of-course examination was a 90-minute practical assessment comprising tasks in Microsoft Word, Excel, and PowerPoint. The Word task assessed text and page formatting, symbol insertion, diagram creation, and the manipulation of shapes and images (10 points). The Excel task assessed spreadsheet construction and formatting, calculations, the use of COUNT, MAX, and AVERAGE functions, and chart creation and formatting (12 points). The PowerPoint task required a five-slide presentation incorporating a theme, formatted text, images, slide transitions, and object animations (8 points). Performance was assessed using a criterion-based scoring checklist with specified point allocations. These points were summed directly to produce an examination score out of 30. The examination checklist was distinct from the five-level analytic rubrics used for practical tasks.

A single instructor assessed the practical tasks, independent assignments, and end-of-course examinations for all 109 students. Subject instructors jointly developed the examination tasks and agreed on the

assessment criteria, point allocations, and scoring principles before marking. This provided a shared framework for scoring; however, inter-rater reliability was not assessed because the students' work was marked by one instructor.

Criterion-level scores were unavailable for the present analysis. Although each seminar contributed a maximum of four points, the complete procedures for converting practical-task rubric or checklist scores into seminar points were not documented in the available materials. The independent assignment conversion and examination point allocations were documented. Performance was therefore summarized using the component-level means reported in the course assessment summary. These means did not establish attainment of individual course learning outcomes or permit analysis of performance on specific rubric or checklist criteria.

Appendix A maps the stated course learning outcomes to relevant learning activities and available assessment evidence. It identifies documented content coverage and score allocations, as well as gaps in outcome-specific assessment documentation. The mapping describes the documented alignment between outcomes, activities, and assessments; it does not establish attainment of individual learning outcomes.

Appendix B presents the end-of-course examination scoring checklist and point allocations.

Appendix C presents English translations of the Word and Excel rubrics provided on the student self-assessment sheets.

Indirect Assessment

Indirect evidence consisted of aggregate summaries from three separate Google Forms questionnaires addressing perceived acquisition of knowledge and skills, teaching practices, and course assessment. The questionnaires were developed by the course instructor and pilot-tested with 15 students before administration. They were administered in Mongolian after course instruction had been completed.

All 109 students enrolled in the course were invited to participate. The skills self-assessment and teaching practices questionnaires each received 70 responses (64.2%), while the course assessment questionnaire received 66 responses (60.6%). Responses across the separate forms were not linked at the individual level; therefore, equal response totals were not interpreted as evidence that the same students completed both questionnaires.

Internal consistency was assessed using responses from the main study. Cronbach's alpha coefficients were 0.85 for the knowledge and skills self-assessment questionnaire, 0.88 for the teaching practices questionnaire, and 0.86 for the course assessment questionnaire. These coefficients provided evidence of internal consistency in the main study sample.

The skills self-assessment questionnaire contained 11 items covering Microsoft Word, Excel, and PowerPoint. Response categories were 5 = acquired very well, 4 = acquired well, 3 = acquired to a moderate extent, 2 = not acquired, and 1 = unable to answer. The "unable to answer" category was treated separately and was not interpreted as the lowest level of competence. An additional question asked which application students wished to strengthen further.

The teaching practices questionnaire addressed group work, peer learning, student engagement, consultation outside class, teacher modeling, collaboration, equitable treatment, practical demonstration, and safety instructions. The course assessment questionnaire addressed clarity, transparency, fairness, assessment methods, point allocation, and the appropriateness of practical, test-based, and criterion-based assessments.

Responses in categories 4 and 5 were combined to summarize favorable perceptions. For the skills self-assessment questionnaire, these categories indicated that students perceived a skill as acquired "well" or "very well." This grouping was used for descriptive reporting and did not constitute a validated threshold for competence or assessment quality.

The results tables present selected items whose response distributions were sufficiently clear in the available charts to reconstruct counts. The equitable treatment item was excluded because its displayed percentages could not be reliably reconciled with the stated response total. Questionnaire findings were interpreted as respondents' perceptions rather than direct evidence of demonstrated achievement and were not generalized to all 109 students.

Data Sources and Procedure

The course was delivered over 14 weeks. Continuous assessment comprised attendance, seminar and practical tasks, a typing assessment, a course test, and an independent assignment. The end-of-course examination constituted the final assessment.

Seminar instruction followed a course handbook containing model tasks and task-specific self-assessment checklists. Activities included identifying computer specifications and configuring Windows and Gmail settings; creating, saving, and formatting Word documents; constructing tables and charts; entering, sorting, and calculating spreadsheet data; preparing PowerPoint slides and animations; and searching for and processing online information. The handbook also provided exercises for repeated practice outside class. Its task-completion checklists were distinct from the post-course Google Forms questionnaires used in this study.

Students submitted seminar tasks and the independent assignment by email, and the instructor assessed the submitted work using the relevant criteria. The independent assignment assessment included acknowledgment of sources through citations and a reference list. The instructor also assessed students' confidence and active engagement during practical exercises, group activities, and presentations. These practices addressed aspects of email communication, responsible use of sources, and engagement in digital learning. However, separate outcome-level scores were unavailable for the present analysis.

The materials reviewed comprised the course syllabus, seminar handbook, course assessment summary, grading scheme, Word and Excel analytic rubrics, task-specific scoring checklists, course test item bank, independent assignment criteria, end-of-course examination tasks and scoring checklist, and questionnaire instruments and aggregate response charts. These materials were used to describe assessment procedures, summarize component-level performance, and map the stated course learning outcomes to relevant activities and available assessment evidence. The questionnaire documentation also included the reported pilot sample size and internal consistency estimates. Individual assessment records and criterion-level scores were unavailable for the present analysis.

Following completion of course instruction, all 109 students were invited to complete three separate Google Forms questionnaires in Mongolian concerning perceived knowledge and skills, teaching practices, and course assessment. Participation was voluntary. The questionnaires yielded 70, 70, and 66 responses, corresponding to response rates of 64.2%, 64.2%, and 60.6%, respectively.

Cronbach's alpha coefficients were calculated from individual questionnaire responses collected in the main study. The descriptive reporting of questionnaire findings in this article, however, relied on aggregate summaries and charts. Where counts were not directly available, they were reconstructed from the displayed percentages and the response total for the relevant questionnaire. Questionnaire responses were not linked to individual course assessment records or matched across forms.

Direct and indirect evidence remained distinct throughout the analysis. Course assessment scores described performance under the conditions of each assessment component, whereas questionnaire responses described students' perceptions. These sources were considered together to identify possible priorities for course improvement, without assuming that they measured identical outcomes or represented matched observations from the same students.

Descriptive Analysis

The analysis was restricted to descriptive summaries supported by the available materials. Course

assessment results were presented as reported component-level mean scores for the 109-student cohort and as percentages of the maximum available scores. Each percentage was calculated by dividing the reported component mean by its maximum score and multiplying by 100. Individual assessment records were unavailable for recalculation; therefore, standard deviations, medians, ranges, score distributions, and component-specific valid observation counts were not reported.

Assessment components were presented as a descriptive performance profile. Expressing scores as percentages of their respective maxima did not establish equivalence across components. Differences in task content, difficulty, weighting, timing, available resources, and scoring conditions limited their comparability. Consequently, differences between component percentages were not interpreted as direct differences in competence or as changes in learning over time. Attendance was treated as an indicator of participation rather than demonstrated knowledge or skills.

Questionnaire findings were summarized using frequencies and percentages based on the response total for each separate form. Where counts were unavailable, they were reconstructed from the percentages displayed in the aggregate charts and the corresponding questionnaire response total. Favorable-response percentages were calculated by combining reconstructed counts for categories 4 and 5, dividing by the relevant response total, and rounding to one decimal place. The “unable to answer” category in the skills questionnaire was retained as a separate response category and was not interpreted as a competence rating. Reconstruction of counts did not recover individual response patterns.

Questionnaire response rates were calculated by dividing the number of responses to each form by the 109 students invited to participate. Findings were interpreted within the responding groups. Potential differences between respondents and nonrespondents could not be assessed using the available aggregate materials.

Direct and indirect findings were considered together descriptively to identify possible priorities for course improvement. No paired-samples tests, correlations, regression analyses, or causal effect estimates were conducted. The analysis did not establish student-level relationships between perceived and demonstrated learning.

Ethical Considerations

The study was reviewed and approved by the Ethics Review Committee of the Department of Engineering at Etugen University at its meeting on October 22, 2025. Students provided informed consent using a participant information and consent form. Participation in the questionnaires was voluntary. Course assessment information was anonymized for analysis, and questionnaire findings were analyzed and reported in aggregate form. The information was used solely for research purposes, and no personally identifiable student information is presented in this article.

RESULTS

The results describe performance across course assessment components, students’ self-reported skill acquisition, and their evaluations of teaching practices and course assessment. Direct and indirect findings are presented separately, followed by a descriptive synthesis of priorities for course improvement.

Course Assessment Results. Table 2 presents the component-level mean scores reported in the assessment summary representing 109 students. Each mean is also expressed as a percentage of the corresponding maximum score.

Table 2. Results Across Course Assessment Components

Assessment component	Mean score	Maximum score	Percentage of maximum score (%)
Attendance	7.63	8	95.4
Seminar and task performance	27.94	32	87.3
Continuous assessment 1: typing performance	7.74	10	77.4

Continuous assessment 2: test	7.08	10	70.8
Independent assignment	9.45	10	94.5
End-of-course examination	21.20	30	70.7

Note. Percentages were calculated by dividing the mean score by the maximum available score and multiplying by 100. Attendance represents participation rather than a direct measure of acquired knowledge or skills.

The reported mean scores were 7.63 out of 8 for attendance (95.4%), 27.94 out of 32 for seminar and task performance (87.3%), 7.74 out of 10 for typing performance (77.4%), 7.08 out of 10 for the course test (70.8%), 9.45 out of 10 for the independent assignment (94.5%), and 21.20 out of 30 for the end-of-course examination (70.7%).

These results describe performance within each assessment component. Although expressing mean scores as percentages provides a common reporting format, the components differed in task requirements and assessment conditions. The percentages therefore do not establish equivalent levels of competence across components. Attendance reflects participation and should be distinguished from assessments of knowledge and skills.

Criterion-level scores were unavailable, preventing identification of specific skills or course learning outcomes attained by individual students. The component-level summaries provide a basis for reviewing course assessment practices, but they do not demonstrate changes in learning over time or explain the causes of the observed performance patterns.

Students' Self-Assessments of Acquired Skills. The skills self-assessment section contained 70 responses. Ratings indicating that a skill had been acquired "well" or "very well" were combined as favorable responses. Table 3 presents the reconstructed counts and corresponding percentages.

Table 3. Students' Self-Assessments of Acquired Skills

Skill	Students reporting acquisition well or very well	Percentage (%)
Processing information in Word	63	90.0
Processing text in Word	62	88.6
Working with images in Word	60	85.7
Working with tables in Word	63	90.0
Working with charts in Word	59	84.3
Configuring page settings in Word	54	77.1
Working with tabular data in Excel	64	91.4
Working with charts in Excel	61	87.1
Using functions in Excel	51	72.9
General proficiency in Excel	59	84.3
Using PowerPoint	57	81.4

Note. Each item received 70 responses. Counts were reconstructed from the percentages displayed in the questionnaire charts and the stated response total. Combined percentages were calculated from these counts and rounded to one decimal place. The results represent perceived skill acquisition rather than demonstrated performance.

Among the 70 respondents to the skills self-assessment questionnaire, favorable responses ranged from 72.9% for Excel functions to 91.4% for tabular data processing in Excel. For Word-related skills, favorable responses ranged from 77.1% for page settings to 90.0% for information processing and table management. Favorable responses for PowerPoint skills accounted for 81.4%.

When asked which application they wished to strengthen further, 32 respondents selected Excel (45.7%), 27 selected PowerPoint (38.6%), and 11 selected Word (15.7%). Excel functions received the lowest proportion

of favorable self-assessments among the reported skills, and Excel was the most frequently selected application for further practice. Together, these findings identify Excel as a priority for instructional review and additional learning support.

These results reflect respondents' perceptions and learning preferences. They do not establish demonstrated proficiency in the applications or a relationship between self-assessed skills and course assessment scores. The questionnaire response rate was 64.2%, and the findings were not generalized to the full cohort of 109 students.

Students' Evaluations of Teaching Practices. The teaching practices section contained 70 responses. Ratings of 4 and 5 were combined as favorable responses, as reported in Table 4.

Table 4. Students' Evaluations of Teaching Practices

Teaching practice	Students providing favorable ratings	Percentage (%)
Organized group work	37	52.9
Organized peer-learning activities	56	80.0
Actively engaged all students	66	94.3
Provided consultation outside class	46	65.7
Served as a positive role model	64	91.4
Worked collaboratively with students	67	95.7
Demonstrated tasks and guided students through them	63	90.0
Provided instructions on safety and appropriate equipment use	68	97.1

Note. Each item received 70 responses. Counts were reconstructed from the chart percentages and the stated response total.

Among the 70 respondents to the teaching practices questionnaire, favorable responses were most frequent for safety instructions (97.1%), collaboration (95.7%), and student engagement (94.3%). Teacher modeling and practical demonstration received favorable responses from 91.4% and 90.0% of respondents, respectively, while peer learning received 80.0%.

Group work (52.9%) and consultation outside class (65.7%) received lower proportions of favorable responses than the other reported items. These patterns suggest that the organization of group activities and opportunities for consultation warrant further review. The questionnaire findings alone, however, do not explain why respondents evaluated these aspects less favorably.

The response rate was 64.2%. These findings describe respondents' perceptions of teaching practices and do not establish the effects of those practices on learning achievement. The equitable treatment item was omitted because its displayed percentages could not be reliably reconciled with the stated response total.

Students' Evaluations of Course Assessment. The course assessment section contained 66 responses. Table 5 summarizes favorable ratings for clarity, transparency, fairness, assessment methods, point allocation, and the appropriateness of specific assessment approaches.

Table 5. Students' Evaluations of Course Assessment

Assessment aspect	Students providing favorable ratings	Percentage (%)
Assessment was clear	60	90.9
Assessment was transparent	59	89.4
Assessment accurately reflected performance	59	89.4
Assessment methods were appropriate	60	90.9
Allocation of assessment points was appropriate	57	86.4
Assessment through practical software tasks was	60	90.9

appropriate		
Test-based assessment was appropriate	60	90.9
Criterion-based assessment of seminar tasks was appropriate	58	87.9

Note. Each item received 66 responses. Favorable evaluations represent ratings of 4 or 5. Counts were reconstructed from the chart percentages and the stated response total.

Among the 66 respondents to the course assessment questionnaire, favorable responses ranged from 86.4% for point allocation to 90.9% for assessment clarity, assessment methods, and the appropriateness of practical and test-based assessments. Transparency and fairness each received favorable responses from 89.4% of respondents, while criterion-based assessment received 87.9%.

Point allocation received the lowest proportion of favorable responses among the reported assessment items, although most respondents evaluated it favorably. This finding identifies the communication of assessment weights and their contribution to the total course score as an area for further review. It does not, by itself, demonstrate that the grading scheme was unclear or inappropriate.

The questionnaire response rate was 60.6%. These findings reflect respondents' perceptions of course assessment and do not independently establish assessment validity, scoring reliability, or attainment of course learning outcomes. Because responses were not linked to individual assessment records, relationships between these perceptions and course performance could not be examined.

Synthesis of Direct and Indirect Assessment Findings. The course assessment summary described performance on assessed tasks, while the questionnaire summaries described perceived skills and course experiences. The sources provided different information that could inform course improvement.

Excel functions received the lowest proportion of favorable skill self-ratings, and Excel was the application most frequently selected for further practice. Group work, consultation outside class, and the allocation of assessment points received comparatively lower favorable ratings within their respective questionnaire sections. These findings identify areas for reviewing practical instruction, learning support, and assessment communication.

Individual questionnaire responses could not be linked to performance records. Agreement between perceived skills and assessed performance was therefore not tested. The synthesis identifies possible improvement priorities without establishing student-level associations or causal explanations.

DISCUSSION

This study examined a course assessment summary representing 109 students alongside aggregate questionnaire findings to identify priorities for improving the Information Technology and Information Systems I course. The assessment summary described performance on course tasks, while the questionnaires captured perceived skills and educational experiences. Examining these sources together provided a broader basis for course review, consistent with approaches that connect assessment evidence with improvement decisions (Banta & Palomba, 2014; Suskie, 2018).

Mean scores for the independent assignment and seminar tasks represented higher percentages of their maximum scores than those for the course test and end-of-course examination. These differences may reflect task demands and assessment conditions as well as differences in performance. Assignments may offer more time, access to resources, feedback, or opportunities for revision. The percentages therefore cannot be interpreted as directly comparable measures of competence. Reviewing task content and scoring criteria would help clarify the skills assessed by each component.

Constructive alignment emphasizes the relationship between assessment tasks and intended learning outcomes (Biggs, 1996; Biggs & Tang, 2011). The component-level means in this study provided an overview of performance, but they did not establish attainment of individual course learning outcomes.

Attendance was treated separately as an indicator of participation. Mapping practical tasks, test items, and rubric criteria to specific outcomes would strengthen future interpretations of achievement.

The available rubrics and scoring checklists clarified the criteria used to assess practical work, the independent assignment, and the end-of-course examination. The examination tasks were jointly developed by subject instructors, who agreed on the criteria, point allocations, and scoring principles. However, criterion-level scores were unavailable, and the procedures for converting practical-task scores into the four-point contribution of each seminar were not fully documented in the available materials. These limitations prevented identification of specific strengths and difficulties within the assessed tasks. One instructor marked all students' work, and inter-rater reliability was not assessed. Explicit criteria support transparent assessment, but evidence supporting score interpretation remains necessary (Brookhart, 2018; Jonsson & Svingby, 2007). Panadero et al. (2023) reported a positive pooled effect of rubric use on academic performance, while estimates for self-regulated learning and self-efficacy remained uncertain. The present descriptive study did not evaluate whether rubric use influenced student performance.

The questionnaire findings suggest that Excel warrants further attention. The proportion of favorable skill self-ratings was lowest for using Excel functions (72.9%), and 45.7% of respondents selected Excel for further practice. These responses indicate perceived learning needs but do not establish a measured proficiency gap. Self-assessed knowledge differs from objectively assessed learning (Sitzmann et al., 2010), and perceived learning may not reflect demonstrated achievement (Deslauriers et al., 2019). A practical diagnostic task could help determine whether students have difficulties understanding functions, selecting formulas, or applying them accurately.

Targeted exercises, worked examples, and opportunities for revision could address difficulties identified through diagnostic tasks. Feedback should clarify the expected standard, help students evaluate their current work, and guide corrective action (Sadler, 1989). Support for interpreting and using feedback may also help students monitor and adjust their learning approaches (Carless & Boud, 2018; Nicol & Macfarlane-Dick, 2006). These activities are proposed improvements; their effectiveness was not evaluated in this study.

Safety instructions, collaboration with students, and student engagement received favorable ratings from more than 94% of respondents. Group work and consultation outside class received comparatively lower ratings, at 52.9% and 65.7%, respectively. These findings suggest examining how group activities and consultation opportunities are organized and communicated. Lower ratings could reflect limited opportunities, scheduling barriers, unclear expectations, or dissatisfaction. Follow-up questions would help distinguish these explanations before changes are introduced to the course.

Teaching ratings describe students' experiences and do not directly measure learning. Reviews have identified limitations in using student evaluations to establish teaching effectiveness or achievement (Spooren et al., 2013; Uttl et al., 2017). Constantinou and Wijnen-Meijer (2022) emphasized interpreting these evaluations alongside other evidence. In the present course, the ratings can guide inquiry into participation and access to support, but they do not explain differences in assessed performance.

Favorable ratings for the selected assessment aspects ranged from 86.4% to 90.9%. Point allocation received the lowest proportion of favorable ratings within this section. This finding warrants examining students' understanding of assessment weights and their rationale. Explaining how the weights relate to task demands and intended outcomes could improve assessment communication. However, perceived fairness and transparency do not establish reliability or validity, which require examination of tasks, criteria, and scoring procedures.

Students may also benefit from opportunities to apply assessment criteria to examples of work. Andrade and Du (2005) reported that students used rubrics to plan and check assignments, although some viewed them mainly as a means of meeting an instructor's requirements. Discussing examples and explaining scoring decisions could support more active engagement with the criteria. Such activities are consistent with the development of evaluative judgement described by Tai et al. (2018), but the present data do not show whether this capacity developed.

Abd Latiff et al. (2025) offered a related example of course evaluation using direct attainment, student feedback, and entrance–exit self-assessments. The present study similarly drew on different evidence sources to identify improvement priorities. Individual questionnaire responses could not, however, be linked to performance records. Correspondence between perceived skills and assessed performance was therefore not tested. The findings provide complementary perspectives rather than statistically verified agreement.

Several limitations affect interpretation. The study examined one course at one institution, limiting generalizability. All 109 students were invited to complete three separate questionnaires, but the skills and teaching practices questionnaires each received 70 responses (64.2%), and the course assessment questionnaire received 66 responses (60.6%). Respondents' experiences may have differed from those of nonrespondents, and nonresponse bias could not be assessed. Responses could not be linked across questionnaires or matched to assessment records. Individual assessment records for the 109-student cohort were unavailable for recalculating the reported means or examining score distributions. The descriptive questionnaire analysis relied on aggregate charts and reconstructed counts, which did not preserve individual response patterns. Although the practical-task rubrics, independent assignment scoring checklist, and examination point allocations were available, criterion-level scores and complete documentation of the conversion of practical-task scores into seminar points were unavailable. Internal consistency estimates for the three questionnaires did not establish their validity or the accuracy of students' self-assessments. Finally, the descriptive design did not support causal conclusions or estimates of the effects of teaching and assessment practices.

Future course evaluation should retain individual assessment records, criterion-level rubric scores, and documented scoring procedures. Where informed consent and data governance arrangements permit, pseudonymous identifiers could link questionnaire responses to performance records. This would support examination of the relationship between perceived and demonstrated skills while protecting student identities.

The findings suggest a structured improvement cycle for future course delivery. Diagnostic tasks could first identify specific learning needs in Word, Excel, and PowerPoint, with particular attention to Excel functions. These results could guide targeted instruction, worked examples, and practice using explicit rubric criteria aligned with course learning outcomes. Feedback and opportunities for revision could then be followed by reassessment using comparable tasks and scoring criteria. Clearly defined group roles, scheduled consultation opportunities, and worked examples of assessment weight calculations could address the other priorities identified by respondents.

Instructors and the department could evaluate this cycle by tracking criterion-level attainment, participation in consultations, and questionnaire response distributions across successive course offerings. Reviewing these indicators could inform decisions about retaining or modifying the changes. Questionnaire response rates and possible nonresponse bias should also be considered in each evaluation. This improvement cycle is proposed for future implementation; its effectiveness was not evaluated in the present study.

CONCLUSION

This study used course assessment summaries and student questionnaire findings to identify priorities for improving the Information Technology and Information Systems I course. Assessment results described performance across course components, while questionnaire responses provided complementary perspectives on perceived skills, teaching practices, and assessment processes.

The findings identified Excel practice, the organization of group activities, consultation opportunities, and communication of assessment weights as areas for further review. These priorities reflect descriptive performance summaries and respondents' perceptions; they do not establish causal effects, changes in

learning over time, or student-level relationships between perceived and demonstrated competence.

Future course evaluation should connect clearly defined learning outcomes with assessment tasks and rubric criteria, retain individual and criterion-level records, and evaluate changes through a structured cycle of diagnostic assessment, targeted instruction, guided practice, and reassessment. This approach could strengthen the evidence available for course improvement and institutional quality assurance.

Data Availability

The aggregate findings used in this study are reported in the article. Individual student assessment records are not publicly available because they contain confidential educational information and were unavailable for the present analysis. Internal consistency coefficients were calculated from individual questionnaire responses, whereas the descriptive questionnaire findings reported here were derived from aggregate summaries and charts. Criterion-level rubric and checklist scores were unavailable for this analysis.

Conflict of Interest

The authors declare no conflicts of interest related to this study.

REFERENCES

1. Abd Latiff, Z. I., M. Zin, N. A., Mohd Atan, F., Ismail, N. A., & Abu Bakar, Z. (2025). Programme outcome measurement through triangulated evidence: Direct attainments and indirect assessment in a Data Communication & Network course. International Journal of Research and Innovation in Social Science, 9(9), 9510–9517. https://doi.org/10.47772/IJRISS.2025.909000783
2. Andrade, H. L. (2019). A critical review of research on student self-assessment. Frontiers in Education, 4, Article 87. https://doi.org/10.3389/feduc.2019.00087
3. Andrade, H., & Du, Y. (2005). Student perspectives on rubric-referenced assessment. Practical Assessment, Research & Evaluation, 10(3), 1–11. https://doi.org/10.7275/g367-ye94
4. Banta, T. W., & Palomba, C. A. (2014). Assessment essentials: Planning, implementing, and improving assessment in higher education (2nd ed.). Jossey-Bass.
5. Biggs, J. (1996). Enhancing teaching through constructive alignment. Higher Education, 32(3), 347–364. https://doi.org/10.1007/BF00138871
6. Biggs, J., & Tang, C. (2011). Teaching for quality learning at university: What the student does (4th ed.). Open University Press.
7. Black, P., & Wiliam, D. (1998). Assessment and classroom learning. Assessment in Education: Principles, Policy & Practice, 5(1), 7–74. https://doi.org/10.1080/0969595980050102
8. Boud, D., & Falchikov, N. (1989). Quantitative studies of student self-assessment in higher education: A critical analysis of findings. Higher Education, 18(5), 529–549. https://doi.org/10.1007/BF00138746
9. Boud, D., & Soler, R. (2016). Sustainable assessment revisited. Assessment & Evaluation in Higher Education, 41(3), 400–413. https://doi.org/10.1080/02602938.2015.1018133
10. Brookhart, S. M. (2018). Appropriate criteria: Key to effective rubrics. Frontiers in Education, 3, Article 22. https://doi.org/10.3389/feduc.2018.00022
11. Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. Assessment & Evaluation in Higher Education, 43(8), 1315–1325. https://doi.org/10.1080/02602938.2018.1463354
12. Constantinou, C., & Wijnen-Meijer, M. (2022). Student evaluations of teaching and the development of a comprehensive measure of teaching effectiveness for medical schools. BMC Medical Education, 22(1), Article 113. https://doi.org/10.1186/s12909-022-03148-6
13. Deslauriers, L., McCarty, L. S., Miller, K., Callaghan, K., & Kestin, G. (2019). Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom. Proceedings of the National Academy of Sciences, 116(39), 19251–19257. https://doi.org/10.1073/pnas.1821936116
14. Falchikov, N., & Boud, D. (1989). Student self-assessment in higher education: A meta-analysis. Review of Educational Research, 59(4), 395–430. https://doi.org/10.3102/00346543059004395

15. Hattie, J., & Timperley, H. (2007). The power of feedback. Review of Educational Research, 77(1), 81–112. https://doi.org/10.3102/003465430298487
16. Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. Educational Research Review, 2(2), 130–144. https://doi.org/10.1016/j.edurev.2007.05.002
17. Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. Studies in Higher Education, 31(2), 199–218. https://doi.org/10.1080/03075070600572090
18. Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. Educational Research Review, 9, 129–144. https://doi.org/10.1016/j.edurev.2013.01.002
19. Panadero, E., Jonsson, A., Pinedo, L., & Fernández-Castilla, B. (2023). Effects of rubrics on academic performance, self-regulated learning, and self-efficacy: A meta-analytic review. Educational Psychology Review, 35(4), Article 113. https://doi.org/10.1007/s10648-023-09823-4
20. Panadero, E., & Romero, M. (2014). To rubric or not to rubric? The effects of self-assessment on self-regulation, performance and self-efficacy. Assessment in Education: Principles, Policy & Practice, 21(2), 133–148. https://doi.org/10.1080/0969594X.2013.877872
21. Purewsuren, B., Tsangia, U., Manaljav, Y., & Lkhagva, N. (2026). Study on the relationship between direct and indirect assessments of a software engineering program. In 2026 6th International Conference on Electrical, Computer and Energy Technologies (ICECET) (pp. 1–7). IEEE. https://doi.org/10.1109/ICECET65726.2026.11633106
22. Reddy, Y. M., & Andrade, H. (2010). A review of rubric use in higher education. Assessment & Evaluation in Higher Education, 35(4), 435–448. https://doi.org/10.1080/02602930902862859
23. Sadler, D. R. (1989). Formative assessment and the design of instructional systems. Instructional Science, 18(2), 119–144. https://doi.org/10.1007/BF00117714
24. Schellekens, L. H., Bok, H. G. J., de Jong, L. H., van der Schaaf, M. F., Kremer, W. D. J., & van der Vleuten, C. P. M. (2021). A scoping review on the notions of assessment as learning (AaL), assessment for learning (AfL), and assessment of learning (AoL). Studies in Educational Evaluation, 71, Article 101094. https://doi.org/10.1016/j.stueduc.2021.101094
25. Sitzmann, T., Ely, K., Brown, K. G., & Bauer, K. N. (2010). Self-assessment of knowledge: A cognitive learning or affective measure? Academy of Management Learning & Education, 9(2), 169–191. https://doi.org/10.5465/amle.9.2.zqr169
26. Spooren, P., Brockx, B., & Mortelmans, D. (2013). On the validity of student evaluation of teaching: The state of the art. Review of Educational Research, 83(4), 598–642. https://doi.org/10.3102/0034654313496870
27. Suskie, L. (2018). Assessing student learning: A common sense guide (3rd ed.). Jossey-Bass.
28. Tai, J., Ajjawi, R., Boud, D., Dawson, P., & Panadero, E. (2018). Developing evaluative judgement: Enabling students to make decisions about the quality of work. Higher Education, 76(3), 467–481. https://doi.org/10.1007/s10734-017-0220-3
29. Tam, M. (2014). Outcomes-based approach to quality assessment and curriculum improvement in higher education. Quality Assurance in Education, 22(2), 158–168. https://doi.org/10.1108/QAE-09-2011-0059
30. Uttl, B., White, C. A., & Gonzalez, D. W. (2017). Meta-analysis of faculty's teaching effectiveness: Student evaluation of teaching ratings and student learning are not related. Studies in Educational Evaluation, 54, 22–42. https://doi.org/10.1016/j.stueduc.2016.08.007

APPENDIX A: MAPPING OF COURSE LEARNING OUTCOMES TO LEARNING ACTIVITIES AND AVAILABLE ASSESSMENT EVIDENCE

Code	Course learning outcome	Relevant learning activities	Available assessment evidence and criteria	Score allocation
K1	Identify the basic components, functions, and	Study of computer components and device functions;	The course test item bank includes questions on computer components,	The 30-item course test was scored out of 30 and divided by three to

	uses of computers and smart devices.	practical identification of computer specifications using the seminar handbook.	device functions, information units, and computer classification. Specific coverage of smart devices was not documented in the available assessment materials.	contribute a maximum of 10 course points. The number of administered items and points attributable to K1 was not documented.
K2	Explain the functions and differences in use of Microsoft Word, Excel, and PowerPoint.	Instruction and practical exercises involving document processing, spreadsheets, and presentations.	The course test item bank includes questions on application functions, commands, and formula syntax. These questions address aspects of this outcome but do not fully assess students' ability to explain differences between the applications.	The course test contributed a maximum of 10 course points. The number of administered items and points attributable to K2 was not documented.
K3	Recognize basic cybersecurity concepts.	Study of cybersecurity, digital ethics, and safe digital practices.	The course test item bank includes questions on information security risks, email-related threats, and malicious software. These provide evidence of relevant item-bank content, rather than verified outcome attainment.	The course test contributed a maximum of 10 course points. The number of administered items and points attributable to K3 was not documented.
S1	Create, edit, and save basic documents using Microsoft Word.	Document creation, saving, and formatting exercises; work with images, tables, headers, footers, and page numbering; preparation of the independent assignment.	The Word practical-task rubric covers page setup, image insertion and formatting, table creation and formatting, and headers, footers, and page numbering. Additional evidence includes the independent assignment scoring checklist and the Word examination checklist.	Word rubric: maximum raw score of 16. Independent assignment Word document: 12 raw points, equivalent to 6 course points. Examination Word task: 10 points. The seminar component totaled 32 points; the allocation specifically attributable to S1 was not documented.
S2	Create basic spreadsheets, use formulas, and produce charts in Microsoft Excel.	Spreadsheet construction, data entry and formatting, calculations, worksheet management, sorting, and chart exercises.	The Excel practical-task rubric covers data entry, formatting and calculations, worksheet management, and chart creation and formatting. The examination checklist covers spreadsheet construction, formatting, calculations, COUNT, MAX, AVERAGE, and charts.	Excel rubric: maximum raw score of 12. Examination Excel task: 12 points. The seminar component totaled 32 points; the allocation specifically attributable to S2 was not documented.
S3	Search for information online and communicate using email.	Online information searching and processing; Gmail activities; submission of	Students submitted seminar tasks and the independent assignment by email, and the instructor assessed the submitted work. This	Submitted work contributed to the seminar component (32 points) and independent assignment (10 points). A separate

		seminar tasks and the independent assignment by email.	documents the use of email for course-related communication. Separate criteria for assessing online searching and the quality of email communication were not documented.	allocation for online searching or email communication was not documented.
A1	Follow safe digital practices and respect and protect others' information.	Cybersecurity learning activities and preparation of citations and reference lists for the independent assignment.	Cybersecurity test items assessed related knowledge. The independent assignment assessed source acknowledgment through citations and a reference list. These addressed aspects of digital safety and respect for others' work, but did not assess all aspects of information protection.	The course test contributed a maximum of 10 course points; the cybersecurity-specific allocation was not documented. The reference-list criterion carried 0.5 raw points, equivalent to 0.25 course points after conversion. A separate allocation for in-text citations was not documented.
A2	Demonstrate confidence and active engagement when learning and working in digital environments.	Practical exercises, group activities, independent work, and assignment presentations.	The instructor assessed students' confidence and active engagement during practical exercises, group activities, and presentations. However, detailed scoring criteria and outcome-specific scores were unavailable for the present analysis.	A separate point allocation for confidence and active engagement was not documented in the materials available for this analysis.

Note. K = knowledge; S = skills; A = attitudes. The seminar component comprised eight seminars, each contributing a maximum of four points, for a total of 32 course points. Rubric levels V, IV, III, II, and I corresponded to 4, 3, 2, 1, and 0 points, respectively. Complete procedures for converting practical-task rubric or checklist scores into seminar points were not documented in the available materials. The independent assignment comprised a Word document (12 raw points) and a PowerPoint presentation (8 raw points); the combined score was divided by two to contribute a maximum of 10 course points. The examination comprised Word (10 points), Excel (12 points), and PowerPoint (8 points), totaling 30 points. PowerPoint production was assessed, although the stated outcomes did not include a separate skill outcome for creating presentations. Component scores are reported for context and should not be interpreted as exclusive allocations to individual outcomes. Item-bank coverage does not establish the distribution of questions in the administered test. This mapping describes documented alignment and gaps in assessment evidence; it does not establish attainment of individual course learning outcomes.

APPENDIX B :END-OF-COURSE EXAMINATION SCORING CHECKLIST

The end-of-course examination was a 90-minute practical assessment in Microsoft Word, Excel, and PowerPoint. Subject instructors jointly developed the tasks and agreed on the assessment criteria, point allocations, and scoring principles. One instructor assessed all 109 students' work.

Application	**Assessment criterion**	**Maximum points**
Microsoft Word	Apply bold, italic, and underline formatting.	1
Microsoft Word	Apply superscript and subscript formatting.	1
Microsoft Word	Apply font color and shading.	1
Microsoft Word	Apply case changes and strikethrough formatting.	1

Microsoft Word	Set the top, bottom, left, and right page margins.	1
Microsoft Word	Set line spacing.	1
Microsoft Word	Set the font and font size.	1
Microsoft Word	Insert symbols.	0.5
Microsoft Word	Insert a diagram.	0.5
Microsoft Word	Format the diagram.	1
Microsoft Word	Insert an image into a shape.	0.5
Microsoft Word	Format shapes.	0.5
Word subtotal		**10**
Microsoft Excel	Construct the required table.	0.5
Microsoft Excel	Apply cell borders.	1
Microsoft Excel	Apply cell fill colors.	0.5
Microsoft Excel	Create a chart.	1
Microsoft Excel	Format chart data labels, axes, and displayed data.	2
Microsoft Excel	Format the chart background.	1
Microsoft Excel	Perform the required calculations.	2
Microsoft Excel	Use the COUNT function.	1
Microsoft Excel	Use the MAX function.	1
Microsoft Excel	Use the AVERAGE function.	1
Excel subtotal		**12**
Microsoft PowerPoint	Create the required slides.	1
Microsoft PowerPoint	Apply a presentation theme.	1
Microsoft PowerPoint	Set the font and font size.	1
Microsoft PowerPoint	Apply text alignment.	1
Microsoft PowerPoint	Set line spacing.	1
Microsoft PowerPoint	Apply slide transitions.	1
Microsoft PowerPoint	Insert images.	1
Microsoft PowerPoint	Apply object animations.	1
PowerPoint subtotal		**8**
Total examination score		**30**

Note. The PowerPoint task required a five-slide presentation. Criterion scores were summed directly to obtain the examination score out of 30; no further conversion was applied. This checklist was distinct from the five-level analytic rubrics used for seminar practical tasks. Criterion-level student scores were unavailable for the present analysis.

APPENDIX C: FIVE-LEVEL RUBRICS FOR STUDENT SELF-ASSESSMENT OF WORD AND EXCEL TASKS

C1. Microsoft Word Rubric

Learning outcome: Develop the ability to prepare and process documents using Microsoft Office applications.

Assessment criterion	Level V (4 points)	Level IV (3 points)	Level III (2 points)	Level II (1 point)	Level I (0 points)
Page setup	Opens Word and sets A4 paper size, page margins, measurement units (cm or mm), and page orientation.	Opens Word and sets A4 paper size, page margins, and measurement units, but does not set page orientation.	Opens Word and sets A4 paper size and page margins, but does not set measurement units.	Opens Word and selects A4 paper size, but does not set page margins.	Opens Word but does not configure page settings.

Image insertion and formatting	Inserts an image, adjusts its size, changes its style, and crops it.	Inserts an image, adjusts its size, and changes its style, but does not crop it.	Inserts an image and adjusts its size, but does not change its style.	Inserts an image but does not adjust its size.	Does not insert an image.
Table creation and formatting	Inserts a table, changes its style, and uses Merge Cells, Split Cells, and cell alignment commands.	Inserts a table, changes its style, and uses Merge Cells and Split Cells, but does not use cell alignment commands.	Inserts a table and changes its style, but does not use Merge Cells or Split Cells.	Inserts a table but does not change its style.	Does not insert a table.
Headers, footers, and page numbering	Inserts headers and footers, adds page numbering including the first page, and inserts a date.	Inserts headers and footers and adds page numbering including the first page, but does not insert a date.	Inserts headers and footers and adds page numbering, but does not number the first page.	Inserts headers and footers but does not add page numbering.	Does not insert headers, footers, or page numbering.

Note. Each criterion is scored from 0 to 4, yielding a maximum raw score of 16.

C2. Microsoft Excel Rubric

Learning outcome: Develop the ability to prepare and process spreadsheet data using Microsoft Office applications.

Assessment criterion	**Level V (4 points)**	**Level IV (3 points)**	**Level III (2 points)**	**Level II (1 point)**	**Level I (0 points)**
Data entry, formatting, and calculations	Enters and formats data using Wrap Text and Orientation, performs calculations, copies data, and completes processing according to the instructions.	Enters and formats data using Wrap Text and Orientation, performs calculations, and copies data.	Enters and formats data using Wrap Text and Orientation and performs calculations.	Enters and formats data using Wrap Text and Orientation.	Enters data only.
Worksheet management	Renames and adds worksheets, saves the workbook, applies formatting, and configures settings.	Renames and adds worksheets, saves the workbook, and applies formatting.	Renames and adds worksheets and saves the workbook.	Renames and adds worksheets.	Renames a worksheet only.
Chart creation and formatting	Creates a chart, changes its style, adds a title and data labels, and changes its location.	Creates a chart, changes its style, and adds a title and data labels.	Creates a chart, changes its style, and adds a title.	Creates a chart and changes its style.	Creates a chart only.

Note. Each criterion is scored from 0 to 4, yielding a maximum raw score of 12. In the original Excel rubric, Level I assigns zero points to basic task completion; this score should not be interpreted as an absence of all relevant skills.

The rubrics were provided as student self-assessment sheets, and the instructor also used the corresponding criteria to assess practical tasks. These task-based self-assessments were distinct from the post-course

Google Forms questionnaires. Criterion-level scores were unavailable for the present analysis, and the procedures for converting raw scores into each seminar's four-point contribution were not fully documented.

APPENDIX D: STUDENT QUESTIONNAIRES

The following items were translated into English from the Mongolian questionnaire items displayed in the supplied Google Forms response summaries. The questionnaires were administered separately in Mongolian after course instruction.

D1. Knowledge and Skills Self-Assessment

Item	Statement
1	I have developed the ability to process information using Microsoft Word.
2	I have learned to process text in Microsoft Word.
3	I have learned to work with images in Microsoft Word.
4	I have learned to work with tables in Microsoft Word.
5	I have learned to work with charts in Microsoft Word.
6	I have learned to configure page settings in Microsoft Word.
7	I have learned to work with tabular data in Microsoft Excel.
8	I have learned to work with charts in Microsoft Excel.
9	I have learned to use functions in Microsoft Excel.
10	I have developed the ability to use Microsoft Excel.
11	I have learned to use Microsoft PowerPoint effectively.

Response categories for Items 1–11: 5 = acquired very well; 4 = acquired well; 3 = acquired to a moderate extent; 2 = not acquired; 1 = unable to answer.

Additional question: Which application would you like to develop your skills in further?

Response options: Microsoft Word; Microsoft Excel; Microsoft PowerPoint.

D2. Evaluation of Teaching Practices

Item	Statement
1	The instructor organized group work.
2	The instructor organized activities that enabled students to learn from one another.
3	The instructor actively engaged all students during class.
4	The instructor provided consultation outside class.
5	The instructor was a positive role model.
6	The instructor was a supportive collaborator.
7	The instructor treated students without discrimination.
8	The instructor demonstrated tasks and guided students in performing them.
9	The instructor provided instructions on safe working practices and the proper care of equipment.

Response options displayed in the supplied summaries: 5, 4, 3, 2, and 1.

Note. Verbal descriptions of these response categories were not shown in the supplied summaries. Item 7 is included here as part of the questionnaire, although its response distribution was excluded from the reported results because the displayed percentages could not be reliably reconciled with the stated response total.

D3. Evaluation of Course Assessment

Original item number	Statement
1	Course assessment was understandable.
2	Course assessment was transparent.
3	Course assessment accurately reflected performance.
4	The methods used to assess the course were appropriate.
5	The allocation of assessment points was appropriate.
6	Assessment through practical tasks using software applications was appropriate.
7	Test-based assessment was appropriate.
8	Criterion-based assessment of seminar tasks was appropriate.

Response options displayed in the supplied summaries: 5, 4, 3, 2, and 1.

Note. The original item numbers are retained. The supplied summaries contained no item numbered 6. Verbal descriptions of the response categories were not shown.

Appendix D presents English translations of the questionnaire items and the response options available in the supplied summaries. Verbal response anchors for the teaching practices and course assessment questionnaires were unavailable in those summaries