Disaggregating Student Achievement in Maintenance Planning, Scheduling and Control: A Retrospective Assessment Study

Ahmad Fadzil Jobli[1], Noor Azland Jainudin[2*], Diana Indim[3], Lannie Francis[4], Farah Hafifee Ahmad[5], Nurul Najihah Mad Rosni[6], Nur Farahiah Ibrahim[7], Anniza Hamdan[8]

[1,2,4]Faculty of Civil Engineering, Universiti Teknologi MARA, Cawangan Sarawak, Kota Samarahan, 94300, Sarawak, Malaysia

[3]Faculty of Applied Sciences, Universiti Teknologi MARA, Cawangan Sabah, 88997 Kota Kinabalu, Sabah, Malaysia

[5,6]Faculty of Civil Engineering, Universiti Teknologi MARA, 40450, Shah Alam, Selangor, Malaysia

[7]Faculty of Electrical Engineering, Universiti Teknologi MARA, Cawangan Sarawak, Kota Samarahan, 94300, Sarawak, Malaysia

[8]Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Cawangan Sarawak, Kota Samarahan, 94300, Sarawak, Malaysia

***Corresponding Author**

DOI: https://doi.org/10.47772/IJRISS.2026.1026EDU0682

Received: 16 October 2026; Accepted: 21 October 2026; Published: 02 November 2026

ABSTRACT

The aggregated grade hides the unevenness with which students develop the skills required by engineers in their professional careers. The purpose of this research is to determine how direct assessments of evidence show students' understanding of knowledge and applied program development, as well as their ability to communicate professionally in the Maintenance Planning, Scheduling and Control course of the Bachelor of Civil Engineering Technology (Asset Management) with Honours programme at UiTM Sarawak. A retrospective census of 18 anonymized student records from the March–July 2026 teaching session was constructed by combining records from two class worksheets in one workbook into one cohort. Assessment components were standardized to a 0–100% scale using the syllabus mapping. The mean scores were 90.06% for procedural knowledge, 67.33% for group project-based applied programme development and 84.78% for professional communication. The overall weighted mean score was 78.17%, and the difference between the reported knowledge and applied domain means was 22.73 percentage points. As there is no specified competency threshold, these comparisons are descriptive rather than evidence of formally defined attainment. Missing individual demonstrations and independent ratings limit validation of tool proficiency. Future assessment should integrate an analytic project rubric, individual verification and formative checkpoints. The study illustrates how competency-specific analysis can make quality-improvement priorities visible without attributing learning differences to any untested instructional cause.

Keywords: assessment alignment; asset management education; engineering competencies; graduate attributes; learning analytics; retrospective cohort

INTRODUCTION

Civil infrastructure maintenance involves much more than knowledge of inspection procedures and terms relating to maintenance. Civil infrastructure maintenance practitioners must take technical information and develop it into feasible work programs. They also select suitable tools and resource requirements for their projects. In addition, they must be able to monitor activities associated with these projects and communicate

decisions made about them to engineering colleagues. When these interdependent capabilities are represented through separate course-level competencies mapped to broader graduate attributes in education, a single weighted course mark may not necessarily indicate that students perform consistently across them.

Outcome-oriented education emphasizes clear expectations and assessment evidence tied to what graduates can demonstrate. Constructive alignment provides a basis for linking intended learning objectives, teaching activities, and assessment (Ramírez De Dampierre et al., 2024) while engineering education scholarship cautions against equating a satisfactory overall grade with evidence of every graduate capability (Gamage et al., 2023). This issue is particularly important when team projects that have been completed by students and which are technically complex contribute a substantial proportion of the final mark (Rajabzadeh et al., 2022).

The current research utilized anonymized assessment data from one Malaysian undergraduate Civil Engineering Technology cohort studying Asset Management. The study evaluated three curriculum mapped areas of competence: (1) knowledge and standard procedures for maintenance; (2) the application of suitable engineering methods and tools to develop maintenance programs; and (3) communication of maintenance-management processes. Instead of separating these data into individual classes, it treats them as a single group.

The research questions were: (i) to what degree do normalized performances vary amongst the three competencies measured; (ii) to what degree will the overall grade mask poor performance in any individual competence area; and (iii) which assessment refinements warrant priority in future course offerings? This paper contributes to the literature by providing an example of a transparent and replicable course level evaluation of student performance. Additionally, it demonstrated how observed scores can differ from claims regarding students' actual proficiency or the potential instructional impact.

LITERATURE REVIEW

Constructive alignment proposes that assessment should elicit performances matching intended learning, rather than merely reward exposure to content (Ramírez De Dampierre et al., 2024). As for application within an engineering program, Gamage et al. (2023) show that graduate competencies can be directly assessed through course design and direct assessments. Recent engineering-education research similarly illustrates the practical value of explicit assessment mappings and scoring criteria (Ravi, 2023). The above studies together support evaluating students based upon their intended competence(s), versus solely on a cumulative weight of all marks.

The individual engineering skills are also not interchangeable. An integrated framework of engineering capabilities was identified through a systematic literature review conducted by Crossin et al. (2023), to include using relevant knowledge, creating effective solutions, managing tasks, and communicating with others. In terms of maintenance, the interconnectedness of engineering methods, decision-making, optimisation and management practices was outlined by Atolagbe and McNeil (2023). The lifecycle approach reflected in recent infrastructure asset-management research further highlights systematic decision-making; this literature is used here as professional context, not as an instructional or accreditation benchmark (Oh et al., 2023).

Assessment design is important when assessing competency to perform task-based activities. Ramírez De Dampierre et al. (2024) provide an overview of the evidence for active and collaborative learning in engineering education; however, this type of evidence cannot confirm whether a group-based project can measure the individual student's competence on its own (Hanley, 2023). Cantera et al. (2021) find that carefully designed, analytic rubrics can strengthen consistency and feedback, while noting that using a rubric alone does not guarantee valid measurement. Ravi (2023) similarly indicates that the purpose of assessment should assist students in developing their own judgments regarding performance as well as how they develop professional practices. This literature supports the importance of developing a method of collecting data from multiple sources and providing clearly defined standards of what is required to complete a specific project.

A methodology issue is that only by normalizing the different maxima assessment weight values will comparative assessments be interpretable. That is, if a student receives a high composite score, this does not mean they performed well on all of their assessments, especially if one or more of those assessments were

significantly weighted. Additionally, the fact that students receive the same group grade, and/or have very small differences in their individual scores, limits how much information regarding individual student ability can be inferred from the data. This study therefore provides a descriptive account of the observed assessment patterns and their implications for implementing a continued cycle of improvement rather than assessing causal impact of instruction.

METHODOLOGY

Design, Sources and Participants

A retrospective (descriptive) document analysis approach is used. The study was conducted in the Maintenance Planning, Scheduling and Control course of the Bachelor of Civil Engineering Technology (Asset Management) with Honours programme at the Faculty of Civil Engineering, UiTM Sarawak. Source documents included a version of the approved syllabus and the end-of-session carry mark workbook for March-July 2026. The workbook contains two administrative class worksheets. Although these worksheets contain separate student record entries, they have been combined prior to analysis. No separate reporting on individual worksheet results occurred. All 18 students with complete assessment results are included in this analysis. Names and student identifiers were removed from the data set and manuscript. No longitudinal or comparison-cohort records were available.

Assessment-to-Competency Mapping

The syllabus indicates that there are four assessments for an overall total of 100% to be achieved. The assignment and final assessment jointly assessed students' knowledge of maintenance principles, components and standard operating procedures (40%). The group project assessed students' development of a maintenance work programme through planning, execution, monitoring, control and relevant tool use (50%). The presentation assessed students' ability to communicate responsibly and effectively about maintenance management (10%). Each dimension has a corresponding programme-level graduate attribute: engineering knowledge, application of modern engineering tools, or professional communication. Table 1 describes mapping in terms of assessment dimensions without administrative outcome codes.

Table 1. Assessment evidence and mapped graduate attributes

Competency dimension	Assessment evidence	Weight	Mapped graduate attribute
Procedural knowledge	Assignment (20%) + final assessment (20%)	40%	Application of engineering knowledge
Applied programme development	Group project	50%	Appropriate engineering methods and modern tools
Professional communication	Presentation	10%	Effective engineering communication

Data Processing and Descriptive Analysis

A dimension score was determined for each student as 100 × (the earned mark(s) in each mapped component / the maximum possible mark(s) for that/those components). Hence, the use of knowledge in this study utilized the assignment plus final total, divided by 40; applied program development utilized project marks divided by 50; and communication utilized presentation marks divided by 10. The overall performance was the same as the sum of the four original weighted component marks. The cohort was summarized with mean, sample standard deviation, minimum, maximum and counts of students scoring at least the illustrative 70% reference value. This reference was used for illustration only; it was not specified as an institutional or accreditation standard in the supplied syllabus.

The differences between the three means were quantified as a difference in percentage points. Since this data set is a census of one course offering with assessment of each student on all components (including group

project), hypothesis testing, explanations for causality, or statements regarding larger student population were also avoided. Workbook totals were verified to equal the total of individual component marks prior to analysis.

Normalized competency means were compared descriptively with the overall weighted course grade; no additional weights were assigned to the competency profile. This prevents counting the 50% project score twice. Scores were interpreted only as evidence from their designated assessment tasks: the written components reflect performance on the knowledge assessment, the project captures performance on a shared applied task, and the presentation reflects assessed communication in that setting. Item-level scores, individual practical demonstrations, assessor-specific ratings and moderation records were unavailable; individual tool proficiency and inter-rater reliability therefore could not be established.

RESULTS AND ANALYSIS

Performance Across Mapped Competencies

The overall composite course mean was 78.17% (standard deviation = 2.42, Observed Range = 71.90-82.00). When broken down into its three parts, the results were somewhat disparate (see Table 2), as shown below. Participants' knowledge concerning maintenance had the highest average at 90.06%, while their ability to communicate professionally averaged at 84.78%. Students did least well on applied program development, with a mean score of 67.33% and the narrowest spread of scores (all of the observed scores were between 65.40% and 69.40%).

Table 2. Descriptive results for the combined cohort (n = 18)

Assessment dimension	Mean (%)	SD	Observed range (%)	At least 70%*
Procedural knowledge	90.06	4.84	77.50–96.75	18/18
Applied programme development	67.33	1.41	65.40–69.40	0/18
Professional communication	84.78	2.67	82.00–92.00	18/18
Overall composite	78.17	2.42	71.90–82.00	18/18

*The 70% reference value is illustrative, not an officially documented competency-attainment target.

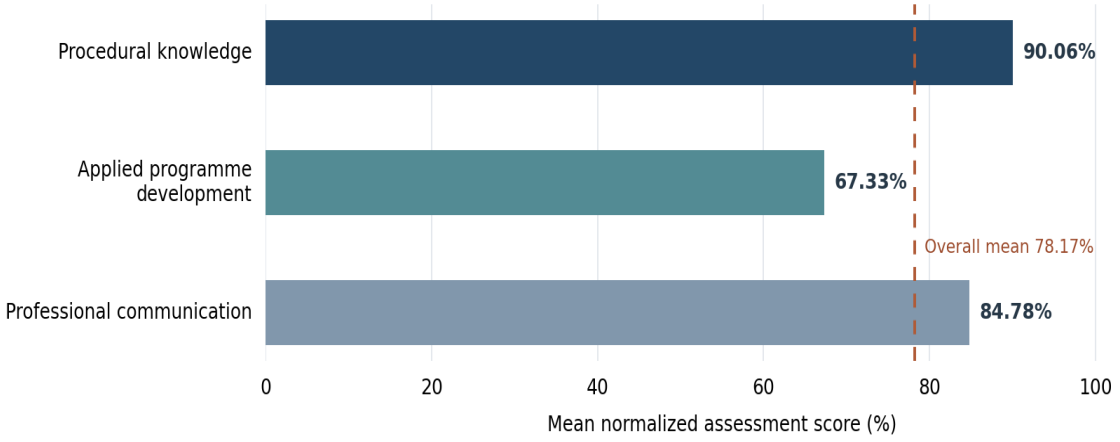


Figure 1. Mean scores for three curriculum-mapped competency dimensions. The dashed line shows the overall weighted course mean, not a competency threshold.

The Masking Effect of Overall Achievement

Based on the reported means, the difference between procedural knowledge and applied programme development was 22.73 percentage points. In addition, the difference between professional communication and

applied programme development was 17.45 percentage points. Although the average mark for the applied domain was at 67.33%, each student received a total score exceeding 70%. Under the purely illustrative 70% reference, no project-derived score reached that value, whereas every knowledge, communication, and overall mark did. This is an example of the descriptive nature on how compensating strengths can lead to good overall grades even though there is one area that has been underdeveloped within the assessment process. However, this does not demonstrate that students have failed to meet an officially defined target.

The group project carried half of the course grade, but its normalized distribution was compressed to a four-point range. Thus, the compression of the data indicates that there is little variation in the student's recorded group project work. However, this does not provide insight as to if an individual student did or did not possess proficiency in using modern tools: group-based assessment could be influenced by how difficult each group's tasks were to complete; what they were being assessed on; and the rubric used when assessing their work. In addition, while the strong and varied results from the knowledge test do reflect some aspects of what was evaluated through written tests, they may not represent all aspects of a student's professional knowledge.

What the Recorded Score Distributions Can and Cannot Establish

Only five distinct project scores were recorded across the 18 students. By comparison, the assignment, presentation and final examination sections showed greater variation in the recorded scores. This limited distribution raises questions about assessment sensitivity and rubric consistency; the workbook cannot distinguish shared grading from restricted score use. Without criterion-level or independently rated practical evidence, the project means cannot verify an individual student's proficiency with engineering tools. High marks on written assessments represent success in that specific task; they cannot automatically translate to an unfamiliar maintenance case. Therefore, the findings should be interpreted as a profile of the Evidence of Assessment, highlighting areas that warrant further investigation.

DISCUSSION

Interpreting the Observed Competency Pattern

The central finding is not merely the lower mean score in one competency dimension but the variation in performance across the assessed graduate attributes. The overall mean obscures a lower applied-project score; however, no institutional competency threshold was documented. Since the assessments have been weighted differently and they measure different student behaviors, only reporting the students' final mark results in removing important data for specific course development. This accords with the rationale for constructive alignment and explicit assessment mapping (Gamage et al., 2023; Ramírez De Dampierre et al., 2024).

Although applied work-programme development incorporates both technique selection and logical activity sequencing along with resource allocation, monitoring, and effective application of engineering tools, weaknesses in the corresponding project-score can be due to difficulties in some or all of the above operations; that is, the scores by themselves cannot identify if there was difficulty in only one area. Literature on asset management stresses that these are interdependent technical and managerial decision-making processes (Atolagbe & McNeil, 2023; Oh et al., 2023). Therefore, another layer of measurement is required before interpreting the results as a specific deficiency in using software skills.

A Practicable Improvement and Re-Assessment Cycle

First, an analytic project rubric should define observable criteria for task breakdown, activity dependencies, resource allocation, schedule updates, risk and constraint management, and interpretation of engineering-tool outputs. The use of analytic descriptors along with example assignments may improve the reliability of assessment and provide more actionable feedback to learners (Cantera et al., 2021). Second, a brief, individually scored practical demonstration or oral defence should complement the group mark. For example, each student could adjust a task sequence or justify resource allocation using the selected tool. The third stage is developing routine practice checkpoints in advance of the completion of the final project. If a student has misinterpreted some aspect of the digital process or control logic, they will have ample opportunity to correct

that misconception by the time the project deadline arrives. The evidence for active collaborative learning supports this formative approach, however it does not indicate the magnitude of improvement within this particular cohort (El-Thalji, 2025; Ramírez De Dampierre et al., 2024).

Fourthly, a full communications evaluation would be required that includes the manner in which the student presents their work as well as the reasons they used certain technical approaches and answers to professional type questions. This responds to the interconnected nature of engineering competencies identified by Crossin et al. (2023). Lastly, an iteration of the course should have analytic criteria that are comparable to those used earlier, and report normalized scores for each offering along with demonstrations that were observed independently. Any comparisons between offerings will need to consider differences in challenge level of the assessments used, marking criteria and cohort composition.

Implications for Future Evaluation

Subsequent offerings should retain the three competency dimensions while collecting criterion-level project scores and independently assessed individual practical demonstrations. Two trained assessors could independently score a sample of project outputs, document their agreement and resolve differences through moderation. A locally approved competency threshold should be defined before assessment; attainment rates could then be reported against that benchmark. Using comparable tasks and rubrics across larger, successive cohorts could support cautious evaluation of formative interventions, accounting for assessment difficulty and cohort composition. Such a design would address the distinction between assessed scores and demonstrated competence discussed in the literature (Cantera et al., 2021; Crossin et al., 2023).

CONCLUSION

The assessment results of the students (N = 18) for their procedural knowledge and professional communication were very high (90.06% and 84.78%, respectively), and the applied programme development score from group-project-based assessments was lower (67.33%). These differences were concealed by the overall weighted course mean of 78.17%. The results indicate there is need for a close evaluation of the current project assessment process and how additional direct, personal evidence of student's utilization of specific engineering tools could be included. The numbers, however, do not constitute proof that students are not proficient or have not met an approved standard. When presented at the competency level, with well-developed rubrics, and with repeated measurements, mark sheets can become a thoughtful, but practical, tool for enhancing the quality of assessments.

REFERENCES

1. Atolagbe, B., & McNeil, S. (2023). Transportation Asset Management Decision Support Tools: Computational Complexity, Transparency, and Realism. Infrastructures, 8(10), 143. https://doi.org/10.3390/infrastructures8100143
2. Cantera, M. A., Arevalo, M.-J., García-Marina, V., & Alves-Castro, M. (2021). A Rubric to Assess and Improve Technical Writing in Undergraduate Engineering Courses. Education Sciences, 11(4), 146. https://doi.org/10.3390/educsci11040146
3. Crossin, E., Richards, J. I., Dart, S., & Naswall, K. (2023). A taxonomy of common engineering activities and competencies. Australasian Journal of Engineering Education, 28(2), 181–193. https://doi.org/10.1080/22054952.2023.2214454
4. El-Thalji, I. (2025). Boosting Active Learning Through a Gamified Flipped Classroom: A Retrospective Case Study in Higher Engineering Education. Education Sciences, 15(4), 430. https://doi.org/10.3390/educsci15040430
5. Gamage, K. A. A., Jeyachandran, K., Dehideniya, S. C. P., Lambert, C. G., & Rennie, A. E. W. (2023). Online and Hybrid Teaching Effects on Graduate Attributes: Opportunity or Cause for Concern? Education Sciences, 13(2), 221. https://doi.org/10.3390/educsci13020221
6. Hanley, K. J. (2023). Group allocation based on peer feedback. European Journal of Engineering Education, 48(2), 284–299. https://doi.org/10.1080/03043797.2022.2106191

7. Oh, W., Cho, C., & Lee, M. (2023). Life Cycle Cost Method for Safe and Effective Infrastructure Asset Management. Buildings, 13(8), 1983. https://doi.org/10.3390/buildings13081983
8. Rajabzadeh, A. R., Long, J., Saini, G., & Zeadin, M. (2022). Engineering Student Experiences of Group Work. Education Sciences, 12(5), 288. https://doi.org/10.3390/educsci12050288
9. Ramírez De Dampierre, M., Gaya-López, M. C., & Lara-Bercial, P. J. (2024). Evaluation of the Implementation of Project-Based-Learning in Engineering Programs: A Review of the Literature. Education Sciences, 14(10), 1107. https://doi.org/10.3390/educsci14101107
10. Ravi, M. (2023). Evolving trends in student assessment in chemical engineering education. Education for Chemical Engineers, 45, 151–160. https://doi.org/10.1016/j.ece.2023.09.003