



INTERNATIONAL JOURNAL OF
MODERN EDUCATION
(IJMoe)
www.gaexcellence.com/ijmoe



PROFILING INNOVATIVE BEHAVIOUR LEVELS AMONG POLYTECHNICS ENGINEERING STUDENTS USING RASCH MEASUREMENT

Aede Hatib Musta'amal @ Jamal^{1*}, Nor Aisyah Che Derasid², Nornazira Suhairom³,
Norzanah Rosmin⁴, Nazaruddin Mohtaram⁵

¹Faculty of Educational Sciences and Technology, Universiti Teknologi Malaysia, (UTM), Malaysia
 aede@utm.my  <https://orcid.org/0000-0002-6271-6536>

²Faculty of Educational Sciences and Technology, Universiti Teknologi Malaysia, (UTM), Malaysia
 noraisyahchederasid@gmail.com  <https://orcid.org/0000-0002-4511-5210>

³Faculty of Educational Sciences and Technology, Universiti Teknologi Malaysia, (UTM), Malaysia
 p-nazira@utm.my  <https://orcid.org/0000-0003-4365-0540>

⁴Faculty of Electrical Engineering, Universiti Teknologi Malaysia, (UTM), Malaysia
 norzanah@utm.my  <https://orcid.org/0000-0002-4084-203X>

⁵Department of Mechanical Engineering, Politeknik Merlimau, Melaka, Malaysia
 nazaruddin@pmm.edu.my  <https://orcid.org/0009-0000-8632-0328>

*Corresponding Author

Article Info:

Article history:

Received date: 03.08.2026
Revised date: 13.08.2026
Accepted date: 23.09.2026
Published date: 28.09.2026

To cite this document:

Jamal, A. H. M., Derasid, N. A. C., Suhairom, N., Rosmin, N., & Mohtaram, N. (2026). Profiling Innovative Behaviour Levels Among Polytechnics Engineering Students Using Rasch Measurement. *International Journal of Modern Education*, 8(31), 1762-1782.

Abstract:

Final-year engineering projects offer opportunities to observe how students recognise problems, develop ideas, and put solutions into practice. This study profiled innovative behaviour among 153 final-year diploma students from three Malaysian Premier Polytechnics, representing Civil, Electrical, and Mechanical Engineering. Students were assessed using the previously developed and validated 26-item Innovative Behaviour Assessment Rubric. Rasch person measures were used to describe overall performance, construct sample-referenced profiles, and examine differences by gender, department, and polytechnic. The mean person measure was 0.02 logits (SD = 2.88). Person reliability was 0.95 and separation was 4.55, indicating approximately six distinguishable ability strata. Using predefined z-score boundaries of ± 0.5 , 49 students (32.0%) were classified as High, 54 (35.3%) as Moderate, and 50 (32.7%) as Low. The distribution was therefore relatively balanced, with a slightly larger Moderate group. A sensitivity analysis using ± 1.0 boundaries placed more students in the Moderate category, showing that the profiles depended on the classification rule. On the common 26-item Wright Map, several of the most difficult criteria concerned advanced problem recognition,

particularly the clarification and justification of project objectives. Several implementation criteria were less difficult. Subgroup tests detected no statistically significant differences by gender, department, or polytechnic; these results do not establish equivalence between groups. The findings identify specific behaviours that may warrant closer attention during project supervision, especially problem framing and objective refinement. The profiles can inform formative feedback and curriculum planning, provided that High, Moderate, and Low are interpreted relative to this sample rather than as externally established proficiency standards.

DOI: 10.35631/IJMOE.831102

Keyword:

Engineering Education, Innovative Behaviour, Performance Assessment, Rasch Measurement, TVET



© The authors (2026). This is an Open Access article distributed under the terms of the Creative Commons Attribution (CC BY NC) (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact ijmoe@gaexcellence.com.

Introduction

Innovation has become a central competency in engineering education as graduates are increasingly expected to respond to complex problems, generate feasible ideas, and implement practical solutions. In technical and vocational education and training (TVET), this competency is particularly important because students' learning outcomes are closely connected to industry readiness, applied problem-solving, and the ability to translate knowledge into practical solutions. Within Malaysia's Premier Polytechnics, final-year projects provide an authentic context for students to demonstrate innovative behaviour through the identification of problems, generation of ideas, and implementation of solutions.

Research on innovation in engineering and TVET education has addressed teaching approaches, student creativity, and the development of assessment instruments. A related question concerns the performance these instruments reveal: which innovative behaviours do students demonstrate, and where do they need further support? A total project score offers limited guidance on these questions. Performance profiles can show how students are distributed across relative ability levels and identify the criteria they find difficult. This study applies an existing rubric to generate such profiles, with the aim of informing final-year project supervision and curriculum planning.

The study examined overall innovative behaviour, its distribution across performance profiles, and variation across rubric dimensions and student groups. It addressed four research questions:

- What is the overall level of innovative behaviour demonstrated by engineering students in Malaysia's Premier Polytechnics?
- How are students distributed across high, moderate, and low innovative behaviour performance profiles?

- Which dimensions of innovative behaviour are most and least demonstrated by engineering students?
- Are there significant differences in innovative behaviour across gender, department, and polytechnic?

The study contributes performance evidence from the application of a previously validated rubric in final-year engineering projects. It identifies patterns in Problem Recognition, Idea Generation, and Idea Implementation, and relates student ability to the difficulty of individual criteria. This application shows how Rasch measures can support decisions about feedback and supervision. The contribution lies in interpreting students' observed performance within the participating polytechnics; the present analysis does not repeat the earlier instrument-development study.

Literature Review

Innovative Behaviour in Engineering Education

Innovative behaviour in engineering education involves the work that leads to a solution as well as the completed output. Students need to recognise a relevant problem, consider possible responses, and develop a solution within technical and contextual constraints. Design thinking, project-based learning, and challenge-based learning give attention to these processes through problem identification, ideation, prototyping, testing, and refinement (Nguyễn et al., 2025; Sukacké et al., 2022; Doulougeri et al., 2024). Assessing innovative behaviour therefore requires evidence of how students develop and apply their ideas throughout a project.

The three IBAR dimensions correspond to observable activities in final-year engineering projects. Problem Recognition concerns how students identify needs, examine constraints, and define a problem. Idea Generation addresses the production, comparison, and refinement of possible solutions. Idea Implementation concerns the application of those ideas through development, testing, and improvement. Project work brings these activities together over an extended period and requires students to draw on technical knowledge while making and explaining design decisions. Studies of practice-oriented and project-based learning have examined their relationships with engineering identity, future readiness, and learning in project management (Ju & Zhu, 2023; Naseer et al., 2025; Afzal & Tumpa, 2025). These settings provide a useful context for examining innovation through students' work, although the opportunities and demands of individual projects may differ.

Performance-Based Assessment in Technical and Vocational Education

TVET assessment places considerable emphasis on what students can demonstrate in practice. Students must apply their knowledge to tasks that reflect the skills and responsibilities of their field. Competency-based assessment addresses this requirement by linking observable performance to intended learning outcomes and professional practice (Yusop et al., 2022). The assessment evidence should therefore show how students use their knowledge, not simply whether they can recall it.

Practical tasks, projects, and products allow educators to examine the application of learning. This relationship between applied learning and competency assessment is reflected in TVET literature on assessment approaches and project-oriented learning (Rahman et al., 2022; Yusop

et al., 2022). An analytic rubric makes the criteria for judging performance explicit and describes what different levels of performance look like (Brookhart, 2013). For innovative behaviour, those criteria can direct attention to how a student frames a problem, evaluates an idea, or tests a proposed solution. The rubric also gives supervisors a shared basis for feedback. Consistent scoring still depends on how raters interpret and apply its descriptors.

Rasch Measurement for Student Performance Profiling

Rubric scores are ordinal: the categories are ordered, but equal numerical differences between scores do not necessarily represent equal differences in ability. When the data adequately fit the model, Rasch analysis supports interval-level interpretation by estimating person ability and item difficulty on a common logit scale (Tennant & Küçükdeveci, 2023; Tesio et al., 2023). This provides a way to examine variation in innovative behaviour across the rubric's observable indicators. Applications in physics education and student engagement illustrate the use of Rasch measurement for interpreting educational assessments (Planinic et al., 2019; Kuswanto et al., 2022).

A Rasch person measure locates a student on the construct defined by the assessment items. Item estimates indicate the relative difficulty of demonstrating each behaviour (Tesio et al., 2023). Person measures may also be grouped for reporting, provided that the classification rule is stated clearly. In this study, High, Moderate, and Low denote positions relative to the sample mean. These categories aid interpretation but are distinct from the Beginning, Developing, and Proficient descriptors used to score individual rubric criteria.

A Wright Map displays person ability and item difficulty on the same scale, allowing the reader to examine both the range of performance and the targeting of the assessment (Tesio et al., 2023). For IBAR, it can identify which criteria are relatively demanding and whether the item set covers the observed range of student ability. The ordering describes relative difficulty within the assessment; it does not, by itself, establish the sequence in which a student develops innovative behaviour. The present study uses this information alongside performance profiles to interpret final-year engineering students' assessment results.

Methodology

This study employed a quantitative performance-based assessment design using Rasch measurement to map the innovative behaviour profile of engineering students in Malaysia's Premier Polytechnics. Since the Innovative Behaviour Assessment Rubric had been validated in a previous phase, the present study focused on applying the rubric to profile students' innovative behaviour in authentic final-year project contexts rather than re-examining rubric development or validation.

Research Design

Innovative behaviour was assessed through observable project performance. This design allowed the study to describe overall ability, identify patterns across rubric dimensions, and compare student groups by gender, department, and polytechnic. Rasch analysis was used to estimate person measures from the ordinal rubric ratings. These measures formed the basis of the overall performance profiles and subgroup analyses.

Participants

The accessible population comprised 866 final-year diploma engineering students across the three participating Premier Polytechnics. A stratified purposive sampling strategy involving intact classes was adopted to obtain coverage across institution and engineering department while maintaining feasibility for performance-based assessment. The final sample comprised 153 students from Politeknik Sultan Salahuddin Abdul Aziz Shah (PSA) ($n = 53$), Politeknik Ibrahim Sultan (PIS) ($n = 47$), and Politeknik Ungku Omar (PUO) ($n = 53$). The sample represented Civil, Electrical, and Mechanical Engineering departments, although Civil Engineering was not offered at PIS. Final-year students were purposively selected because their project work provided authentic evidence for assessing innovative behaviour across the final-year project process.

The adequacy of the sample size was evaluated using both Rasch-oriented and power-analysis considerations. The primary purpose of the study was to profile students' innovative behaviour using Rasch person measures; therefore, the sample needed to provide sufficient variation in student ability to support stable estimates of person ability, item difficulty, and rating-scale functioning. The final sample of 153 students across three Premier Polytechnics and three engineering departments provided adequate heterogeneity for this profiling purpose. This was further supported by the observed person reliability of 0.95 and person separation of 4.55, indicating high measurement precision across the sampled ability range.

As a supplementary check for the inferential subgroup analyses, statistical power was examined using medium-effect assumptions. For the gender comparison, the achieved sample exceeded the approximate minimum sample size required for detecting a medium effect at $\alpha = .05$ and power = .80. Using the actual gender distribution in this study, the achieved power remained acceptable for detecting a medium gender effect. For the ANOVA-based comparisons involving polytechnic and department, the achieved sample was close to the recommended size for detecting medium main effects, although it was more limited for detecting interaction effects. Therefore, subgroup comparisons were interpreted cautiously and treated as supplementary evidence to the main Rasch-based performance profiling. Overall, the sample size was considered methodologically acceptable for the study's primary profiling aim and adequate for supporting cautious interpretation of subgroup patterns.

Instrument

The Innovative Behaviour Assessment Rubric (IBAR) underwent sequential validation before its full-scale application. Content validation was initially conducted on a 46-item pool by eight experts with relevant expertise in engineering education, final-year project supervision, innovation, curriculum development, rubric design and psychometrics. Item relevance was evaluated using the Item-Level Content Validity Index (I-CVI), Scale-Level Content Validity Index (S-CVI/Ave) and modified kappa. The overall S-CVI was 0.85, with dimensional S-CVI values of 0.83 for Problem Recognition, 0.80 for Idea Generation and 0.91 for Idea Implementation. Following expert review and item refinement, the content-validation phase produced a 35-item rubric (Che Derasid et al., 2025). The rubric refinement involved two main stages. First, the 35-item pool produced from content validation was reviewed again by experts to examine item clarity, contextual suitability, descriptor quality, and logical progression across performance levels. Items that were overlapping, unclear, or less strongly aligned with the construct were revised, combined, or removed, reducing the rubric to 30 items. Second, the 30-

item rubric was piloted and analysed using the Many-Facet Rasch Measurement model to examine item functioning, rater consistency, rating-scale performance, and fairness. Based on the psychometric evidence and practical considerations, the rubric was further refined into the final 26-item IBAR with a three-level rating scale: Beginning, Developing, and Proficient.

Table 1: Structure of the Innovative Behaviour Assessment Rubric

Dimension	Number of Items	Main Focus
Problem Recognition	13	Problem finding, fact finding, problem definition, user needs, and objective clarification
Idea Generation	7	Idea exploration, solution evaluation, novelty, feasibility, and user-centred design
Idea Implementation	6	Action planning, resource use, prototyping, functionality, adaptability, and user alignment

Rater consistency had been examined during the prior calibration phase using Many-Facet Rasch Measurement with six raters. Individual rater fit statistics were generally close to the model expectation, while observed exact agreement was 59.7% compared with 47.6% expected agreement. Rater separation was 1.42 with separation reliability of 0.67, and the fixed chi-square test indicated some systematic variation in rater severity, $\chi^2(5) = 18.0$, $p < .05$. Accordingly, these statistics were interpreted as evidence that raters generally applied the rubric in model-consistent ways while still exhibiting some differences in scoring severity. Rater briefing, calibration exercises, a scoring manual and continued access to the researcher were therefore used during full-scale administration to promote consistent interpretation of the performance criteria.

Data Collection

Data were collected from Jun 2025 to November 2025, spanning Academic Session I 2024/2025 and Academic Session II 2025/2026 in the Malaysian polytechnic academic calendar. The two-semester administration followed the natural progression of the same student cohort throughout the final-year project process. During Academic Session I 2024/2025, when the participants were in Semester 4, the Problem Recognition dimension was assessed during the problem exploration and project-definition stage. The same students were subsequently followed into Semester 5 during Academic Session II 2025/2026, when the Idea Generation and Idea Implementation dimensions were assessed as the projects progressed towards solution development, prototyping, testing, and completion.

Formal permission was obtained from the relevant authorities and participating institutions before data collection. The assessment was coordinated with Heads of Department, Unit Penyelidikan, Inovasi dan Komersial coordinators, and final-year project coordinators to ensure that rubric administration was aligned with the students' project milestones. Trained supervisors assessed student performance using multiple sources of evidence, including project reports, logbooks, prototypes, and observed project activities. Rater briefing and calibration were conducted through face-to-face and online sessions, supported by a rater manual to promote consistent rubric application.

Data Analysis

Data were analysed using WINSTEPS Rasch Software version 5.11.1, IBM SPSS Statistics version 29, and Microsoft Excel 365 version 2604. Rasch analysis was first conducted to generate person ability measures from the 26 rubric items. These measures provided calibrated logit estimates of students' innovative behaviour and addressed the research question on overall performance level.

Next, Rasch person logits were standardised into z-scores to develop performance profiles. Students with z-scores above +0.5 standard deviation were classified as High, those between -0.5 and +0.5 standard deviation as Moderate, and those below -0.5 standard deviation as Low. This classification was applied to the overall innovative behaviour profile and to dimension-level profiles for Problem Recognition, Idea Generation, and Idea Implementation.

To facilitate educational interpretation, Rasch person measures were standardised as z-scores and classified into three relative performance profiles. Students with $z > +0.5$ were classified as High, those with $-0.5 \leq z \leq +0.5$ as Moderate, and those with $z < -0.5$ as Low. The ± 0.5 SD boundaries were specified a priori as a sample-referenced descriptive rule to distinguish students performing above, around, or below the sample mean while retaining an interpretable central band. This classification was used to support diagnostic profiling of students' innovative behaviour and to help translate Rasch logits into categories that are more meaningful for educators and curriculum stakeholders. Therefore, the categories should not be interpreted as fixed national proficiency standards or universal Rasch cut-off points.

Dimension-level analysis was conducted to identify which aspects of innovative behaviour were most and least demonstrated. The Wright Map was generated in Winsteps using all 26 items to compare student ability measures and item difficulty estimates on the same logit scale. This supported interpretation of both student performance distribution and the hierarchy of innovative behaviour indicators.

Subgroup comparisons were conducted using SPSS. An independent samples t-test was used to compare male and female students, while two-way analysis of variance was used to examine differences across polytechnic and department, including their interaction effect. Assumptions of normality and homogeneity of variance were checked before analysis. Statistical significance was interpreted with effect sizes, using Cohen's d for gender comparison and partial eta squared for polytechnic and department comparisons. Microsoft Excel was used to prepare descriptive statistics and visual displays, including frequency distributions, bar charts, and radar charts.

Validity and Methodological Rigor

The assessment design linked the construct to evidence that students produced during their projects. IBAR supplied explicit criteria and performance descriptors, while briefing and calibration supported raters' interpretation of those descriptors. Rasch analysis related person measures to item difficulty, and the Wright Map helped identify demanding criteria. These procedures supported the interpretation of the profiles, subject to the quality of the ratings and the fit of the data to the model. As contextualised assessment can include variation unrelated to the intended construct, the influence of assessment conditions also warrants attention (Zhai et al., 2021).

Findings

Demographic Profile

Table 2 summarises the 153 participants. PSA and PUO each contributed 53 students, and PIS contributed 47. The sample covered Civil, Electrical, and Mechanical Engineering, with no Civil Engineering group at PIS. There were 95 male and 58 female students. The representation of institutions and departments provided coverage for the study's descriptive comparisons, although subgroup sizes were unequal.

Table 2: Demographic of Respondents

Institution	Department	Male (n)	Female (n)	Total (n)	% of Total
PSA	Civil Engineering	5	14	19	35.85%
	Electrical Engineering	11	7	18	33.96%
	Mechanical Engineering	11	5	16	30.19%
Total PSA		27	26	53	34.64%
PIS	Electrical Engineering	17	9	26	55.32%
	Mechanical Engineering	17	4	21	44.68%
Total PIS		34	13	47	30.72%
PUO	Civil Engineering	9	9	18	33.96%
	Electrical Engineering	14	6	20	37.74%
	Mechanical Engineering	11	4	15	28.30%
Total PUO		34	19	53	34.64%
Overall		95	58	153	100%

Overall Innovative Behaviour Performance

Person measures from the 26-item IBAR had a mean of 0.02 logits (SD = 2.88), ranging from -5.28 to +4.86 logits (Table 3). This spread indicates variation in the innovative behaviour demonstrated by the sampled students. Person separation was 4.55 and reliability was 0.95. The separation index corresponds to approximately six distinguishable ability strata. These strata describe measurement discrimination and should be distinguished from the three categories used to report the performance profiles.

Table 3: Descriptive Statistics of Students' Innovative Behaviour

Summary	Mean	SD	Logit Min	Logit Max	Person Separation	Person Reliability
Person	0.02	2.88	-5.28	4.86	4.55	0.95
Item	0.00	2.74	-4.83	4.99	10.86	0.99

Item estimates ranged from -4.83 to +4.99 logits. Item separation was 10.86 and item reliability was 0.99, supporting the reproducibility of the estimated item ordering within the analysed data. The range included both relatively easy and demanding behaviours. These indices complement the item hierarchy but do not, on their own, establish model fit or validity.

Student Performance Profile

The overall profile used the predefined z-score boundaries of ± 0.5 . High, Moderate, and Low therefore describe students' positions relative to the distribution of person measures in this sample.

Using the predefined ± 0.5 SD classification, 49 students (32.0%) were classified as High, 54 students (35.3%) as Moderate, and 50 students (32.7%) as Low. Because the classification is sample referenced, a sensitivity analysis was conducted using more conservative ± 1.0 SD boundaries. Under this alternative rule, 34 students (22.2%) were classified as High, 94 students (61.4%) as Moderate, and 25 students (16.3%) as Low. The main interpretation changed from a relatively balanced three-way distribution to a more middle-concentrated profile, although the findings still indicate meaningful variation across students' innovative behaviour levels. The results therefore indicate that the performance categories should be interpreted as relative diagnostic profiles rather than naturally occurring or externally standardised proficiency groups.

The distribution was relatively balanced, although the largest proportion of students was in the Moderate category. This suggests that students' innovative behaviour was spread across different ability levels rather than concentrated in a single category.

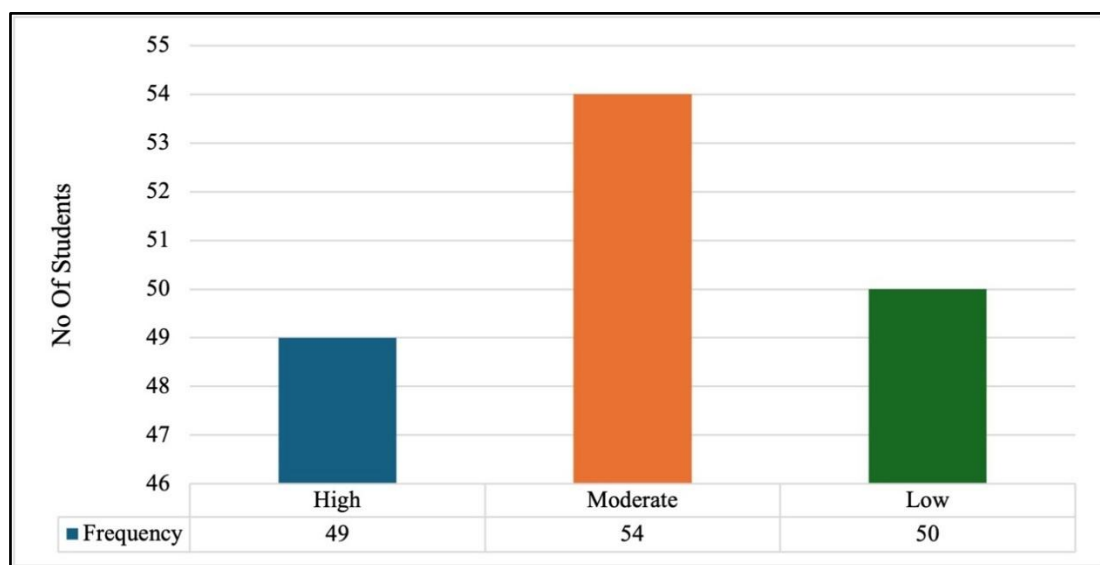


Figure 1: Level of Innovative Behaviour

The profile by polytechnic showed that PIS had the highest proportion of High performers, PSA had the highest proportion of Moderate performers, and PUO had the highest proportion of Low performers. By department, Electrical Engineering recorded the highest number of High performers, Mechanical Engineering recorded the highest proportion of Moderate performers, and Civil Engineering recorded the highest proportion of Low performers. By gender, male students showed a slightly higher proportion in the High category, while female students showed a higher proportion in the Low category. However, these descriptive patterns were further examined through inferential statistical analysis.

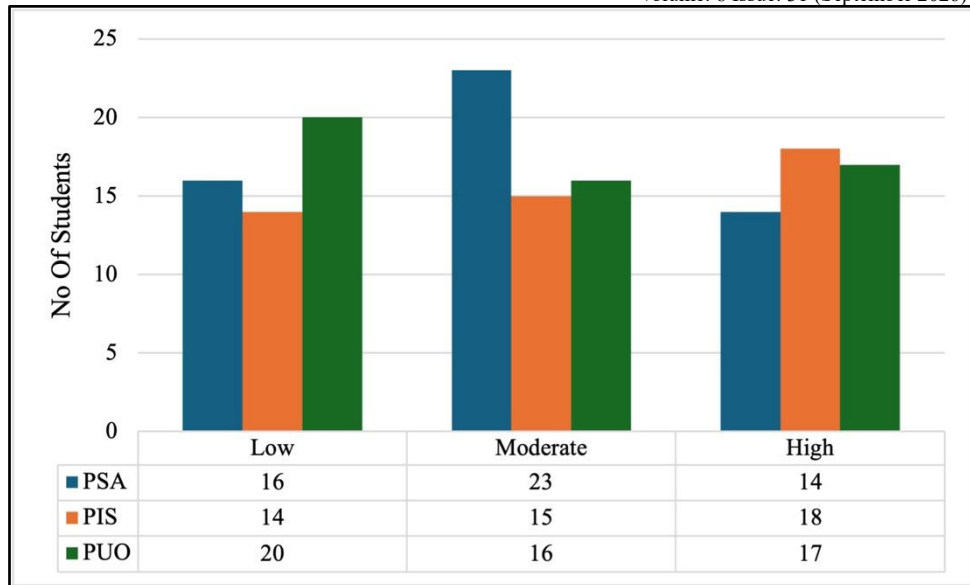


Figure 2: Level of Innovative Behaviour Across Polytechnic

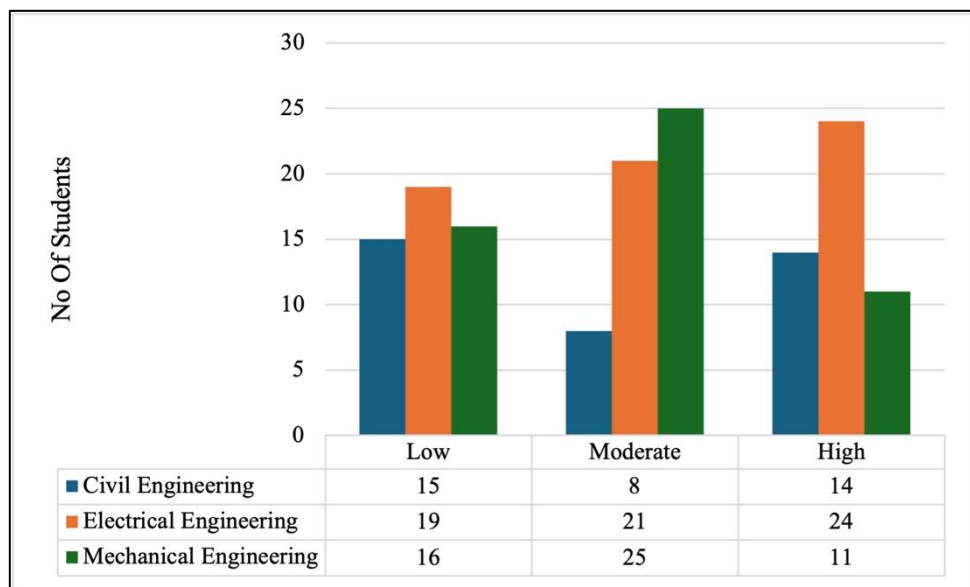


Figure 3: Level of Innovative Behaviour Across Department

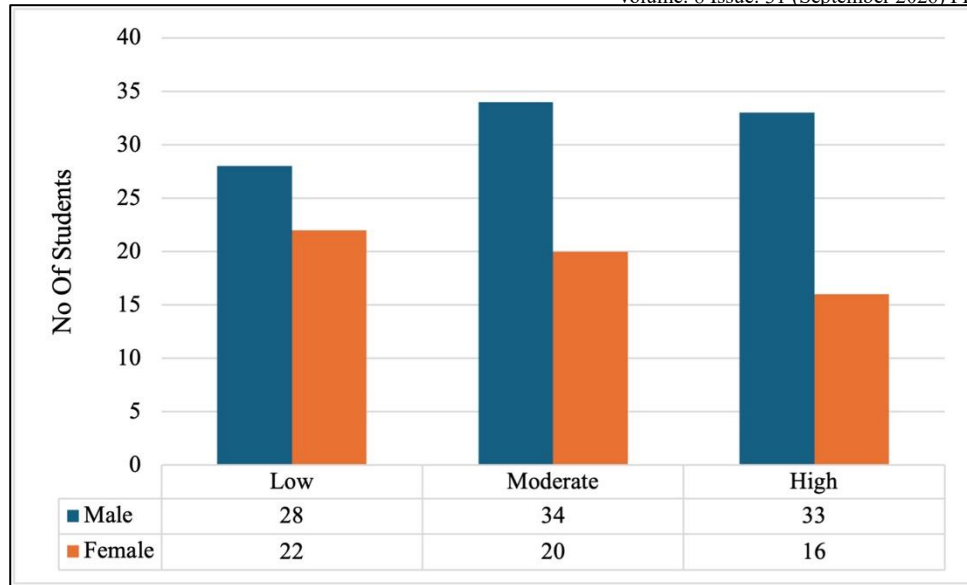


Figure 4: Level of Innovative Behaviour Across Gender

Dimension-Level Performance

Dimension-level analysis was conducted to determine which dimensions of innovative behaviour were most and least demonstrated during final-year project implementation. The analysis focused on three dimensions: Problem Recognition, Idea Generation, and Idea Implementation. As shown in Table 4, all three dimensions demonstrated acceptable person reliability and separation, supporting their use for diagnostic profiling.

Problem Recognition recorded a mean person measure of -0.75 logits, with person separation of 2.95 and reliability of 0.90. Idea Generation recorded a mean person measure of -0.11 logits, with person separation of 2.65 and reliability of 0.88. Idea Implementation recorded the highest mean person measure, with person separation of 2.78 and reliability of 0.89.

Table Error! No text of specified style in document.: Descriptive Statistics By Rubric Dimension

Dimension	Summary	Mean	SD	Logit Max	Logit Min	Person Separation	Person Reliability
Problem Recognition	Person	-0.75	2.75	4.48	-5.91	2.95	0.90
	Item	0.00	2.89	3.72	-5.05	11.86	0.99
Idea Generation	Person	-0.11	5.18	6.97	-10.97	2.65	0.88
	Item	0.00	2.02	2.65	-3.23	6.84	0.98
Idea Implementation	Person	14.61	16.10	41.88	-4.43	2.78	0.89

	Item	0.00	12.09	20.88	-11.89	6.92	0.98
--	------	------	-------	-------	--------	------	------

The reported dimension profiles also differed. Problem Recognition had the largest proportion classified as Low. Idea Generation included a larger High group, while a notable proportion remained in the Low category. Idea Implementation was concentrated in the Moderate category, with relatively few students classified as Low. These patterns describe distributions under the dimension-specific classification rules. Evidence about the relative difficulty of particular behaviours is considered next using the combined 26-item Wright Map.

Wright Map and Item-Person Distribution

The Wright Map places student ability and item difficulty on the common scale from the 26-item analysis. The person mean of 0.02 logits was close to the item mean of 0.00 logits. The highest person measure, +4.86 logits, was close to the most difficult item estimate, +4.99 logits. At the lower end, the minimum person measure of -5.28 logits extended below the easiest item at -4.83 logits. The item locations therefore covered much of the observed ability range. The means and endpoints suggest broadly suitable targeting, although they do not establish that measurement precision was uniform across the scale.

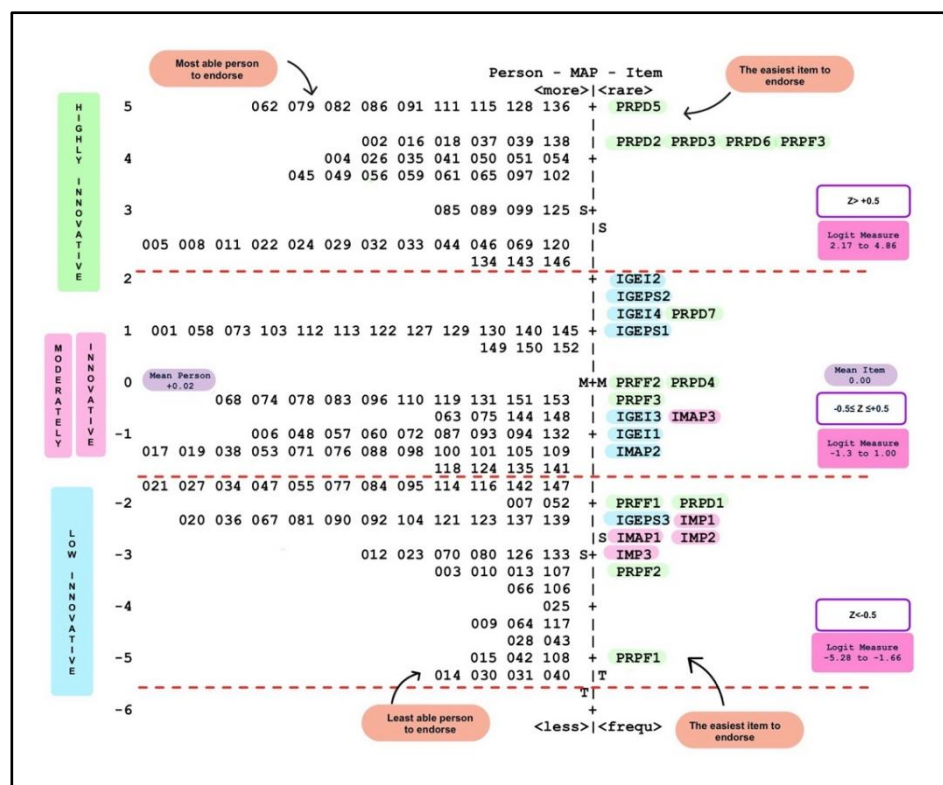


Figure 5: Person Map Item

Several of the most difficult criteria belonged to Problem Recognition. PRPD5, Clarification of Objectives: Timeliness, was the most difficult. It was followed by PRPD3, Clarification of Objectives: Achievability; PRPD6, Integration of New Information; PRPF3, Justification of Problem Framing; and PRPD2, Clarification of Objectives: Measurability. This ordering

identifies advanced problem definition and objective refinement as demanding areas of project performance in the sample.

Idea Generation items generally occupied the middle of the hierarchy, while several Idea Implementation items were less difficult. The easiest item, however, was PRPF1, Limitations of Existing Product, which also belongs to Problem Recognition. Difficulty therefore varied within dimensions as well as across them. Students found some basic problem-recognition behaviours relatively accessible, while the more advanced clarification and justification criteria were demanding.

Subgroup Differences

Subgroup analyses were conducted to determine whether students' innovative behaviour differed significantly across gender, polytechnic, and department. For gender, an independent-samples t-test showed no statistically significant difference between male and female students, $t(151) = 0.96$, $p = .339$. Male students recorded a mean logit score of 0.20, while female students recorded a mean score of -0.27. The effect size was small, with Cohen's $d = 0.16$, indicating a negligible practical difference.

Table 5: Estimated Marginal Means of Innovative Behaviour by Polytechnic and Department

Polytechnic	Department	N	Mean (Logit)	SE	95% CI Lower	95% CI Upper
PSA	Civil Engineering	19	0.50	0.67	-0.82	1.82
	Electrical Engineering	18	-0.39	0.69	-1.75	0.97
	Mechanical Engineering	16	-0.50	0.73	-1.5	1.4
PIS	Electrical Engineering	26	0.54	0.57	-0.6	1.66
	Mechanical Engineering	21	-0.17	0.64	-1.42	1.09
PUO	Civil Engineering	18	-0.82	0.69	-2.18	0.54
	Electrical Engineering	20	0.65	0.65	-0.64	1.94
	Mechanical Engineering	15	-0.47	0.75	-1.96	1.02

Note. SE = Std. Error; 95% CI = 95% Confidence Interval

A two-way analysis of variance was conducted to examine differences by polytechnic and department, including the interaction between both factors. Although some descriptive variation was observed across groups, the differences were not statistically significant. The analysis showed no significant main effect of polytechnic, $F(2,145) = 0.19$, $p = .83$, partial $\eta^2 = .003$, and no significant main effect of department, $F(2,145) = 0.43$, $p = .65$, partial $\eta^2 = .006$. The interaction between polytechnic and department was also not significant, $F(3,145) = 1.06$, $p = .37$, partial $\eta^2 = .022$.

Overall, the subgroup findings indicate that students' innovative behaviour did not differ meaningfully across gender, department, or polytechnic. Although descriptive patterns were observed, the inferential results suggest broadly comparable innovative behaviour performance among engineering students across the participating Premier Polytechnics.

Discussion

Interpreting the Innovative Behaviour Profile

The Rasch-based profile generated in this study shows that innovative behaviour among engineering students in Malaysia's Premier Polytechnics is measurable, differentiated, and unevenly developed. Although students demonstrated innovative behaviour across High, Moderate, and Low levels, the largest proportion was located in the Moderate category, with the remaining students almost equally distributed between the High and Low groups. This pattern suggests that innovative behaviour should not be viewed as a uniform competency among final-year engineering students, but as a continuum of ability in which some students demonstrate stronger innovation-related behaviours, some perform at a developing level, and a substantial group still requires pedagogical support. This finding is consistent with innovation theory, which recognises that innovative behaviour is not evenly demonstrated across individuals; only some learners consistently initiate, generate, and implement ideas at a high level (Anderson et al., 2014; Janssen, 2000; Scott & Bruce, 1994). In TVET engineering education, this is significant because the goal is not only to produce technically competent graduates, but also to develop students who can engage with authentic problems, generate relevant ideas, and implement practical solutions.

The moderate-to-high tendency in the overall profile indicates that final-year project experiences provide meaningful opportunities for students to demonstrate innovation-related competencies. However, the relatively balanced distribution across performance levels also suggests that project exposure alone may not be sufficient to elevate all students to a high level of innovative behaviour. This pattern resembles findings from other Rasch-based educational assessments, where most learners cluster around moderate competency, while smaller groups demonstrate either advanced or weaker performance (Duckett et al., 2019; Planinic et al., 2019). Therefore, the profile should not be interpreted only as evidence of student achievement, but also as diagnostic evidence of unequal readiness. Students at the Low level may require structured scaffolding, those at the Moderate level may need targeted enrichment, and High performers may benefit from more complex innovation challenges. This interpretation aligns with the argument that differentiated instruction is necessary when assessment evidence reveals varied student ability levels (Brookhart et al., 2024).

The broad dispersion of innovative behaviour observed in the present study reinforces the value of performance-based assessment in engineering education. Final-year and capstone projects require students to integrate problem analysis, design decisions, and practical implementation, making rubric-based assessment particularly suitable for capturing competencies that cannot be represented adequately by written tests alone. Previous engineering capstone work has likewise used rubrics to assess graduate attributes such as design, problem analysis, and investigation across proposal and final-project activities. Within Malaysian TVET, Ab Rahman et al. similarly demonstrated the usefulness of rubric-based performance assessment analysed through Many-Facet Rasch Measurement, while also showing that differences in rater severity can persist despite common assessment criteria. Together, these studies support interpreting IBAR as a diagnostic performance-assessment framework rather than merely a mechanism for assigning project marks.

No statistically detectable differences in overall innovative behaviour were found by gender, polytechnic, or department, and the observed effect sizes were small. These findings describe broadly similar group means within the sample. They do not demonstrate that the groups are equivalent. Purposive sampling, unequal subgroup sizes, and limited power for interaction effects restrict that conclusion. The absence of a Civil Engineering group at PIS also makes the institution–department design incomplete. In this dataset, individual profiles and specific rubric criteria offer more direct guidance for feedback than the demographic groupings alone.

Although descriptive variations were observed, the inferential results did not indicate systematic group-based performance gaps. This may reflect the influence of standardised curriculum structures, common final-year project requirements, and shared learning outcomes across Malaysian Premier Polytechnics. The non-significant gender difference is also notable because engineering remains male-dominated in participation; however, the findings suggest that male and female students demonstrated comparable levels of innovative behaviour when assessed through performance evidence. This challenges assumptions that innovation ability differs inherently by gender in technical fields and supports the view that innovation-related performance may be shaped more by learning opportunities and project experiences than by demographic characteristics. This finding aligns with studies reporting no significant subgroup effects in Rasch-calibrated assessments (Duckett et al., 2020; Kuswanto et al., 2022), although it differs from studies that found gender differences in creative self-efficacy or innovation confidence (Tekmen-Araci, 2019). Nevertheless, the absence of significant subgroup differences does not imply that all students require the same type of support. The individual and dimension-level profiles still show meaningful variation, indicating that interventions should be designed around performance needs rather than broad demographic categories.

Problem Recognition as the Critical Developmental Gap

Among the three dimensions, Problem Recognition emerged as the most challenging aspect of innovative behaviour. This finding is important because problem recognition represents the front end of innovation: students must first understand limitations, user needs, contextual constraints, and project objectives before they can generate or implement meaningful solutions. In final-year project work, weak problem recognition can affect the entire innovation process because unclear or poorly framed problems may lead to technically completed but less relevant solutions. The result is consistent with engineering and STEM education literature showing that students often struggle more with problem framing than with later solution development, particularly when dealing with ill-structured or authentic problems (Ernawati et al., 2024; Ibrahim et al., 2024; Ridwan et al., 2025; Setiono & Widiningtyas, 2023).

The item hierarchy provides a deeper explanation of this challenge. The most difficult items were related to advanced problem clarification and objective refinement, including timeliness, achievability, measurability, integration of new information, and justification of problem framing. These criteria require students not merely to identify a problem, but to evaluate whether the problem is realistic, measurable, evidence-informed, and worth solving within project constraints. Such performance demands cognitive and metacognitive control because students must regulate their thinking, evaluate evidence, and justify the relevance of their project direction. Previous studies suggest that students often move too quickly toward solutions before sufficiently analysing root causes or contextual complexity (Tekmen-Araci, 2019; Zhai et al., 2021). This may explain why simpler recognition tasks, such as identifying

limitations of existing products or assessing user needs, were easier than more advanced tasks involving justification and objective clarification.

The relative difficulty associated with Problem Recognition is also consistent with engineering design research showing that problem framing itself is a learned and demanding competency. Svihla et al. used a performance-based assessment of engineering students' design problem-framing skills and found that carefully designed authentic challenges supported growth in attention to underlying needs, contexts of use, and design function. This comparison is particularly relevant to the present finding that advanced clarification, justification, and refinement of project problems were among the more demanding IBAR behaviours. Rather than indicating a simple lack of creativity, the result suggests that students may require stronger scaffolding in the analytical front end of innovation before moving into solution generation and implementation.

This finding suggests that the main challenge in developing innovative behaviour is not simply whether students can complete or implement a project solution, but whether they can first recognise, justify, and frame an authentic problem. In this sense, Rasch-based profiling makes visible a developmental gap that may remain hidden if final-year projects are assessed only through completed products or overall scores. Polytechnic students may be comfortable with practical and technical tasks, but they still require stronger guidance in the earlier analytical stage of innovation. Although final-year project structures already position problem definition and planning as important early-stage activities, the present findings suggest that students need more explicit support in needs analysis, problem validation, project objective refinement, and evidence-based justification. Scaffolding tools such as SMART objectives, root-cause analysis, user interviews, problem validation templates, and formative supervisor feedback could strengthen students' ability to frame problems before moving into ideation and implementation (Amiruddin et al., 2024; Li & Wang, 2021). Without such support, students may produce technically functional projects that are weakly grounded in authentic user, industry, or contextual needs.

In contrast, Idea Implementation appeared more consistently demonstrated, with most students located at the Moderate level and only a small proportion in the Low category. This suggests that polytechnic engineering students are generally able to translate ideas into practical outputs, reflecting the applied and hands-on orientation of TVET learning. However, stronger implementation performance should be interpreted carefully. It does not necessarily mean that students achieved exemplary innovation. Rather, many students appeared able to reach adequate implementation, while fewer demonstrated advanced refinement, systematic testing, optimisation, and iterative improvement. This pattern reflects the "execution gap" in engineering education, where students can produce functional outcomes but may struggle to move beyond minimum viability toward high-quality innovation (McDonald, 2018; Valtonen et al., 2023). Similar findings have been reported in project-based and TVET contexts, where learners often complete prototypes but face difficulties in rigorous validation, multidisciplinary integration, and sustained refinement (Becker, 2020; Marks & Chase, 2019; Setiono & Widiningtyas, 2023). Therefore, implementation should not be assessed merely as product completion, but as the ability to apply ideas through planning, testing, adaptation, feedback integration, and alignment with authentic needs. This reinforces the need for final-year project supervision to include iterative checkpoints, prototype testing, user-centred evaluation, and evidence-based refinement.

Rasch-Based Profiling for TVET Curriculum Improvement

Rasch-based profiling related student performance to the demands of individual rubric criteria. Person reliability of 0.95 and separation of 4.55 indicated that the assessment distinguished approximately six ability strata in the sample. This information complements the three reporting categories, which deliberately simplify the distribution for interpretation. The value of the analysis lies in retaining access to the underlying person estimates and item hierarchy when those broad categories are insufficient. Rasch guidance emphasises the interpretation of measures, item functioning, and the relationship between persons and items (Bond & Fox, 2013; Boone & Staver, 2020). High reliability and separation should be considered alongside those other sources of evidence.

The Wright Map connected the overall profile to particular teaching concerns. Student and item means were close, and the item range covered much of the observed person range. Within that range, advanced clarification of problems and objectives was demanding, while several implementation criteria were less difficult. A supervisor can use this ordering to identify criteria that deserve closer attention across the cohort, then examine each student's work before deciding what support to provide. The hierarchy offers an empirical starting point for curriculum discussion. It does not imply that every student experiences the same difficulty or that item order represents a fixed developmental sequence.

The High, Moderate, and Low profiles offer a compact way to communicate the Rasch results. Their usefulness depends on reporting the classification rule and acknowledging what is lost when continuous measures are grouped. In the sensitivity analysis, the Moderate group increased from 35.3% under ± 0.5 boundaries to 61.4% under ± 1.0 boundaries. This change reflects the reporting rule, not a change in student performance. For educational decisions, the category should therefore be read alongside the person measure, relevant uncertainty, and criterion-level evidence. The profiles are suited to describing relative needs within the sample; they should not be used as universal pass standards.

The findings support closer attention to the reasoning that precedes solution development. Final-year project assessment can make this reasoning visible by requiring evidence of problem exploration, user needs, objective clarification, and the consideration of alternative ideas. Later checkpoints can focus on testing and the use of feedback to improve a prototype. Structured ideation activities, including mind mapping, brainstorming, peer discussion, or TRIZ, are possible teaching approaches, but their suitability and effects would need to be evaluated in the programme concerned. The contribution of the present profile is to identify where that evaluation and instructional attention might begin, particularly in advanced problem recognition.

The study has several limitations. Stratified purposive sampling of intact classes provided coverage of the participating institutions and departments, but the findings cannot be assumed to represent all Malaysian polytechnic engineering students. The performance categories were sample referenced and sensitive to the chosen boundaries. Some rater severity differences remained in the earlier calibration, and those findings do not establish the absence of rater effects during full-scale administration. The relationship between the overall construct and its dimensions also requires further examination, including the comparability of dimension-specific calibrations. Finally, the incomplete institution–department design and the available sample limit subgroup inference, especially for interactions. Future studies could use larger

independently sampled cohorts, examine changes over time, and undertake criterion-referenced standard setting to develop educationally defensible proficiency levels.

Conclusion

This study mapped the innovative behaviour of engineering students in Malaysia's Premier Polytechnics using a Rasch-based performance assessment approach. The findings showed that students generally demonstrated moderate to high levels of innovative behaviour, with approximately six strata of ability identified through Rasch person measures. The common item hierarchy identified advanced problem recognition, particularly the clarification and justification of project objectives, as a demanding area. Several implementation criteria were less difficult, although difficulty also varied within dimensions. Subgroup tests detected no statistically significant differences by gender, department, or polytechnic. The findings can inform formative feedback and curriculum planning by directing attention to specific behaviours observed during final-year projects. Their interpretation remains sample referenced and should draw on the underlying project evidence. Future research may extend the profiling approach to larger samples, different engineering programmes, and longitudinal assessment to examine how innovative behaviour develops over time.

-
- Acknowledgements:** The authors would like to express their sincere gratitude to the Ministry of Higher Education Malaysia, the Department of Polytechnic and Community College Education, and the participating Premier Polytechnics for their support and facilitation throughout this study. The authors also extend sincere appreciation to the lecturers, supervisors, coordinators, expert panel members, and students who contributed to the data collection, rubric validation, and implementation process.
- Funding Statement:** This research received financial support from Fundamental Research Grant Scheme (FRGS) Grant Number FRGS/1/2023/SS107/UTM/02/14. The funding body had no role in the design of the study, data collection, analysis, interpretation of results, or the decision to publish this manuscript.
- Conflict of Interest Statement:** The authors declare that there is no conflict of interest regarding the publication of this paper. All authors have contributed to this work and approved the final version of the manuscript for submission to the International Journal of Modern Education (IJMOE).
- Ethics Statement:** This study was conducted in accordance with ethical research standards. All procedures involving human participants were reviewed and approved by the Research and Innovation Centre, Polytechnics and Community College Education Department, Ministry of Higher Education, approval number JPPKK/PPI/700-1/1 Jld.35 (15) dated 21st February 2025. Informed consent was obtained from all participants prior to data collection. Participation was voluntary, and respondents were assured of confidentiality and anonymity. The data collected were used solely for academic purposes.
- Author Contribution Statement:** All authors contributed significantly to the development of this manuscript. Aede Hatib Musta'amal @ Jamal and Norzanah Rosmin supervised the overall study and contributed to methodology, data interpretation, and critical revision. Nor Aisyah Che Derasid was responsible for conceptualisation, methodology, data collection, data analysis, and original draft preparation. Nornazira Suhairom contributed to the literature review, drafting, and critical revision of the manuscript. Nazaruddin Mohtaram contributed to data collection, validation, review, and editing. All authors read and approved the final version of the manuscript prior to submission.
-

References

- Afzal, F., & Tumpa, R. (2025). Project-based group work for enhancing students learning in project management education: an action research. *International Journal of Managing Projects in Business*. <https://doi.org/10.1108/ijmpb-06-2024-0150>
- Anderson, N., Potočník, K., & Zhou, J. (2014). Innovation and Creativity in Organizations: A State-of-the-Science Review, Prospective Commentary, and Guiding Framework. *Journal of Management*, 40(5), 1297-1333. doi:10.1177/0149206314527128
- Becker, C. R. (2020). *Learn Human-Computer Interaction: Solve human problems and focus on rapid prototyping and validating solutions through user testing*: Packt Publishing Ltd.
- Bond, T. G., & Fox, C. M. (2013). *Applying the Rasch model: Fundamental measurement in the human sciences*: Psychology Press.
- Boone, J. W., Staver, J. R., & Yale, M. S. (2020). *Advances in Rasch analyses in the human sciences*. Cham, Switzerland: Springer. <https://doi.org/10.1007/978-3-030-43420-5>
- Brookhart, S. M. (2013). *How to create and use rubrics for formative assessment and grading*: Ascd.
- Brookhart, S. M. (2024). Educational assessment knowledge and skills for teachers revisited. *Education Sciences*, 14(7), 751. <https://doi.org/10.3390/educsci14070751>
- Che Derasid, N. A., Musta'amal @ Jamal, A. H., Suhairom, N., & Suswanto, H. (2025). Content validity of an innovative behavior rubric for polytechnic engineering students. *International Journal of Evaluation and Research in Education (IJERE)*, 14(6), 4307–4319. <https://doi.org/10.11591/ijere.v14i6.33370>
- Doulougeri, K., Vermunt, J. D., Bombaerts, G., & Bots, M. (2024). Challenge-based learning implementation in engineering education: A systematic literature review. *Journal of Engineering Education*. <https://doi.org/https://doi.org/10.1002/jee.20588>
- Duckett, T. R., Liberatore, M. W., Asogwa, U., Mentzer, G. A., & Malefyt, A. P. (2019). *Assessing the reliability of a chemical engineering problem-solving rubric when using multiple raters*. Paper presented at the 2019 ASEE Annual Conference & Exposition.
- Janssen, O. (2000). Job demands, perceptions of effort-reward fairness and innovative work behavior. *Journal of Occupational and organizational psychology*, 73(3), 287-302.
- Ju, T., & Zhu, J. (2023). Exploring senior engineering students' engineering identity: the impact of practice-oriented learning experiences. *International Journal of STEM Education*. <https://doi.org/https://doi.org/10.1186/s40594-023-00439-2>
- Kuswanto, H., Pratidhina, E., & Rosana, D. (2022). Developing Students' Engagement in Online Learning Questionnaire using Rasch Measurement Model. *KnE Social Sciences*.
- Marks, J., & Chase, C. C. (2019). Impact of a prototyping intervention on middle school students' iterative practices and reactions to failure. *Journal of Engineering Education*, 108(4), 547-573.
- McDonald, K. (2018). Post hoc evaluation of analytic rating scales for improved functioning in the assessment of interactive L2 speaking ability. *Language Testing in Asia*, 8(1), 19.
- Naseer, F., Tariq, R., Alshahrani, H., Alruwais, N., & Al-Wesabi, F. (2025). Project based learning framework integrating industry collaboration to enhance student future readiness in higher education. *Scientific Reports*. <https://doi.org/10.1038/s41598-025-10385-4>
- Nguyễn, L. C., Hoa, H., & Hien, L. H. P. (2025). Integrating design thinking into STEM education: Enhancing problem-solving skills of high school students. *Eurasia Journal of Mathematics, Science and Technology Education*. <https://doi.org/10.29333/ejmste/16084>

- Planinic, M., Boone, W. J., Susac, A., & Ivanjek, L. (2019). Rasch analysis in physics education research: Why measurement matters. *Physical Review Physics Education Research*, 15(2), 020111.
- Rahman, K. A. A., Rozali, M. Z., Samah, N. A., Bakar, M. A. A., Ahmad, N., Gerijih, D. D., & Zakariah, S. H. (2022). Conceptual Model of Video Learning based on Project-Oriented Problem-Based Learning and Competency-Based Education for Technical and Vocational Education. *Journal of Technical Education and Training*. <https://doi.org/10.30880/jtet.2022.14.01.004>
- Ridwan, Sura, S. A., Koerniadi, D., Zaradiva, L., Soeharto, S., Najira, . . . Muhammad, F. (2025). Investigating High School Student Misconceptions: A Rasch-Based Three-Level Diagnostic Evaluation of Osmoregulation and Excretion Systems. *Emerging Science Journal*.
- Scott, S. G., & Bruce, R. A. (1994). Determinants of innovative behavior: A path model of individual innovation in the workplace. *Academy of management journal*, 37(3), 580-607.
- Setiono, I. A., & Widiningtyas, A. (2023). Characteristics of Students' Cognitive Ability on the Hyperopia Concept: Rasch Analysis. *Berkala Ilmiah Pendidikan Fisika*.
- Sukackè, V., Guerra, A., Ellinger, D., Carlos, V., Petroniené, S., Gaižiūnienė, L., Gelabert, S. B., Tallada, A. M., & Brose, A. (2022). Towards Active Evidence-Based Learning in Engineering Education: A Systematic Literature Review of PBL, PjBL, and CBL. Sustainability. <https://doi.org/https://doi.org/10.3390/su142113955>
- Svihla, V., Chen, Y., & Kang, S. “. (2022). A funds of knowledge approach to developing engineering students' design problem framing skills. *Journal of Engineering Education*, 111(2), 308–337. <https://doi.org/10.1002/jee.20445>
- Tekmen-Araci, Y. (2019). Teaching risk-taking to engineering design students needs risk-taking. *Art, Design & Communication in Higher Education*.
- Tennant, A., & Küçükdeveci, A. A. (2023). Application of the Rasch measurement model in rehabilitation research and practice: early developments, current practice, and future challenges. *Frontiers in Rehabilitation Sciences*. <https://doi.org/https://doi.org/10.3389/fresc.2023.1208670>
- Tesio, L., Caronni, A., Kumbhare, D., & Scarano, S. (2023). Interpreting results from Rasch analysis 1. The “most likely” measures coming from the model. *Disability and Rehabilitation*. <https://doi.org/10.1080/09638288.2023.2169771>
- Valtonen, A., Kimpimäki, J. P., & Malacina, I. (2023). From ideas to innovations: The role of individuals in idea implementation. *Creativity and innovation management*, 32(4), 636-658.
- Yusop, S. R. M., Rasul, M., Yasin, R. M., Hashim, H., & Jalaludin, N. (2022). An Assessment Approaches and Learning Outcomes in Technical and Vocational Education: A Systematic Review Using PRISMA. Sustainability. <https://doi.org/10.3390/su14095225>
- Zhai, X., Haudek, K. C., Wilson, C., & Stuhlsatz, M. (2021). *A framework of construct-irrelevant variance for contextualized constructed response assessment*. Paper presented at the Frontiers in Education.