Psychometric Quality Assessment of a 21st Century Skills Instrument in a Design Thinking Context Among Final-Year Polytechnic Students Using the Rasch Measurement Model

Aede Hatib Musta'amal[1,*], Nornazira Suhairom[1], Mohd Safarin Nordin[1], Nor Aisyah Che Derasid[1], Roslaili Anuar[2], Siti Zuraida Maaruf[2] and Rozita Jayus[3]

[1] University of Technology Malaysia, Faculty of Educational Sciences and Technology, Johor Bahru, Malaysia
[2] Universiti Teknologi MARA, Faculty of Education, Puncak Alam, Selangor, Malaysia
[3] Faculty of Business, Economics and Social Development, Universiti Malaysia Terengganu, Kuala Terengganu, Malaysia
*aede@utm.my

Abstract. The development of valid and reliable instruments is essential for evaluating 21st-century skills within innovative pedagogical approaches such as Design Thinking (DT), particularly in Technical and Vocational Education and Training (TVET) contexts. This study aimed to develop and assess the psychometric quality of an instrument measuring 21st-century skills (21CS) and Design Thinking experience among Malaysian polytechnic students engaged in Design Thinking-driven Final Year Projects (FYPs) using the Rasch Measurement Model. A total of 213 final-year engineering diploma students from Premier and Conventional Polytechnics participated in a survey comprising 34 items measuring 21CS and 25 items measuring DT experience. Rasch analysis demonstrated strong psychometric properties for both instruments. For the 21CS scale, item reliability was 0.91 and person reliability was 0.95, with item and person separation indices of 3.10 and 4.41, respectively. The DT experience scale showed satisfactory reliability (item = 0.71; person = 0.94). Group comparisons revealed significant differences by polytechnic type, semester of study, and academic program, with Premier polytechnic and later-semester students demonstrating higher levels of 21CS and DT experience. The findings provide evidence supporting the psychometric quality of the instruments and highlight the effectiveness of Design Thinking pedagogy in fostering essential future-oriented competencies within the TVET context.

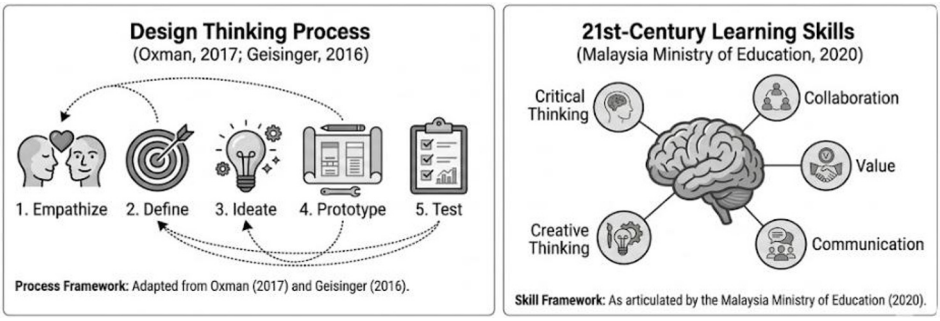
Keywords: 1st Century Skills, Design Thinking, Polytechnics, TVET, Rasch Analysis, Psychometric Quality Assessment.

© The Author(s) 2026
C. Y. Lye et al. (eds.), *Proceedings of the Pacific Rim Objective Measurement Symposium (PROMS 2025)*, Atlantis Highlights in Social Sciences, Education and Humanities 56,
https://doi.org/10.2991/978-94-6239-834-4_14

1 Introduction

Malaysia's Ministry of Education (MOE) has consistently emphasized the integration of 21st-century learning for both teachers and students as a central pillar of national education reform [1]. This emphasis was formally articulated in the Malaysia Education Development Plan (MEDP) 2013-2025, which advocates for learner-centred pedagogies, innovative teaching approaches, and enriched learning environments that foster higher-order thinking and lifelong learning competencies [2]. The overarching goal of this policy direction is to enhance the quality of graduates so that they are competitive at the global level and responsive to contemporary industrial and societal demands. The Ministry's focus on 21st-century learning extends across all educational sectors, including Technical and Vocational Education and Training (TVET). Within the TVET ecosystem, graduate quality is increasingly evaluated not only in terms of technical proficiency but also through the possession of transferable and transversal skills. Employers expect graduates to demonstrate competencies including communication, collaboration, creative and critical thinking, adaptability, and ethical reasoning [3,4,5]. These competencies, commonly referred to as 21st-century skills, are widely recognised as essential for navigating the complexities of Industry 4.0 and the knowledge-based economy [6,7].

Despite strong policy endorsement, previous studies indicate persistent gaps between the intended curriculum and actual classroom practice, particularly in TVET contexts where instructional approaches often remain teacher-centred and assessment-driven [8,9]. Consequently, there is an increasing need for pedagogical models that can authentically embed 21st-century skills within real-world, practice-oriented learning environments. Design thinking has gained prominence as a promising pedagogical approach. Originating from product and industrial design disciplines, design thinking emphasises human-centred problem-solving through iterative processes such as empathising, defining problems, ideating, prototyping, and testing solutions. Empirical studies suggest that design thinking-based learning can effectively support the development of critical and creative thinking, teamwork, and communication skills – key dimensions of 21st-century competencies [12,13].



Source: Oxman (2017), Geisinger (2016), Malaysia Ministry of Education (2020)

Fig 1. Integration of Design Thinking Process and the 21st Century Learning Skills

The theoretical foundation of this study draws on the 21st-Century Learning Skills framework outlined by the Ministry of Education Malaysia, encompassing collaboration, creative thinking, critical thinking, communication, and values. These competencies are mapped onto the Design Thinking process comprising the stages of Empathize, Define, Ideate, Prototype, and Test (Fig 1). The framework positions Design Thinking experience as a pedagogical mechanism through which 21CS are developed during the implementation of Final Year Projects [1,14,15].

1.1 Problem Background

Within TVET and engineering education, design thinking aligns particularly well with project-based and problem-based learning structures, such as Final Year Projects, where students engage with authentic industry or community challenges. However, while design thinking is increasingly adopted as an instructional strategy, there remains a lack of rigorously validated instruments to measure students' acquisition of 21st-century skills within this pedagogical framework, especially in the Malaysian polytechnic context. Addressing this gap is critical to ensuring that policy aspirations translate into measurable learning outcomes.

1.2 Research Objectives

This study was guided by the following objectives:

1. To develop and assess the psychometric quality of an instrument measuring 21st-century skills within a Design Thinking context using the Rasch Model.
2. To examine differences in students' mastery of 21st-century skills and Design Thinking experience across polytechnic types, semesters of Final Year Project implementation, and engineering diploma programs.
3. To determine the practical implications of these differences for curriculum design and institutional development in Malaysian polytechnics.

2 Literature Review

2.1 Polytechnics, 21st-Century Skills, and the Need for Robust Measurement

Polytechnics under Malaysia's Ministry of Higher Education (MOHE) represent a core pillar of the national Technical and Vocational Education and Training (TVET) system, with the mandate to produce skilled and semi-skilled technical graduates to support national economic development. However, rapid advancements in industrial practices, digitalisation, and information technology have significantly transformed workplace demands. Contemporary industries require employees who are not only technically competent but also adaptable, communicative, collaborative, and capable of higher-order thinking in dynamic and uncertain environments.

In this context, the employability of polytechnic graduates increasingly depends on the mastery of soft skills, often conceptualised as 21st-century skills. Despite longstanding recognition of their importance, employer dissatisfaction with graduates' soft skills remains a persistent issue. Early evidence revealed that many employers were dissatisfied with TVET graduates' communication abilities, interpersonal competence, and critical thinking skills [16-17]. More recent studies suggest that these challenges continue to persist, indicating a structural gap between curriculum intentions and graduate outcomes. The framework of 21st-century learning provides a comprehensive guide for addressing this gap by emphasising a set of transferable competencies essential for lifelong learning and employability. These competencies typically include communication, collaboration, critical thinking, creative thinking, and moral or ethical values [18].

Previous studies [19] described this persistent mismatch between graduate capabilities and workplace expectations as the "twenty-first-century gap," identifying deficiencies in oral and written communication, critical thinking, problem-solving, professionalism, teamwork, technological application, and leadership. Empirical evidence from Malaysian polytechnics supports these concerns. For example, [20] reported that a proportion of polytechnic graduates had not fully mastered interpersonal skills upon graduation, while [21-22] found that students' levels of critical and creative thinking remained only moderate.

Challenges in collaborative competence have also been widely reported. [23, 24, 25] observed that many higher education students demonstrate passive collaboration, characterised by listening rather than active communication, negotiation, or shared decision-making. Effective collaboration, however, requires a broader set of abilities, including respect for diversity, emotional regulation, constructive dialogue, and shared responsibility. At the same time, global employer reports consistently rank creativity and critical thinking among the most important skills for innovation and workforce readiness [7, 26]. Beyond cognitive and interpersonal skills, ethical reasoning and values-based behaviour remain critical concerns. [27] reported that ethical behaviour among Malaysian polytechnic engineering students was only moderate, echoing findings by [28], who noted moderate levels of soft skill implementation from students' perspectives. These findings suggest that, despite policy emphasis and curriculum reforms, the holistic development of 21st-century competencies in polytechnics remains uneven.

Given these challenges, educators and institutions are increasingly seeking pedagogical approaches that can authentically integrate multiple dimensions of 21st-century learning. No single instructional model can fully address all aspects of 21st-century skills; therefore, educators are encouraged to adopt flexible, student-centred approaches that align pedagogy, assessment, and real-world application. In the context of polytechnic education, 21st-century skills closely mirror learning outcomes commonly associated with design-oriented disciplines, including innovation, problem-solving, communication, collaboration, and digital literacy [19]. This alignment makes design-based pedagogies particularly suitable for TVET settings.

2.2 Design Thinking as a Pedagogical Environment for 21st-Century Skills

Design Thinking has gained prominence as a structured yet flexible problem-solving approach that integrates analytical and creative thinking. It involves iterative cycles of understanding user needs, defining problems, generating ideas, developing prototypes, and testing solutions through feedback [29]. While Design Thinking does not have a single standardised model, most frameworks share common stages that emphasise human-centred inquiry and experiential learning [30]. This study adopts the Design Thinking model proposed by the Hasso Plattner Institute of Design, which comprises five stages: Empathise, Define, Ideate, Prototype, and Test [31]. Each stage involves specific tasks, interactions, and learning environments that promote collaboration, communication, creativity, critical reflection, and self-directed learning. When embedded within project-based courses such as Final Year Projects, Design Thinking provides an authentic context for students to apply technical knowledge while simultaneously developing 21st-century skills. From an educational perspective, Design Thinking aligns strongly with student-centred learning principles and constructivist theories, encouraging learners to actively construct knowledge through experimentation, reflection, and social interaction [32, 33]. As such, it holds significant potential as a pedagogical environment for addressing persistent gaps in 21st-century skill development among polytechnic students.

2.3 Utilization of Rasch Analysis

Many studies on 21st-century skills rely on raw scores from Likert-scale instruments without adequately examining their measurement properties. The Rasch Measurement Model provides a robust methodological framework for addressing these limitations. Unlike classical test theory, Rasch analysis transforms ordinal survey data into interval-level measures, allowing for meaningful comparisons of student ability and item difficulty [34].

Critically, this study situates its Rasch-based analyses within Messick's (1989) unified validity framework [35], which encompasses multiple inter-related sources of validity evidence including content, structural, substantive, generalisability, external, and consequential aspects. The present study provides primarily structural validity evidence (via Rasch-based dimensionality and fit analyses) and content-related evidence (via systematic item development from the 21CS and DT literature). A complete validity argument would require additional evidence such as convergent/discriminant validity and consequential validity, which are identified as directions for future research.

By integrating Rasch measurement with Design Thinking pedagogy, this study contributes methodologically and substantively to TVET research, offering an evidence-based approach for evaluating skill development aligned with national policy goals and industry expectations.

3 Methods

3.1 Research Design

A quantitative research design was employed using a survey-based instrument development approach.

3.2 Participants

The participants comprised 213 final-year diploma students enrolled in engineering programs at Malaysian polytechnics. Students were drawn from both Premier and Conventional Polytechnics and were involved in Design Thinking implementation as part of their Final Year Projects. Participants were categorized into Semester 1 (initial phase of FYP) and Semester 2 (completion phase).

3.3 Instrument Development

The instrument was developed based on an extensive review of the literature on 21st-century skills and Design Thinking. It consisted of three sections: Section A: demographic information; Section B: 34 items measuring self-perceived 21st-century skills; and Section C: 25 items measuring Design Thinking implementation experience. All items were rated on a five-point Likert scale ranging from 1 (Strongly Disagree) to 5 (Strongly Agree).

3.4 Data Analysis

Rasch analysis was conducted to examine item and person reliability, separation indices, item fit, polarity, and dimensionality. Group comparisons were performed using Rasch-calibrated person measures, supported by Welch's t-tests, one-way ANOVA, and chi-squared tests

3.5 Validation Framework

This study explicitly adopts Messick's (1989) unified validity framework as its theoretical foundation [35]. The Rasch-based analyses reported here provide evidence of internal structure - a recognised and important facet of construct validity - and are complemented by content validity evidence derived from systematic item development grounded in the 21CS and DT literature. Bond and Fox (2015) [36] and Wilson (2005) support the use of the Rasch model as a means of providing construct validity evidence through examination of fit, dimensionality, and item hierarchy. It is acknowledged that further evidence (convergent/discriminant, predictive, and consequential validity) would be needed for a comprehensive validity argument.

4 Results

4.1 Demographic Information

The demographic profile of the study participants reveals a total sample size of 213 respondents. The gender distribution shows a predominance of male participants (n = 160, 75%), while female participants constitute a smaller proportion (n = 53, 25%). In terms of institutional affiliation, the majority are from Politeknik Ibrahim Sultan (PIS), accounting for 60.6% (n = 130), whereas 39.4% (n = 83) are from Politeknik Merlimau Melaka (PMM). Over half of the participants are from the Department of Electrical Engineering (JKE) with 51% (n = 108), followed by the Department of Mechanical Engineering (JKM) at 30.3% (n = 65), and the Department of Civil Engineering (JKA) at 18.8% (n = 40). Students from Semester 4 form the majority (n = 119, 55.8%), while those in Semester 5 comprise 44.2% (n = 94).

Table 1. Demographic Information

Demographic Variable	Category	Frequency (n)	Percentage (%)
Gender	Male	160	75.0
	Female	53	25.0
Polytechnics	Premier	130	60.6
	Conventional	83	39.4
Department	Electrical & Electronics Engineering	108	51.0
	Mechanical Engineering	65	30.3
	Civil Engineering	40	18.8
Semester	4	119	55.8
	5	94	44.2

4.2 Psychometric Quality of the 21st-Century Skills Scale

Research Question 1: What is the psychometric validation of an instrument measuring 21CS in the DT context using the Rasch Model?

Item and Person Reliability and Separation

The analysis demonstrated strong internal consistency of the instrument. The item reliability was 0.91, indicating high replicability of the item hierarchy across different samples, while the person reliability was 0.95, showing acceptable differentiation among respondents based on their ability levels. An item separation index of 3.10 suggested that the sample size was sufficient to confirm the stability of item difficulty estimates. The person separation index of 4.41 supported the instrument's capacity to distinguish between at least three distinct strata of student ability.

Item Fit Statistics

The analysis of the 34 items reveals that most items fall within the acceptable range for both Infit and Outfit MNSQ (0.5 to 1.5) as recommended by Wright and Linacre (1994) [37] for survey-type rating scale instruments. The ZSTD threshold of $|Z| <= 2.0$ follows Bond and Fox (2015) [36], and the Point-Measure Correlation threshold (positive values indicating expected polarity) is referenced to Linacre (2002) [38]. The mean Infit MNSQ of 1.00 and Outfit MNSQ of 1.02 suggest that the items align well with the construct being measured. The observed match percentage of 68.1% further supports the validity of the model.

Unidimensionality

The Rasch model accounts for 43.1% of the raw variance, meeting the commonly accepted threshold for unidimensionality (≥40%). Although the eigenvalue of 5.7 for the 1st contrast exceeds the threshold of 3.0, disattenuated correlations between all cluster pairs ranged from 0.68 to 0.79, with most pairs surpassing the 0.70 criterion typically regarded as evidence of a shared underlying construct (Smith, 2002; Tennant & Conaghan, 2007). The borderline value for Clusters 1–3 ($r = 0.68$) suggests these subsets are largely overlapping, consistent with the theoretically interrelated nature of 21st-century skills. All secondary contrasts (2nd to 5th) yielded eigenvalues between 1.6 and 2.5, indicating no strong secondary factors. Taken together, the evidence supports treating the instrument as essentially unidimensional for measurement purposes, while acknowledging the construct's inherent complexity. Within Messick's framework, this analysis provides structural validity evidence supporting the internal coherence of the instrument.

Table 2. Item Dimensionality for 21CS

Table of STANDARDIZED RESIDUAL variance (in Eigenvalue units)					
		-- Empirical --			Modeled
Total raw variance in observations	=	59.8	100.0%		100.0%
Raw variance explained by measures	=	25.8	**43.1%**		43.5%
Raw variance explained by persons	=	15.8	26.4%		26.7%
Raw Variance explained by items	=	10.0	16.7%		16.9%
Raw unexplained variance (total)	=	34.0	56.9%	100.0%	56.5%
Unexplned variance in 1st contrast	=	**5.7**	9.5%	16.7%	
Unexplned variance in 2nd contrast	=	2.5	4.2%	7.5%	
Unexplned variance in 3rd contrast	=	2.3	3.9%	6.8%	
Unexplned variance in 4th contrast	=	2.0	3.4%	6.0%	
Unexplned variance in 5th contrast	=	1.6	2.8%	4.9%	

Further analysis is conducted to review items contributing to the 1st contrast to see if they cluster around specific constructs (e.g., creativity or ethics).

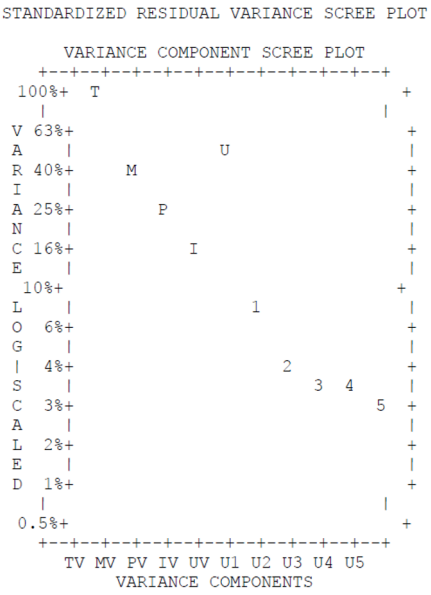


Fig 2. Scree Plot of PCA of Residuals

The Standardized Residual Variance Scree Plot indicates that the Rasch model explains a substantial portion of the total variance, with the first contrast (U1) accounting for less than 5%, suggesting minimal unexplained variance and no strong secondary dimension.

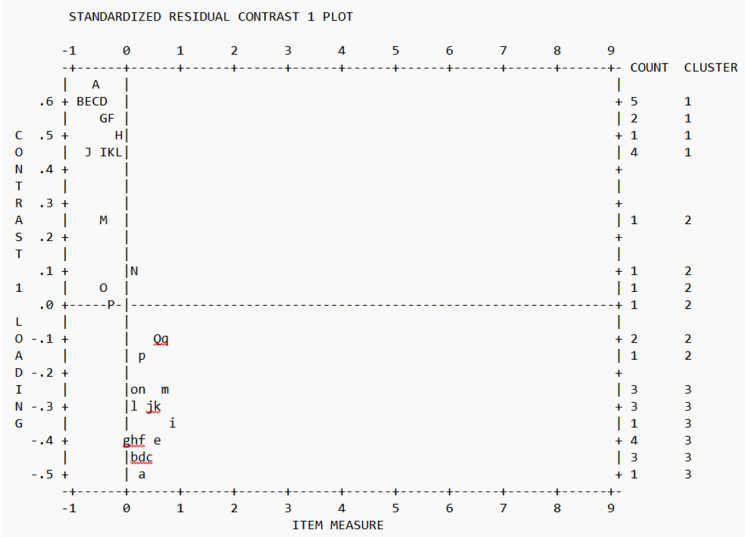


Fig 3. Standardized Residual Contrast 1 Plot

The steep drop in variance after the model component and the low values for subsequent contrasts (U2, U3, etc.) further support the unidimensionality of the instrument. This implies that the items in the scale are largely measuring a single underlying construct, confirming the validity of the measurement model used.

Table 3. Item Clustering for 21CS

Cluster	Loading Range	Items (Code)	Count	Interpretation	Possible Trait Theme
Cluster 1 (Positive)	+0.42 to +0.62	B41, B42, B43, B44, B45, B46, B47, B51–B55 (A–L)	~13	High-loading, harder items—possibly measuring a more advanced or different subskill	Measuring *advanced, individual-level thinking* such as COLLABORATION and VALUES
Cluster 2 (Neutral)	–0.14 to +0.22	B21–B27, B24, B23, B25, B26 (M–P, p, q, Q)	~7	Neutral to low-loading items—likely aligned with primary construct	Aligned with the *primary dimension*; anchor items (CREATIVITY)
Cluster 3 (Negative)	–0.48 to –0.30	B11–B18, B31–B37, B12, B13, B14, B15, B17, B18 (a–o)	~14	Easier, lower-loading items—may represent a simpler subskill	Focused on *social or foundational aspects* like COMMUNICATION and CRITICAL THINKING.

4.3 Properties of the Design Thinking Experience Scale

Research Question 2: What is the psychometric validation of an instrument measuring Design Thinking experience among students engaged in Final Year Project implementation?

Item and Person Reliability and Separation

The Rasch analysis for the Design Thinking Experience scale demonstrated strong internal consistency. The item reliability was 0.71 and the person reliability was 0.94. An item separation index of 1.58 and a person separation index of 3.90 supported the instrument's capacity to distinguish between at least three distinct strata of student ability.

Item Fit Statistics

The analysis of the 25 items reveals that most items fall within the acceptable Infit and Outfit MNSQ range (0.5 to 1.5) following Wright and Linacre (1994) [37]. The mean Infit MNSQ of 1.00 and Outfit MNSQ of 0.99 suggest that items align well with the construct. ZSTD threshold of |Z| <= 2.0 (Bond and Fox, 2015) [36] and Point-Measure Correlation polarity (Linacre, 2002) [38] were applied. The observed match percentage of 69.7% supports the measurement quality of the model.

Unidimensionality

The Rasch model accounts for 43.1% of the raw variance, meeting the unidimensionality threshold (>=40%). The 1st contrast eigenvalue of 3.0 indicates no substantial secondary dimension. The scale exhibits unidimensionality, providing structural validity evidence within Messick's framework.

Table 4. Item Dimensionality for DTx

Table of STANDARDIZED RESIDUAL variance (in Eigenvalue units)

	Eigenvalue	Observed		Expected
Total raw variance in observations =	43.9028	100.0%		100.0%
Raw variance explained by measures =	18.9028	**43.1%**		42.9%
Raw variance explained by persons =	16.5078	37.6%		37.5%
Raw Variance explained by items =	2.3949	5.5%		5.4%
Raw unexplained variance (total) =	25.0000	56.9%	100.0%	57.1%
Unexplned variance in 1st contrast =	**3.0582**	7.0%	12.2%	
Unexplned variance in 2nd contrast =	2.5865	5.9%	10.3%	
Unexplned variance in 3rd contrast =	1.9597	4.5%	7.8%	
Unexplned variance in 4th contrast =	1.7744	4.0%	7.1%	
Unexplned variance in 5th contrast =	1.4661	3.3%	5.9%	

4.4 Group Comparisons

Research Question 3: How do students from Premiere and Conventional Polytechnics differ in their Rasch calibrated person measures of 21st-century skills and design thinking?

A comparison of 21st-century skills across polytechnic types revealed a statistically significant difference. Welch's t-test indicated that Conventional Polytechnic students scored, on average, 0.89 logits lower than their Premier counterparts, $t(174) = -2.85$, $p = .005$, Cohen's $d = 0.41$. This finding was corroborated by one-way ANOVA, $F(1, 205) = 8.16$, $p = .005$, and confirmed by chi-squared test, $x2(1) = 8.09$, $p = .004$. Collectively, these results indicate that institutional type has a significant impact on students' acquisition of 21st-century skills. A comparison of design thinking experience revealed that Premier Polytechnic students scored, on average, 1.03 logits higher than their Conventional counterparts, $t(181) = -3.14$, $p = .002$, Cohen's $d = 0.44$. One-way ANOVA, $F(1, 203) = 9.70$, $p = .0025$, and chi-squared test, $x2(1) = 9.86$, $p = .0017$, further confirmed that polytechnic type significantly predicts students' design thinking experience.

Research Question 4: How do students from Semester 1 and Semester 2 differ in their Rasch-calibrated person measures of 21st-century skills and design thinking?

A comparison of 21st-century skills by semester revealed that Semester 2 students scored on average 1.03 logits higher than Semester 1 students, $t(198) = -3.37$, $p = .001$, Cohen's $d = 0.47$. ANOVA confirmed $F(1, 205) = 11.24$, $p = .0013$, and chi-squared $x2(1) = 11.35$, $p = .0008$, providing clear evidence of significant skill development across the FYP progression. For design thinking experience, Semester 2 students scored 1.48 logits higher, $t(196) = -4.65$, $p < .001$, Cohen's $d = 0.65$. ANOVA $F(1, 203) = 21.46$, $p < .0001$ and chi-squared $x2(1) = 21.67$, $p < .0001$ confirmed substantial growth in DT experience over the two-semester FYP.

Research Question 5: How do students from different academic engineering programs differ in their Rasch-calibrated person measures of 21st-century skills and design thinking?

A comparison of 21st-century skills across engineering programs showed a statistically significant effect of program type, $F(2, 204) = 5.12$, $p = .0066$, $x2(2) = 9.84$, $p = .0071$. Students in Diploma in Electronics Engineering scored significantly lower than those in Mechanical Engineering (mean difference = -0.83 logits, $p = .013$, Cohen's $d = 0.39$) and Civil Engineering (mean difference = -1.15 logits, $p = .012$, Cohen's $d = 0.50$). No significant difference was observed between Mechanical and Civil Engineering students. For design thinking experience, the ANOVA indicated $F(2, 202) = 7.55$, $p = .0007$, $x2(2) = 14.55$, $p = .0007$. Electronics Engineering students scored significantly lower than Mechanical Engineering (mean difference = -1.06 logits, $p = .003$, Cohen's

d = 0.47) and Civil Engineering (mean difference = -1.47 logits, p = .002, Cohen's d = 0.61) students.

5 Discussions

This study sought to assess the psychometric quality of an instrument measuring 21st Century Skills among polytechnic students engaged in Design Thinking-driven final-year projects. Using the Rasch Measurement Model, the instrument demonstrated strong psychometric properties across multiple parameters, including reliability, item fit, unidimensionality, and rating scale functionality.

5.1 Implications for TVET and Design Thinking Pedagogy

The findings affirm the theoretical alignment between the Design Thinking framework and the targeted 21st Century Skills. Students engaged in DT projects are challenged to empathise, ideate, and co-create. These are the processes that inherently cultivate communication, collaboration, critical thinking, and creativity [29,32]. These results suggest that Design Thinking not only serves as a practical instructional model but also a measurable pathway to holistic skill development within polytechnic environments.

5.2 Psychometric Contributions

From a methodological perspective, this study contributes evidence of psychometric quality for a measurement tool addressing 21st Century Skills in higher education. By applying Rasch analysis within Messick's (1989) validity framework [35], the research provides structural validity evidence and content-related evidence, addressing long-standing concerns over the adequacy of soft skill assessments in polytechnic education [28]. It is acknowledged that this study provides evidence for two components of Messick's framework-structural and content validity-while recognising that convergent/discriminant, generalisability, and consequential validity evidence would be required for a comprehensive validity argument.

5.3 Limitations and Future Research

While the instrument showed high psychometric quality in this study, several limitations warrant consideration:

- The sample was drawn from only two polytechnic institutions; further validation is needed across more diverse institutional settings.
- The self-reported nature of the instrument may introduce bias; future studies should triangulate with performance-based assessments or peer evaluations.
- The correlation between actual DT project quality and measured skill levels was not examined.
- Future work should include convergent/discriminant and consequential validity studies to complete the validity argument under Messick's framework.

- Future work could also include longitudinal studies to track skill development over time or explore the instrument's utility in other TVET sectors such as community colleges or industrial training institutes.

6 Conclusion

This study assessed the psychometric quality of a newly developed instrument designed to measure 21st Century Skills-specifically communication, collaboration, critical thinking, creativity, and ethics-among final-year polytechnic students engaged in Design Thinking-based projects. Using the Rasch Measurement Model and situated within Messick's (1989) validity framework, the instrument demonstrated strong psychometric properties, including high item and person reliability, acceptable fit statistics, unidimensionality supported by disattenuated correlation analysis, and well-functioning rating scale categories. These findings provide structural and content-related validity evidence for the instrument's use in assessing essential soft skills within a Design Thinking pedagogical context in Malaysian TVET institutions.

As polytechnics aim to produce work-ready graduates aligned with the demands of Industry 4.0 and lifelong learning, the availability of a reliable and psychometrically sound measurement tool offers significant value for curriculum development, instructional improvement, and educational research. Future research should explore longitudinal applications, incorporate multi-method assessment approaches including convergent/discriminant and consequential validity investigations, and expand the instrument's validation across broader TVET and higher education settings.

Acknowledgments. This research was funded by the Universiti Teknologi MARA (UiTM) External Research Grant (PY/2022/01868) and supported by the UiTM–Universiti Teknologi Malaysia (UTM) Matching Grant (PY/2022/03144).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Malaysia Education Development Plan (MEDP) 2013-2025. https://www.moe.gov.my
2. Jamaluddin, N. S. A., et al. (2024). Seeking higher order construction of cognitive abilities in a psychosocial learning environment. Int J Eval & Res Educ, 13(3), 1979-1987.
2. O'Dea, X., et al. (2025). Graduate attributes in the age of disruptive technology. In Artificial Intelligence in Education (Vol. 1, No. 1, pp. 144-158). Emerald Publishing Limited.
3. Al-Qaimari, G., & Khan, M. A. (2025). Future Trends in Higher Education. In Strategic Planning for Higher Education Institutions (pp. 329-343). Springer Nature Switzerland.
4. Lam, M. H. A. (2024). Exploring the future-ready graduate attributes. International Journal of Chinese Education, 13(1), 2212585X241234338.

5. van Laar, E., et al. (2022). Developing policy aimed at 21st-century digital skills. Journal of Education and Work, 35(2), 195-209.
6. World Economic Forum (2020). The Future of Jobs Report 2020. https://www.weforum.org
7. Koya, C. F. (2015). Pedagogical practices in Fiji schools. Transforming teaching and learning in Asia and the Pacific, 22-43.
8. Voss, R., Lynch, J., & Herbert, S. (2023). Teacher concerns about competency-based mathematics education. Journal of Vocational Education & Training, 75(5), 867-889.
9. Torrisi-Steele, G. (2020). Facilitating the shift from teacher centred to student centred university teaching. IJAET, 11(3), 22-35.
10. Razzouk, R., & Shute, V. (2012). What is design thinking and why is it important? Review of Educational Research, 82(3), 330-348.
11. Henriksen, D., Richardson, C., & Mehta, R. (2017). Design thinking: A creative approach. Thinking Skills and Creativity, 26, 140-153.
12. Gronman, S., & Lindfors, E. (2021). The process models of design thinking. Techne serien, 28(2), 110-118.
13. Oxman, R. (2017). Thinking difference: Theories and models of parametric design thinking. Design Studies, 52, 4-39.
14. Geisinger, K. F. (2016). 21st century skills: What are they and how do we assess them? Applied Measurement in Education, 29(4), 245-249.
15. Husain, M. Y., et al. (2010). Importance of employability skills from employers' perspective. Procedia-Social and Behavioral Sciences, 7, 430-438.
16. Abdullah, W. F. W., et al. (2022). Employability skills in the TVET trainer training program. Journal of Technical Education and Training, 14(1), 150-157.
17. Suto, I., & Eccles, H. (2014). The Cambridge approach to 21st Century skills. IAEA conference, Singapore, 1-10.
18. Trilling, B., & Fadel, C. (2009). 21st century skills: Learning for life in our times. John Wiley & Sons.
19. Zainuddin, Z. A. A., & Selamat, S. (2012). Efficacy of Polytechnic Students' Interpersonal Communication Skills. Advances in Language and Literary Studies, 3(2), 76-86.
20. Esa, A., et al. (2014). Applications of soft skills in engineering programme at polytechnic Malaysia. Procedia-Social and Behavioral Sciences, 140, 115-120.
21. Deep, S., et al. (2019). Study on problem-based learning towards improving soft skills. International Journal of Innovation and Learning, 25(1), 17-34.
22. Choi, H., & Hur, J. (2023). Passive participation in collaborative online learning activities. Online Learning Journal, 127.
23. Agapova, T. V., & Aisner, L. Y. (2019). Basic forms of interaction and teaching methods in higher school. Pedagogicheskij zhurnal, 9(1-1), 269-275.
24. Somaratne, P. (2021). Identify reasons for passive behaviour of Asian students during collaborative work. IRSFOE 2021, 348.
25. Tushar, H., & Sooraksa, N. (2023). Global employability skills in the 21st century workplace. Heliyon, 9(11).
26. Ahmad, A. M., et al. (2017). Level of Ethical Behavior among Engineering Students: Polytechnic Malaysia. JoHDeC, 6, 1-14.
27. Ismail, A., et al. (2019). Effectiveness of entrepreneurship programmes in developing entrepreneurship skills. Journal of Technical Education and Training, 11(1).

28. Razzouk, R., & Shute, V. (2012). What is design thinking and why is it important? Review of Educational Research, 82(3), 330-348.
29. Waidelich, L., et al. (2018). Design thinking process model review. IEEE International Conference on Engineering, Technology and Innovation, 1-9.
30. Dam, R. F., & Siang, T. Y. (2020). Design thinking: Get started with prototyping. Interaction Design Foundation.
31. Searle, R. (2023). Exploring the application of Design Thinking to develop a work-based learning programme for hospitality and tourism.
32. Grau, S. L., & Rockett, T. (2022). Creating student-centred experiences: Using design thinking. The Journal of Entrepreneurship, 31(2_suppl), S135-S159.
33. Fan, J., & Bond, T. (2019). Applying Rasch measurement in language assessment. In Quantitative Data Analysis for Language Assessment Volume I (pp. 83-102). Routledge.
34. Messick, S. (1989). Validity. In R. L. Linn (Ed.), Educational Measurement (3rd ed., pp. 13-103). American Council on Education/Macmillan.
35. Bond, T. G., & Fox, C. M. (2015). Applying the Rasch Model: Fundamental Measurement in the Human Sciences (3rd ed.). Routledge.
36. Wright, B. D., & Linacre, J. M. (1994). Reasonable mean-square fit values. Rasch Measurement Transactions, 8(3), 370.
37. Linacre, J. M. (2002). What do infit and outfit, mean-square and standardized mean? Rasch Measurement Transactions, 16(2), 878.
38. Tennant, A., & Conaghan, P. G. (2007). The Rasch measurement model in rheumatology. Arthritis Care & Research, 57(8), 1358-1362.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

