JTET

JOURNAL OF TECHNICAL EDUCATION AND TRAINING
ISSN: 2229-8932 e-ISSN: 2600-7932

Vol. 18 No. 1 (2026) 212-229
https://publisher.uthm.edu.my/ojs/index.php/ijie

Development and Validation of Visualisation Ability Test for Technical Drawing Students Using Rasch Models

Mohd Salehudin Marji[1]*, Aede Hatib Musta'amal[1], Nor Aisyah Che Derasid[1] Leong Ting Chee[2]

[1] *Department of Advanced Technical and Vocational Education and Training, Universiti Teknologi Malaysia, Jalan Iman, 81310 Skudai, Johor Darul Ta'zim, 81310, MALAYSIA*

[2] *Riam Institute of Technology, Department of Mechanical Engineering, Lot 1305, Mile 3.5, Riam Avenue, Road, Miri, Sarawak, 98000, MALAYSIA*

*Corresponding Author: salehudinmarji@gmail.com
DOI: https://doi.org/10.30880/jtet.2026.18.01.015

Article Info

Received: 5 January 2026
Accepted: 30 March 2026
Available online: 23 June 2026

Keywords

Visualization ability test, technical drawing, Rasch model

Abstract

This study aimed to develop a visualization ability test for technical drawing students and to provide preliminary evidence of its validity and reliability using the Rasch model. The main problem is the lack of a specific measuring tool to assess students' visualization abilities in the context of technical drawing in Malaysia. The study employed a development approach, involving four experts in technical drawing, engineering, and architecture, as well as 34 Form 4 *Grafik Komunikasi Teknikal* students from two secondary schools in Johor as the sample. Expert validation showed content validity, with an overall Fleiss' Kappa of 0.85, while Rasch analysis indicated good item reliability (0.84) and a Cronbach's Alpha of 0.77. The analysis showed that the instrument consistently and significantly distinguished different levels of students' visualization ability. This study provides preliminary evidence that the developed test has acceptable validity and reliability for assessing visualization ability in technical drawing learning. Recommendations are offered for further research with larger samples and for integrating technology to improve the application and generalizability of results.

1. Introduction

Students in the field of technical drawing still struggle to interpret, visualize, and manipulate three-dimensional objects accurately, even after instruction based on the basic concepts of engineering drawing (Abbasian, 2024; Ilanković et al., 2025; Yuan et al., 2024). This problem causes students to often make mistakes in reading views, producing orthographic drawings, understanding the transformation of shapes when objects are rotated or reoriented, and affecting their performance in courses related to technical graphics (Guo et al., 2025; Setiyawan et al., 2025). This situation highlights the need to develop a test that can measure students' visualization ability accurately and reliably.

Visualization ability is a core element of spatial skills, involving an individual's ability to process, organize, and manipulate visual representations in the mind (Tandon et al., 2023; Tiwari et al., 2024). Visualization ability has long been recognised as a key component of spatial cognition. Early studies by Shepard and Metzler (1971) demonstrated that individuals mentally rotate objects to solve spatial problems, forming the basis of mental rotation theory. Later, Lohman (2013) conceptualised spatial ability as a multidimensional construct consisting of spatial visualization, mental rotation, and spatial relations. In engineering and technical education contexts, Sorby (2009) further emphasized that spatial visualization ability plays a crucial role in understanding graphical representations and interpreting engineering drawings. These theoretical perspectives suggest that visualization ability is not a single skill but a set of related cognitive processes that support the interpretation and manipulation of spatial information. Grounded in spatial cognition and visualization theory, visualization ability in this study is treated as a theoretically defined construct rather than a purely operational skill. In this study, visualization ability is conceptualized as a multidimensional construct encompassing several spatial sub-skills, including mental rotation, spatial visualization, two-dimensional to three-dimensional transformation, and spatial relations. These dimensions represent the cognitive processes required when interpreting and constructing technical drawings. In this study, visualization ability is conceptualized as a multidimensional construct encompassing several spatial sub-skills, including mental rotation, spatial visualization, two-dimensional to three-dimensional transformation, and spatial relations. These dimensions reflect the cognitive processes required when interpreting and constructing technical drawings. In the context of technical drawing, this skill is essential for understanding the relationships among views in orthographic systems, determining the positions of objects in three-dimensional space, and interpreting the structures of mechanical and engineering components (Mackiewicz, 2025). Students with poor visualization abilities typically show confusion in connecting plan, front, and side views, difficulty reading complex drawings, and errors in matching three-dimensional shapes to 2D representations (Bartlett & Dorribo Camba, 2023; Korkut & Surer, 2023; Sudirman et al., 2023). The implications for learning are significant, as these weaknesses also affect the overall design skills and production of technical drawings.

Although several instruments measure visualization ability, such as the Mental Rotation Test and the Purdue Spatial Visualization Test, studies show that these instruments are not fully suited to the context of learning technical drawing, particularly in the local education system (Bartlett, 2024). Many of these instruments were developed for general spatial ability assessment, are not explicitly grounded in technical drawing curricula, and do not sufficiently reflect the specific cognitive demands of graphical communication tasks. Furthermore, some instruments lack rigorous psychometric validation within applied technical education settings (Bartlett & Camba, 2023). This has led to an urgent need for instruments specifically designed for technical drawing students. In the context of technical drawing education in Malaysia, the subject of technical drawing, or better known as *Grafik Komunikasi Teknikal* (GKT), was only introduced in 2017. There is still a lack of studies on this subject, especially regarding teaching aids, teaching modules, and the measurement of students' visualization ability (Kiong et al., 2023).

Based on these requirements, this study focuses on the development of visualization ability tests tailored to the context of technical drawing education, covering the process of construct determination, item construction, expert evaluation, preliminary testing, and empirical analysis. The conceptual framework underpinning the instrument links curriculum demands in technical drawing with core visualization sub-skills and their observable manifestations in students' responses. The construct framework guiding item development integrates key dimensions of visualization ability relevant to technical drawing, namely mental image manipulation, spatial rotation, three-dimensional representation, and geometric interpretation within orthographic systems (Gummaluri et al., 2025; Piri & Cagiltay, 2024). Each dimension is operationalized through test items designed to elicit specific visualization behaviors required in authentic technical drawing tasks. In addition, expert ratings are used to assess the suitability of the content, the clarity of the instructions, and the accuracy of the graphical representation.

To ensure the psychometric quality of the instrument, this study adopts the Rasch model as the primary analytical framework. The Rasch model enables detailed examination of item functioning, item difficulty, student ability, and scale structure. Compared to classical test theory, which relies heavily on sample-dependent statistics such as total scores and item difficulty indices, Rasch modelling offers sample-independent item calibration and provides interval-level measurement. This characteristic makes Rasch model particularly suitable for instrument development, as it allows researchers to evaluate whether each item consistently measures the intended visualization construct and to identify misfitting or poorly functioning items (Van Zile-Tamsen, 2017). The use of the Rasch model also allows researchers to construct an objective hierarchy of item difficulty levels, thereby ensuring that tests can distinguish between low-, moderate-, and high-ability students (He et al., 2023).

In addition, Rasch's analysis produces indicators such as item polarity, fit statistics, item-person maps, and reliability, which help refine the instrument to ensure validity and reliability are at a satisfactory level (Combrinck, 2020). In the context of technical education, previous research has also recommended the use of the Rasch model because it provides a more accurate picture of the relationship between students' abilities and teaching needs (Chakraburty et al., 2025; Suryadi et al., 2025). Through valid and empirically robust measurement, educators can design more targeted interventions, including the use of three-dimensional models, digital simulations, and technology-enhanced visual learning strategies.

In this regard, this study aims to develop a visualization ability test for technical drawing (GKT) students and to provide preliminary calibration evidence of its measurement properties using the Rasch model, with the aspiration to provide a standard instrument for academic assessment, research, and pedagogical development, especially in Malaysia. This instrument is expected to help teachers and lecturers identify students' weaknesses in visualization ability and support the improvement of spatial literacy and technical drawing skills. With the availability of valid, uniform instruments, improvement efforts in the teaching of technical drawings can be carried out more systematically and based on solid empirical data.

2. Method

The study employed an instrument development and validation design, which was suitable for developing new tests or measurement tools through a systematic process, including item construction, expert content validation, pilot testing, and psychometric analysis. The validation evidence reported in this study was intended to be preliminary, as the instrument was evaluated at the pilot-testing stage rather than through large-scale confirmatory validation. The development and validation of the instrument were conducted through three sequential phases. The overall research procedure is illustrated in Figure 1.

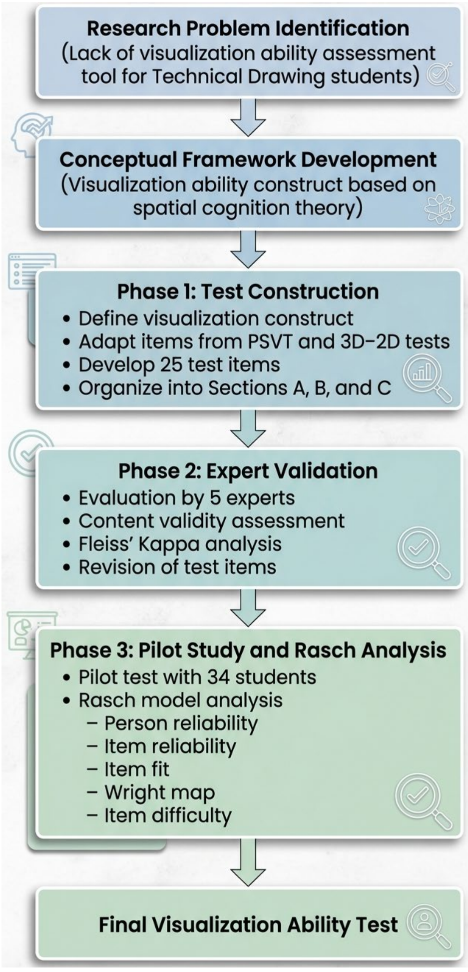

Fig. 1 *Development and validation procedure of the visualization ability test*

The overall research procedure consisting of Phase 1 (test construction), Phase 2 (expert validation), and Phase 3 (pilot testing and Rasch analysis) is illustrated in Figure 1. Phase 1, test construction, involved defining the construct of visualization ability, preparing the test items, and conducting an initial review by the researcher. Phase 2, content validation, involved evaluation by five experts in technical drawing, engineering, and architecture, with their feedback used to refine the items. Phase 3 comprised a pilot study followed by Rasch model analysis, where the test was administered to 34 students, and the data were analyzed using the Rasch model to assess the instrument's validity and reliability. This study employed a quantitative approach involving the collection of numerical data and statistical analysis using the Rasch model.

The study sample was selected according to the research phases. In Phase 2, five experts were chosen to evaluate the suitability of the items in the developed test. The experts were selected based on several criteria to ensure their relevance and expertise in the field. The selection criteria included (1) having professional expertise in technical drawing, engineering, or architecture; (2) possessing at least a bachelor degree or equivalent professional qualification; (3) having a minimum of five years of teaching or professional experience related to technical drawing or graphical communication; and (4) having experience in curriculum development, technical drawing instruction, or assessment. The aspects assessed by the experts included the test format, item definitions, alignment of items with the target sample, font size, instructions, illustration sizes in the questions, and the allotted time to complete the test. The five experts were selected based on pre-determined criteria set by the researcher, with two experts were lecturers from public universities in Malaysia, while the remaining three were experienced technical drawing (GKT) teachers and the remaining three being technical drawing teachers (GKT) with extensive experience in the field. A list of experts and their respective areas of expertise is presented in Table 1.

Table 1 *List of experts for the evaluation of the visualization ability test*

Expert	Areas of Expertise	Position	Teaching experience
1	Engineering	Public university lecturer (PhD)	15 years
2	Architecture	Public university lecturer (PhD)	8 years
3	Technical drawing	Teacher	27 years
4	Technical drawing	Teacher	23 years
5	Technical drawing	Teacher	15 years

Meanwhile, in Phase 3, to evaluate the validity and reliability of the visualization ability test, 34 Form 4 students enrolled in the Technical Drawing (GKT) subject from two secondary schools in Johor were selected as the study sample using purposive sampling. According to Aziz et al. (2017), a sample size between 16 and 36 respondents is considered adequate for preliminary Rasch analysis in pilot studies. This sample size is also consistent with recommendations for pilot studies in instrument development using the Rasch model, where smaller samples are acceptable for preliminary calibration and item functioning analysis (Boone, 2016). The primary objective at this stage was to examine the initial measurement properties of the instrument rather than to achieve population generalization. The Rasch model was used to evaluate item functioning and scale structure, with item retention guided by empirical evidence and expert judgement. Ethical clearance and official permission to conduct this study were obtained from the Educational Planning and Research Division, Ministry of Education Malaysia (Reference No.: KPM.600-3/2/3-eras(16028), dated 7 April 2023) and the Johore State Education Department (Reference No.: JPNJ.100-12/2/1 Jld 3 (59), dated 13 April 2023). Both approvals were granted subject to the permission and consideration of the participating school administrators. Participation in the study was entirely voluntary, and informed consent was obtained from all participants before data collection commenced. Students were informed of the purpose of the study and were assured that their responses would be kept confidential and used solely for research purposes. All data were anonymized and managed in accordance with relevant research ethics and institutional guidelines.

3. Results and Discussion

This section primarily reports the empirical findings from the Rasch analysis, including reliability indices, item fit statistics, Wright map distribution, and item difficulty measures. Interpretations of these findings and their implications for visualization ability assessment are discussed in relation to relevant literature within the discussion components of this section.

3.1 Phase 1

Researchers developed the visualization ability test to measure students' level of visualization ability. The test was developed by adapting existing tests, namely the Purdue Spatial Visualization Test for Development and the 3D-to-2D Transformation Test. Researchers commonly use both tests to assess students' visualization skills, as seen in studies by Mohd Safarin (2009), Rohani (2018), and Thompson (2012). The test consists of 25 items and is divided into three sections. Table 2 provides a detailed breakdown of the visualization ability test, while Tables 3, 4, and 5 indicate the difficulty level of each test item in Sections A, B, and C, respectively.

Table 2 *Visualization ability test details*

Section	Objective	Number of questions	Sorting questions according to difficulty level	
Section A 3D to 2D Development Test	To assess students' ability to visualize 2D development drawings from the given 3D objects	10	1 &2	Very easy
			3 & 4	Easy
			5 & 6	Moderate
			7 & 8	Difficult
			9 &10	Very Difficult
Section B 2D to 3D Development Test	To assess students' ability to visualize a 3D object from the given 2D development drawing	10	1 & 2	Very Easy
			3 & 4	Easy
			5 & 6	Moderate
			7 & 8	Difficult
			9 & 10	Very Difficult
Section C 3D to 2D Orthographic Test	To assess students' ability to visualize the appearance of a 3D object in 2D form when viewed from the given orthographic projection	5	1	Very easy
			2	Easy
			3	Moderate
			4	Difficult
			5	Very difficult

Table 3 *Classification of questions by degree of difficulty. Section A*

Degrees Difficulty	Question number	Characteristics of questions
1	1&2	• Simple 3D objects. • 2D development drawing not to exceed 5 surfaces. • The base of the 3D object is in the same position for all 5 answer options.
2	3&4	• A given 3D object has 7-8 surfaces. • One base of the 3D object among the 5 answer options is rotated differently from the others.
3	5&6	• The given 3D object is a combination of two 3D objects. • 2 bases of the 3D object among the 5 answer options are rotated differently from the others.
4	7&8	• The given 3D object is a combination of two to three 3D objects. • 3 bases of the 3D object among the 5 answer options are rotated differently from the others.
5	9&10	• The given 3D object is a combination of three to four 3D objects. • All 5 answer options have bases that rotate at different rates.

Table 4 *Classification of questions by degree of difficulty. Section B*

Degrees Difficulty	Question number	Characteristics of questions
1	1&2	• Simple 3D objects. • 2D development drawing not to exceed 5 surfaces. • All 5 3D objects in the answer options are drawn on the same isometric view.
2	3&4	• The given 3D object has 7-8 surfaces. • One of the 5 3D objects in the answer options is drawn in an isometric view that differs from the others.
3	5&6	• The given 3D object is a combination of two 3D objects. • The given 3D object has 8-9 surfaces. • Two of the 5 3D objects in the answer options are drawn in an isometric view that differs from the others.
4	7&8	• The given 3D object is a combination of two to three 3D objects. • The 2D development drawing given has 9-10 total surfaces. • The position of the base in the 2D development drawing is rotated differently from the base positions of the 3D objects in all five answer choices.
5	9&10	• The given 3D object is a combination of three to four 3D objects. • The 2D development drawing given has 12-13 total surfaces. • The position of the base in the 2D development drawing is rotated differently from the base positions of the 3D objects in all five answer choices.

Table 5 *Classification of questions by degree of difficulty. Section C*

Degrees Difficulty	Question number	Characteristics of questions
1	1&2	• Simple 3D object. • There are only 2 to 3 surfaces on the answer options. • 3 of the 5 answer options are orthographic views that differ from the view according to the given arrow.
2	3&4	• Simple 3D object • There are 3 to 5 surfaces in the answer option • 3 of the 5 answer options are orthographic views that differ from the views according to the given arrow. • Introduction to hidden lines.
3	5&6	• 3D objects with curves and circles are introduced • Use of hidden lines to show hollow/ • 2 of the 5 answer options are orthographic views that differ from the view according to the given arrow.
4	7&8	• The use of more complex 3D objects as a result of a combination of straight lines, curves, holes, and circles/ • 1 of the 5 answer options is an orthographic view that differs from the view according to the given arrows. • The use of complex hidden lines.
5	9&10	• The use of more complex 3D objects as a result of a combination of straight lines, curves, holes, and circles/ • All 5 answer options show the same orthographic view, as indicated by the given arrow. • The use of complex hidden lines.

J. of Technical Education and Training Vol. 18 No. 1 (2026) p. 212-229

Meanwhile, Figures 2, 3, and 4 show examples of questions in sections A, B, and C.

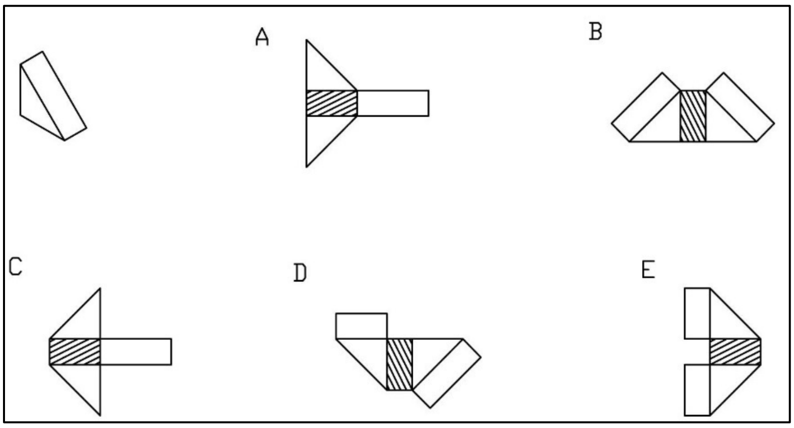

Fig. 2 *Sample section A questions*

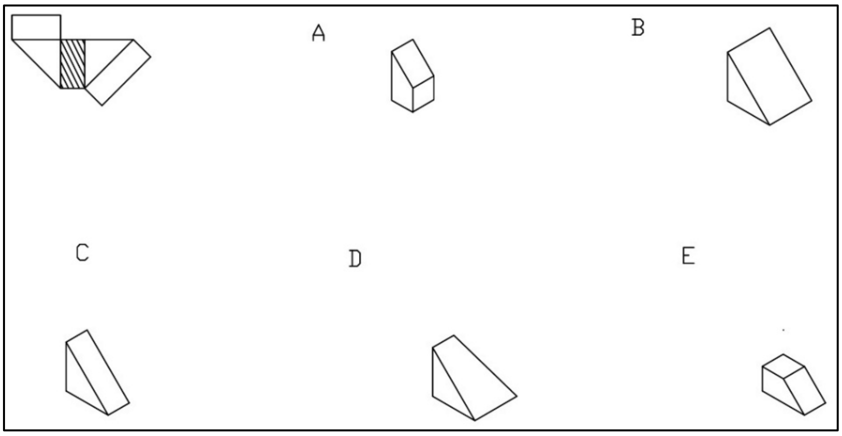

Fig. 3 *Sample section B questions*

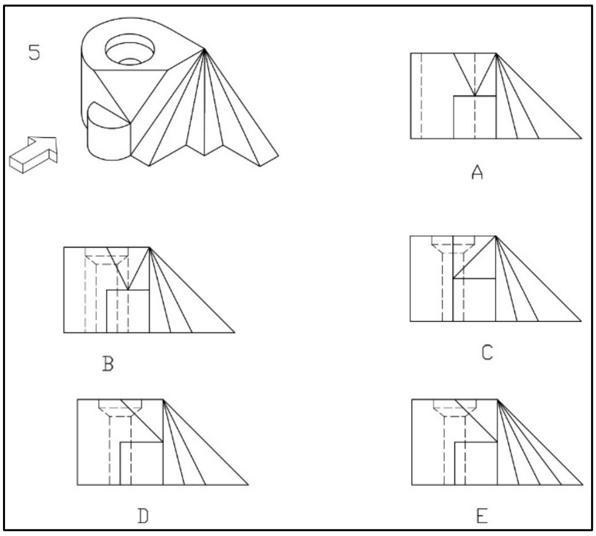

Fig. 4 *Sample section C questions*

3.2 Phase 2

The findings from Phase 2 are presented in Table 6. The evaluations from the five experts were analyzed using Fleiss' kappa analysis.

Table 6 *Fleiss' Kappa analysis for expert evaluation of the visualization ability test*

No	Criteria assessed	P1	P2	P3	P4	P5	Kappa Values (k)
1	The test format is appropriate.	/	/	/	/	/	1.00
2	Each question item is clearly defined.	/	/	/	/	/	1.00
3	The tests developed are suitable for the study sample.	/	/	/	/	/	1.00
4	The item is appropriately aligned with the research instrument's construct.	/	/	/	/	/	1.00
5	Font size is suitable and easy to read.	/	/	/	/	/	1.00
6	Clear and easy-to-understand instructions.	/	/	/	/	/	1.00
7	The illustration size is clear and appropriate.	X	/	X	/	/	0.40
8	The time given to answer this test is sufficient.	X	X	/	/	/	0.40
	Overall Kappa (k) value						0.850

Note: P1–P5 represent the five expert evaluators; "/" indicates agreement and "X" indicates disagreement.

Table 6 presents the Fleiss' Kappa analysis for the expert evaluation of the items in the visualization ability test. The findings show that 6 of the 8 evaluated criteria achieved perfect agreement among all experts, with a kappa value of 1.00. Meanwhile, two criteria, Criterion 7 and Criterion 8, showed weak kappa values of 0.40. For Criterion 7, Expert 1 suggested that the researcher use a larger paper size (A3) to print the test questions. However, due to high costs and logistical considerations, the researcher decided to continue using A4-sized paper but to increase the size of the illustration instead. Regarding the test duration in Criterion 8, the researcher revised the original time allocation from 25 minutes to 30 minutes for students to answer all 25 questions. Overall, the evaluation yielded a kappa of 0.850, indicating a very good level of agreement among all experts.

3.3 Phase 3

Following expert validation and subsequent revisions based on their feedback, a pilot study was conducted to assess the validity and reliability of the developed visualisation ability test. The study involved 34 Form 4 technical drawing (GKT) students from two secondary schools in Johor, who were selected to complete the test. Rasch model was employed to examine the data collected from these students, providing a robust assessment of the test's psychometric properties. The results of the Rasch model are presented as follows. The following subsections present the empirical measurement results of the instrument, including reliability, item fit, Wright map distribution, and item difficulty.

3.3.1 Person Reliability Analysis of the Entire Test

Table 7 *Person reliability analysis of the visualization ability test*

Measured Person	34
Alpha Cronbach value	0.77
Person reliability value	0.62

Table 7 presents the person reliability analysis of the visualization ability test. The Rasch analysis of responses from 34 students yielded a person reliability value of 0.62, indicating moderate response consistency, while Cronbach's Alpha was 0.77, reflecting good internal consistency (Sumintono & Widhiarso, 2014). This level of person reliability suggests that the instrument can differentiate students into general ability strata. However, it may not yet optimally separate individuals with closely adjacent ability levels, which may be due to a partial mismatch between the distribution of item difficulty and the ability range of the pilot sample, as well as variation in students' prior exposure to technical drawing. Future studies may improve person separation and measurement precision by increasing the number of moderate-difficulty items and expanding the sample size.

These reliability findings provide an initial indication of the instrument's measurement consistency, which will be further interpreted in relation to visualization ability assessment in the discussion. From a Rasch measurement perspective, person reliability reflects the extent to which the instrument can consistently distinguish respondents according to their ability levels. A value of 0.62 indicates a moderate level of separation among students, suggesting that the instrument is capable of differentiating broad categories of visualization ability, such as low, moderate, and high performers. However, this value also implies that the precision of measurement may still be limited when distinguishing students with closely similar ability levels. Such a pattern is common in pilot studies with relatively small samples and a limited range of respondent abilities (Boone, 2016; Sumintono & Widhiarso, 2014). Increasing the sample size and introducing additional items with varying levels of difficulty could enhance person separation and improve the stability of the reliability estimate in future studies.

3.3.2 Item Reliability Analysis of the Entire Test

Table 8 *Item reliability analysis for the visualization ability test*

Measured Item	25 questions
Alpha Cronbach value	0.77
Person reliability value	0.62

Table 8 presents the item reliability analysis of the visualization ability test. The Rasch model of all 25 test items yielded an item reliability of 0.84, indicating good item quality. Additionally, Cronbach's Alpha for this analysis was 0.77, indicating good internal consistency (Sumintono & Widhiarso, 2014). The high item reliability and Cronbach's Alpha values demonstrate that the developed test is of good quality and reliable.

3.3.3 Item Fit Order Analysis for the Entire Test

Table 9 presents the item fit analysis for the entire visualization ability test. As explained by Aziz et al. (2017), if an item exhibits MNSQ and PT Measure CORR values that do not meet the required criteria. However, the ZSTD value falls within the acceptable range, the item is still considered to fit and can therefore be retained (Sumintono & Widhiarso, 2014). Accordingly, the researcher examined the Table to identify any instances of misfit based on the MNSQ, ZSTD, and PT Measure CORR values for each test item.

Table 9 *Item fit order analysis for the visualization ability test*

Item	Outfit		Infit		PT Measure CORR
	MNSQ	ZSTD	MNSQ	ZSTD	
A1	0.99	0.10	0.97	0.14	0.30
A2	0.84	0.11	0.20	-0.45	0.42
A3	1.07	0.34	1.24	0.65	0.23
A4	0.96	-0.05	0.99	0.13	0.30
A5	0.94	-0.49	1.16	0.90	0.37
A6	1.07	0.65	1.04	0.27	0.29
A7	0.95	-0.32	0.93	-0.27	0.39
A8	1.29	2.39	1.48	2.49	0.02
A9	1.10	0.67	1.09	0.43	0.22
A10	0.84	-1.42	0.81	-1.11	0.53
B3	1.12	0.39	1.09	0.41	0.13
B4	0.89	-0.07	0.68	-0.25	0.39
B5	0.95	0.05	1.24	0.54	0.26
B6	0.81	-1.66	0.77	-1.34	0.57
B7	0.93	-0.50	0.87	-0.68	0.45
B8	0.78	-0.71	0.69	-0.67	0.56
B9	1.04	0.70	1.04	0.30	0.28
B10	0.98	-0.13	1.08	0.48	0.36
C1	0.77	-0.63	0.65	-0.63	0.55
C2	1.28	0.87	1.41	0.89	0.02
C3	0.87	-0.34	0.83	-0.28	0.45
C4	1.09	0.52	1.04	0.24	0.28
C5	1.10	0.82	1.08	0.48	0.26

Note: MNSQ = mean square fit statistic; ZSTD = standardized fit statistic; PT Measure Corr = point–measure correlation.

Items B1 and B2 are not listed in the Table because both received perfect scores, with all 34 students answering these questions correctly. These two items fall into the extreme category, as indicated in the Winsteps output. Consequently, they do not have infit or outfit values for measurement. Nevertheless, these items were not removed because the researcher had intentionally designed the test to include two very easy questions in each section, as shown in Table 2, where Items B1 and B2 are classified as very easy. Therefore, it is reasonable that all students answered these items correctly. Furthermore, before conducting the Rasch model using the pilot test data, the visualization ability test had already been validated by five experts, and no comments were raised regarding these two items.

Meanwhile, Item A2 yielded an MNSQ value outside the acceptable range; however, its PT Measure CORR and ZSTD values met the required criteria, indicating that Item A2 remains a suitable (fit) item. In addition, several items did not meet the PT Measure CORR threshold, namely Items A1, A3, A4, A5, A6, A7, A9, B3, B4, B5, B9, B10, C2, C4, and C5. Despite this, these items are still considered fit because their MNSQ and ZSTD values fall within the acceptable ranges of 0.5 to 1.5 for MNSQ and −2.0 to +2.0 for ZSTD (Aziz et al., 2017). Only one item, Item A8, exhibited dual misfit indicators, namely unacceptable ZSTD and PT Measure CORR values. Item A8 recorded an outfit MNSQ of 2.39 and an infit MNSQ of 2.49, both exceeding the maximum permissible value of +2.00. Its PT Measure CORR value was 0.02, which is also below the acceptable lower limit. Based on these results, Item A8 could reasonably be considered for removal. However, the item was provisionally retained because it represents a critical visualization task involving complex spatial rotation and the interpretation of multiple surfaces in a 3D object. Such cognitive processes are central to visualization ability in technical drawing, particularly in tasks that require students to mentally manipulate object orientation and translate between spatial representations. The item was also judged as conceptually relevant by all five experts during the content validation stage. Therefore, Item A8 was retained to preserve construct coverage, although it will be revised and re-evaluated in future studies with larger samples. Overall, the item fit analysis provides empirical evidence regarding the functioning of the developed items within the Rasch measurement framework.

3.4 Wright Map

Figure 5 presents the Wright map generated from the conducted Rasch model.

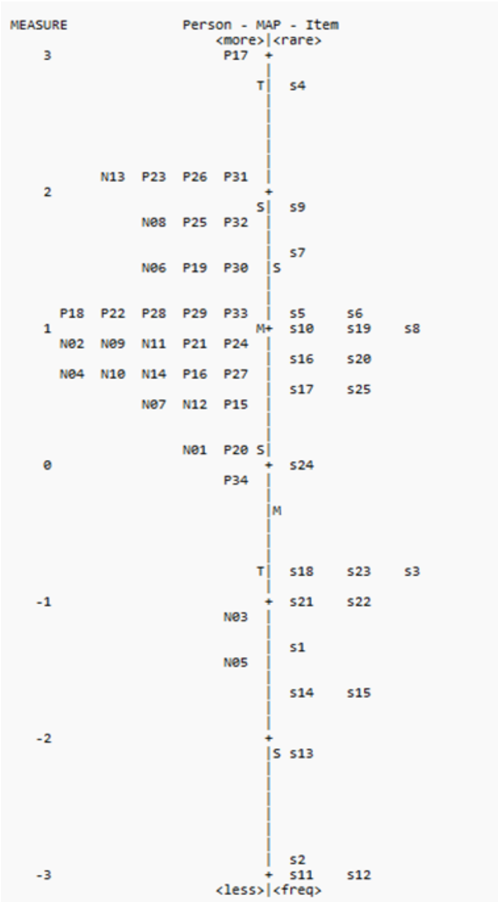

Fig. 5 *Wright map*

The right side of the Wright Map represents the distribution of test items by difficulty level, while the left side shows the distribution of student abilities. Based on the Wright Map, Item 4 (A4) is identified as the most difficult item compared to all other items. In contrast, Items 11 (B1) and 12 (B2) are the easiest among all test items. In terms of student ability, Student P17 demonstrates the highest level among the 34 students, whereas Student N05 shows the lowest. The Wright Map also indicates that none of the items fall outside the two standard deviation limits (T), which suggests that no test item is excessively easy or excessively difficult. The Wright map thus illustrates the relative distribution between item difficulty and student ability within the pilot sample.

3.5 Item Measure

The item measure is divided into 3 sections, corresponding to the three sections of the visualization ability test: A, B, and C. The difficulty level of each question was determined from the logit values, as shown in Table 10.

Table 10 *Item measure. Section A*

Question number	Number of correct answers	Logit Value	Logit Value mean	Degrees Difficulty
1	30/34	-1.30	-2.11	Easy
2	33/34	-2.91		
3	28/34	-0.77	0.98	Difficult
4	6/34	2.73		
5	16/34	1.13	1.13	Difficult
6	16/34	1.13		
7	13/34	1.55	1.28	Difficult
8	17/34	1.00		
9	11/34	1.84	1.42	Difficult
10	17/34	1.00		
Logit value Mean of the entire item			0.54	Difficult

Table 10 presents the item difficulty measures for the questions in Section A. During the test development process, the researcher initially divided the 10 questions into five pairs based on difficulty, ranging from very easy to very difficult, as shown in Table 2. The criteria determining difficulty, such as object shape, number of surfaces, rotation of the base in the orthographic drawing, and others, were agreed upon by the participating experts. Questions 1 and 2 had the lowest mean logit value of –2.11, indicating that these items were at an easy difficulty level. Question 4 was the most difficult item among the remaining nine questions, with a logit value of 2.73. However, after averaging the logit values of Question 4 and its paired question, Question 3, the resulting mean logit was 0.98, indicating that both questions fall within the difficult level. For Questions 5 and 6, the mean logit value was 1.13, also indicating a difficult level. Similarly, Questions 7 and 8 had a mean logit of 1.28, while the final pair, Questions 9 and 10, recorded the highest mean logit value of 1.42, classifying both items as difficult. Although all paired questions were classified as difficult level except for Questions 1 and 2, there is a clear progression in difficulty from Questions 3 and 4 through Questions 9 and 10, as indicated by the logit values. Overall, the items in Section A are dominated by difficult-level questions, with an average mean logit value of 0.54.

Table 11 *Item measure. Section B*

Question number	Number of correct answers	Logit Value	Logit Value mean	Degrees Difficulty
1	34/34	-4.15	-4.15	Easy
2	34/34	-4.15		
3	32/34	-2.14	-1.90	Easy
4	31/34	-1.66		
5	31/34	-1.66	-0.46	Moderate
6	19/34	0.73		
7	20/34	0.59	-0.09	Moderate
8	28/34	-0.77		
9	17/34	1.00	0.87	Difficult
10	19/34	0.73		
Logit value Mean of the entire item			1.15	Moderate

Table 11 presents the item difficulty measures for the questions in Section B. As in Section A, the researcher divided the 10 test questions into five pairs, ordered from easiest to most difficult, as shown in Table 2. Questions 1 and 2 were the easiest pair, with all 34 students answering both questions correctly. This pair is classified as easy, with a mean logit value of –4.15. Questions 3 and 4 are also categorized as easy, with a mean logit of -1.90, as only 2 students answered Question 3 incorrectly and 3 answered Question 4 incorrectly. Questions 5 and 6 fall into the moderate difficulty category, with a mean logit of –0.46, while Questions 7 and 8 are in the same category, with a mean logit of –0.09. Question 9 is the most difficult level in Section B, with a logit value of 1.00. After averaging with Question 10, the pair is classified as difficult, with a mean logit of 0.87. Similar to Section A, Section B also shows a progression in mean logit values from Questions 1 and 2 to Questions 9 and 10, indicating an increase in item difficulty according to the question sequence. On average, the difficulty level of all items in Section B is classified as moderate, with a mean logit of –1.15 (Sumintono & Widhiarso, 2014).

Table 12 *Item measure. Section C*

Question Number	Number of correct answers	Logit Value	Degrees Difficulty
1	29/34	-1.02	Moderate
2	29/34	-1.02	Moderate
3	28/34	-0.77	Moderate
4	24/34	-0.10	Moderate
5	20/34	0.59	Difficult
Logit value Mean of the entire item		-0.46	Moderate

Table 12 presents the item difficulty measures for the questions in Section C. As in Sections A and B, the questions in Section C were organized into five difficulty levels. However, Section C contains only five questions, which the researcher arranged by difficulty level according to the criteria outlined in Table 2. Questions 1 and 2 share the same logit value of -1.02, classifying both items as moderate in difficulty. Question 3, with a logit value of –0.77, is also categorized as moderate, as is Question 4, which has a logit value of –0.10. Question 5 has the highest logit value of 0.59, making it the most difficult itemin Section C and placing it in the difficultlevel category. Overall, Section C is dominated by moderate-difficulty questions, with an average mean logit value of –0.46 (Sumintono & Widhiarso, 2014).

3.6 Person Measure

The person-measure analysis was conducted to evaluate students' ability to answer the test items. This analysis also provides a detailed description of students' ability levels, as illustrated in the Wright Map. Table 13 presents the person measures of the 34 students who completed the test.

Table 13 *Person measure*

Student number	Total score	Number of questions	Logit value	Person Measure
P17	23	25	3.01	High
N13	21	25	2.08	High
P23	21	25	2.08	High
P26	21	25	2.08	High
P31	21	25	2.08	High
N8	20	25	1.74	High
P25	20	25	1.74	High
P32	20	25	1.74	High
N6	19	25	1.44	High
P19	19	25	1.44	High
P30	19	25	1.44	High
P18	18	25	1.16	High
P22	18	25	1.16	High
P28	18	25	1.16	High
P29	18	25	1.16	High
P33	18	25	1.16	High
N2	17	25	0.90	Moderate
N9	17	25	0.90	Moderate
N11	17	25	0.90	Moderate
P21	17	25	0.90	Moderate
P24	17	25	0.90	Moderate
N4	16	25	0.65	Moderate
N10	16	25	0.65	Moderate
N14	16	25	0.65	Moderate
P16	16	25	0.65	Moderate
P27	16	25	0.65	Moderate
N7	15	25	0.40	Moderate
N12	15	25	0.40	Moderate
P15	15	25	0.40	Moderate
N1	14	25	0.15	Moderate
P20	14	25	0.15	Moderate
P34	13	25	-0.09	Moderate
N3	9	25	-1.13	Low
N5	8	25	-1.43	Low

The letters “N” and "P" in the student numbers represent their respective schools: 14 students from School N and 20 from School P. The student with the highest score was P17, who scored 23 and had a logit value of 3.01. Meanwhile, students N13, P23, P26, and P31 all obtained the same logit value of 2.08, with scores of 21, ranking second-highest after P17. The lowest-scoring student was N5, with a score of 8 and a logit value of –1.43. Overall, 16 students were classified as high-ability, with scores ranging from 18 to 23. Among these, P18, P22, P28, P29, and P33 had the lowest logit value in this category at 1.16. Similarly, 16 students were classified in the moderate-ability category, with logit values ranging from –0.9 to 0.90. Students N2, N9, N11, P21, and P24 achieved the highest logit values in this category at 0.90. Only two students fell into the low-ability category: N3 and N5, with logit values of –1.13 and –1.43, respectively. The determination of student ability levels based on logit values follows the recommendations of Sumintono and Widhiarso (2014).

3.7 Person Fit Order

Rasch's modeling can also detect whether individuals exhibit inappropriate responses given their abilities. This is evident in the individual not meeting the three-person fit criteria, namely the MNSQ, ZSTD, and PT Measure CORR values. If the student does not meet at least two criteria, namely PT Measure Corr and MNSQ, then the student’s response is considered not to meet the criteria, and the scalogram tests should be conducted (Boone, 2016; Sumintono & Widhiarso, 2014). Meanwhile, based on Table 14, three students (N8, P19, and P33) did not meet the MNSQ score or the PT Measure CORR. This means there is a discrepancy between the answers given and the students' ability. Therefore, a scalogram test was conducted on these three students to identify the cause of this incompatibility.

Table 14 *Person fit order analysis*

Student number	Infit		Outfit		PT Measure CORR
	MNSQ	ZSTD	MNSQ	ZSTD	
N1	1.27	1.13	1.29	0.86	0.45
N2	1.01	0.13	0.85	-0.17	0.51
N3	1.05	0.28	0.97	0.10	0.60
N4	0.88	-0.51	0.73	-0.57	0.59
N5	1.22	0.83	1.50	0.93	0.51
N6	0.92	-0.27	0.68	-0.36	0.50
N7	0.92	-0.28	0.86	-0.27	0.58
N8	1.12	0.50	1.93	1.24	0.30
N9	0.92	-0.31	1.08	0.33	0.52
N10	0.84	-0.70	0.74	-0.54	0.60
N11	0.77	-1.08	0.59	-0.88	0.61
N12	1.15	0.72	1.15	0.51	0.48
N13	1.24	0.78	1.33	0.63	0.26
N14	0.75	-1.19	0.58	-1.04	0.65
P15	0.97	-0.05	0.86	-0.27	0.57
P16	0.75	-1.19	0.58	-1.04	0.65
P17	1.13	0.41	0.71	0.24	0.25
P18	1.03	0.23	0.97	0.12	0.46
P19	1.52	1.93	2.04	1.48	0.16
P20	0.98	-0.01	1.30	0.90	0.54
P21	0.67	-1.65	0.50	-1.15	0.66
P22	0.88	-0.50	0.62	-0.63	0.55
P23	0.74	-0.73	0.45	-0.47	0.50
P24	1.20	0.95	1.13	0.42	0.42
P25	0.81	-0.64	0.54	-0.49	0.51
P26	0.88	-0.26	0.55	-0.31	0.45
P27	1.06	0.34	0.98	0.10	0.50
P28	1.12	0.61	0.98	0.14	0.43
P29	0.84	-0.67	0.60	-0.68	0.56
P30	1.38	1.48	1.32	0.67	0.28
P31	0.76	-0.66	0.47	-0.44	0.49
P32	0.84	-0.51	1.05	0.31	0.45
P33	1.16	0.75	1.98	1.50	0.36
P34	1.17	0.75	1.21	0.67	0.50

Note: MNSQ = mean square fit statistic; ZSTD = standardized fit statistic; PT Measure Corr = point–measure correlation.

3.8 Scalograms Test

Based on Figure 6, Student N8 failed to answer Questions B5, C4, A8, A7, and A4, which have logit values of -1.66, -0.01, 1.00, 1.55, and 2.73, respectively. Meanwhile, Student N8 has a person measure logit of 1.74. This indicates that Questions B5, C4, A8, and A7 should have been correctly answered by Student N8, as this student's logit value exceeds the difficulty of these items.

```
    |11 111 22 1221212 11
    |1223451123834756080956794
    |-------------------------
 17 +1111111111111111101111011  P17
 13 +1111111110111110110111101  N13
 23 +1111111111111111101110100  P23
 26 +1111111111111111101010110  P26
 31 +1111111111111111110011100  P31
  8 +1111101111110111101111010  N08
 25 +1111111111111011101101100  P25
 32 +1111111101111111111110000  P32
  6 +1111111111110101111001100  N06
 19 +1111110111001111101011101  P19
 30 +1111111101111001100110111  P30
 18 +1111111011111001111100010  P18
 22 +1111111111111001001011100  P22
 28 +1111111110110110011101010  P28
 29 +1111111111111011001010010  P29
 33 +1110111111111110101101000 1  P33
  2 +1111111111110100011110010  N02
  9 +1111011111111110110000100  N09
 11 +1111111111111111000001010  N11
 21 +1111111111111101001100000  P21
 24 +1111111100111001011101100  P24
  4 +1111111111100111011000000  N04
 10 +1111111011111101000101000  N10
 14 +1111111111110100110001000  N14
 16 +1111111111111001010011000 0  P16
 27 +1111110111110100100011100  P27
  7 +1111110111111001110000010 0  N07
 12 +1111101110011110110100000  N12
 15 +1111111110110110010010100 0  P15
  1 +1111011101011010010010010  N01
 20 +1111111111101001000000011  P20
 34 +1110111010111000010101000  P34
  3 +1101111010001010000000000  N03
  5 +1111000001100000001001000 0  N05
    |-------------------------
    |11 111 22 1221212 11
    |1223451123834756080956794
```

Fig. 6 *Scalograms test*

Similarly, Student P19 failed to answer Questions A1, B8, C3, A8, B9, and A9, which have logit values of –1.30, –0.77, –0.77, 1.00, 1.00, and 1.84, respectively. Table 13 shows that Student P19 has a person measure logit of 1.44. This suggests that P19 should have been able to correctly answer Questions A1, B8, C3, A8, and B9, since the student's logit value is higher than those of these items. Student P33 failed to answer Questions B3, C5, B10, A9, A6, A7, and A9, which have logit values of –2.14, 0.59, 0.73, 1.00, 1.13, 1.55, and 1.84, respectively. Meanwhile, Student P33 has a person measure logit of 1.66. This indicates that P33 should have been able to correctly answer Questions B3, C5, B10, A9, A6, and A7, as their logit value exceeds the difficulty logits for these items. The scalogram analysis indicates inconsistencies between the expected response patterns based on person measures and the actual responses of Students N8, P19, and P33 on several items. These discrepancies suggest a mismatch between item difficulty and response behavior for these individuals. The observed person-misfit patterns identified through the scalogram analysis should be interpreted with caution. Such inconsistencies indicate a mismatch between expected response patterns and the actual responses of several individuals, but the underlying causes cannot be determined from the present data. Further investigations using additional data, such as response-time analysis or qualitative follow-up, may help clarify the factors contributing to these response patterns.

4. Conclusion

This study successfully developed a visualization ability test for the technical drawing (GKT) subject in Malaysia and provides preliminary evidence of its validity and reliability by adapting several standardized international assessments. Expert evaluation demonstrated a very high level of agreement on content validity, with an overall Fleiss' Kappa value of 0.85. Reliability analysis using the Rasch model indicated good item reliability (0.84) and satisfactory internal consistency (Cronbach's Alpha = 0.77), while person reliability was moderate (0.62). Further analysis showed that the test items effectively distinguished students' visualization abilities and consistently measured the intended construct, although a few items were retained based on expert justification. The findings provide preliminary calibration evidence suggesting that the instrument may function as a valid and reliable measure of spatial visualization skills in technical drawing contexts. At the secondary school level, it provides educators with a diagnostic tool to identify students' strengths and weaknesses in visualizing three-dimensional objects and their orthographic projections, enabling targeted instructional interventions. It can also serve as an evaluative tool to monitor the development of spatial skills and inform curriculum design. The instrument also has potential applicability at the university level across disciplines that require advanced spatial reasoning, such as architecture, engineering, industrial design, and other STEM programs, where it can support the assessment and development of students' spatial abilities.

Despite these strengths, the study has several limitations. First, the sample size for the pilot study was relatively small, limited to two schools, which may limit the generalizability of the findings. Second, while the Rasch model provided a robust psychometric evaluation, it did not capture qualitative insights into the cognitive processes students employ when solving visualization tasks. Third, some items exhibited extreme scores or minor misfit, highlighting the need for careful item review and potential refinement. In addition, differences in instructional approaches, curriculum pacing, and students' prior familiarity with technical drawing content across schools may have influenced performance independently of visualization ability, which was not explicitly controlled in this study. These results provide an overview of response pattern inconsistencies among several respondents within the pilot dataset.

Future research should address these limitations by administering the instrument to larger and more diverse samples across multiple educational contexts to enhance external validity. Incorporating qualitative methods, such as think-aloud protocols, could provide a deeper understanding of students' reasoning strategies and sources of errors. Longitudinal studies are recommended to investigate how spatial visualization skills develop over time and how instructional interventions influence these skills. Moreover, integrating technology-based tools, such as 3D simulations or virtual manipulatives, is encouraged to enrich assessment and provide more authentic measures of visualization capabilities, aligning with forward-looking trends in educational assessment. Finally, continued use of the Rasch model is advised to refine items, enhance discriminatory power, and ensure that the instrument remains sensitive across varying levels of student ability. Overall, this study provides a robust methodological framework for developing and validating psychometric tools in educational settings. Its implications extend beyond technical drawing, offering a reference for assessment design in STEM and vocational education at both secondary and tertiary levels, and potentially guiding curriculum development, instructional practices, and targeted development of spatial reasoning skills.

Acknowledgement

We would like to express our appreciation for the support provided by all parties directly or indirectly involved in this research. We also thank our colleagues and collaborators for their support and assistance throughout the research process.

Conflict of Interest

The authors declare that there is no conflict of interest regarding the publication of the paper.

Author Contribution

The authors are responsible for the study conception, research design, data collection, data analysis, result interpretation and manuscript drafting.

References

Abbasian, M. (2024). Enhancing understanding of the rotating magnetic field in electric machines through active learning and visualization. Engineering Reports, 6(11), e12976. https://doi.org/doi.org/10.1002/eng2.12976

Andrés Rodríguez-Nieto, C., Bongani Dhlamini, Z., Singh Chauhan, A., Baltaeva, U., Abubakar, A., Dejarlo, J. O., & Andriani, M. (2023). Ways of thinking 3D geometry: Exploratory case study in junior high school students. Polyhedron International Journal in Mathematics Education, 1(1), 15–34. https://doi.org/10.59965/pijme.v1i1.5

Aziz, A. A., Masodi, M. S., & Zaharim, A. (2013). Asas model pengukuran rasch pembentukan skala & struktur pengukuran (Edisi Pert). Penerbit Universiti Kebangsaan Malaysia.

Aziz, A. A., Zaharim, M. S., & Azmai, M. (2017). Asas model pengukuran Rasch: Pembentukan skala dan struktur pengukuran (3rd ed.). Penerbit UKM (ISBN 9789674121822).

Bartlett, K. A. (2024). The Politics of the Purdue Spatial Visualization Test of Rotations (PSVT: R) and its use in engineering education. Engineering Studies, 16(1), 56–77. https://doi.org/10.1080/19378629.2023.2297958

Bartlett, K. A., & Camba, J. D. (2023). Is the PSVT:R suitable for evaluating spatial skill in design? A critique. In J. S. Gero (Ed.), Design Computing and Cognition'22 (pp. 115–126). Springer International Publishing. https://doi.org/10.1007/978-3-031-20418-0_8

Bartlett, K. A., & Dorribo Camba, J. (2023). The role of a graphical interpretation factor in the assessment of Spatial Visualization: A critical analysis. Spatial Cognition & Computation, 23(1), 1–30. https://doi.org/10.1080/13875868.2021.2019260

Boone, W. J. (2016). Rasch model for instrument development: Why, when, and how? CBE Life Sciences Education, 15(4), 1–7. https://doi.org/10.1187/cbe.16-04-0148

Chakraburty, S., Glazewski, K. D., Hmelo-Silver, C. E., Svetina Valdivia, D., Ottenbreit-Leftwich, A., Mott, B., & Lester, J. (2025). Measuring upper-elementary students' understanding of AI concepts – a Rasch model analysis. Information and Learning Sciences, 126(7–8), 445–471. https://doi.org/10.1108/ILS-10-2023-0164

Combrinck, C. (2020). Is this a useful instrument? An introduction to Rasch measurement models. In S. Kramer, S. Laher, A. Fynn, & H. H. V. J. Vuuren (Eds.), Online Readings in Research Methods (pp. 1–26). Psychological Society of South Africa. https://doi.org/10.17605/OSF.IO/ BNPFS

Gummaluri, V. S. S. S., Kumar, S., Chamarty, S., & Surendra Babu, B. (2025). Spatial visualization as the impetus for additive manufacturing education – An insight from computer-aided design perspective. International Journal of Mechanical Engineering Education, 1(2025), 03064190251330196. https://doi.org/10.1177/03064190251330196

Guo, J., Lyu, B., Zhang, J., & Qi, X. (2025). The impact of error diagnosis in handwriting practice on learners' Chinese character production and orthographic awareness. Computer Assisted Language Learning, 1(2025), 1–31. https://doi.org/10.1080/09588221.2025.2505917

He, P., Zhai, X., Shin, N., & Krajcik, J. (2023). Applying rasch measurement to assess knowledge-in-use in science education. In X. Liu & W. J. Boone (Eds.), Advances in Applications of Rasch Measurement in Science Education (pp. 315–347). Springer International Publishing. https://doi.org/10.1007/978-3-031-28776-3_13

Ilanković, N., Živanić, D., & Gašić, V. (2025). The shift to 3D education in mechanical engineering: Adapting to generation Z visualization needs. International Journal of Mechanical Engineering Education, 1(2025), 1–17. https://doi.org/10.1177/03064190251329578

Kiong, T. T., Yee, C. S., Wulansari, R. E., Kaur, C., Azid, N., & Hanapi, Z. (2023). Challenges and Teaching Needs for Technical Communication Graphics Teachers at National Secondary Schools. ONLINE JOURNAL FOR TVET PRACTITIONERS, 8(2). https://doi.org/10.30880/ojtp.2023.08.02.005

Korkut, E. H., & Surer, E. (2023). Visualization in virtual reality: A systematic review. Virtual Reality, 27(2), 1447–1480. https://doi.org/10.1007/s10055-023-00753-8

Lohman, D. F. (2013). Spatial ability and g. In I. Dennis & P. Tapsfield (Eds.), Human Abilities: Their Nature and Measurement (1st ed., pp. 94–114). Psychology Press. https://doi.org/10.4324/9780203774007.

Mackiewicz, J. (2025). Learning Technical Drawings. In Learning Skilled Trades in the Workplace (pp. 123–144). Springer Nature Singapore. https://doi.org/10.1007/978-981-96-3396-8_8

Nordin, M. S. (2009). Kesan kaedah pengajaran lukisan kejuruteraan menggunakan pendekatan serentak-pemodelan bongkah tiga dimensi terhadap kemahiran visualisasi pelajar. Universiti Teknologi Malaysia.

Piri, Z., & Cagiltay, K. (2024). Can 3-dimensional visualization enhance mental rotation (MR) ability?: A systematic review. International Journal of Human–Computer Interaction, 40(14), 3683–3698. https://doi.org/10.1080/10447318.2023.2196161

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. Science, 171(3972), 701–703. https://doi.org/10.1126/science.171.3972.701

Setiyawan, A., Soeharto, S., Yunianto, W., Cahyono, A. N., Kasti, H., & Lavicza, Z. (2025). How orthographic projection engineering drawing support VHS students in their future career: An industrial perspective. International Journal of Mechanical Engineering Education, 1(2025), 1–27. https://doi.org/10.1177/03064190251377899

Sorby, S. A. (2009). Educational research in developing 3-D spatial skills for engineering students. International Journal of Science Education, 31(3), 459–480. https://doi.org/10.1080/09500690802595839

Sumintono, B., & Widhiarso, W. (2014). Aplikasi model rasch untuk penelitian ilmu-ilmu sosial (B. Trim (ed.)). Trim Komunikata Publishing House.

Suryadi, N. R. S., Nurmegawati, L., Mu'aziyah, S. E. S., & Perdani, A. S. (2025). Rasch model for analysis of scientific attitude instruments in the context of secondary school science education. Equator Science Journal, 3(2), 98–106. https://doi.org/10.61142/esj.v3i2.234

Tandon, S., Abdul-Rahman, A., & Borgo, R. (2023). Measuring effects of spatial visualization and domain on visualization task performance: A comparative study. IEEE Transactions on Visualization and Computer Graphics, 29(1), 668–678. https://doi.org/10.1109/TVCG.2022.3209491

Thompson, A. C. (2012). The effect of enhanced visualization instruction on first grade students' scores on the north carolina. standard course assessment. Liberty University.

Tiwari, S., Shah, B., & Muthiah, A. (2024). A global overview of SVA — spatial – visual ability. Applied System Innovation, 7(3), 1–28. https://doi.org/10.3390/asi7030048

Van Zile-Tamsen, C. (2017). Using Rasch model to inform rating scale development. Research in Higher Education, 58(8), 922–933. https://doi.org/10.1007/s11162-017-9448-0

Wahab, R. A. (2018). Pembelajaran geometri tiga dimensi melalui perisian sketch up make bagi mempertingkatkan kemahiran visual spatial dan tahap pemikiran geometri. Universiti Teknologi Malaysia.

Yuan, Q., Shou, K., Xu, J., Yao, Z., & Pan, Z. (2024). Examining the impact of head-mounted augmented reality on learning engineering drawings: A case study for three-view drawing. Computer Applications in Engineering Education, 32(3), e22716. https://doi.org/10.1002/cae.22716