ORIGINAL ARTICLE

Volume 18 No. 3 2026

DOI: 10.21315/eimj2026.18.3.4

ARTICLE INFO

Submitted: 19-01-2025
Accepted: 08-02-2026
Online: 30-09-2026

Development of a Rubric for Deep and Interactive Learning in the Pharmacy Virtual Simulation Environment for College Students

Weinie Li[1,2], Adibah Abdul Latif[2]

[1]*Hainan Medical University, Hainan, CHINA*

[2]*Faculty of Educational Sciences and Technology, Universiti Teknologi Malaysia, Skudai, Johor, MALAYSIA*

To cite this article: Li W, Latif AA. Development of a rubric for deep and interactive learning in the pharmacy virtual simulation environment for college students. Education in Medicine Journal. 2026; 18(3):47–69. https://doi.org/10.21315/eimj2026.18.3.4

To link to this article: https://doi.org/10.21315/eimj2026.18.3.4

ABSTRACT

Pharmacy virtual simulation experiments have become a technical means in higher education to simulate the real world and help students acquire skills or experience. They help students understand complex concepts more deeply through visualisation and interactive operations and enhance skills in a safe, controlled setting. However, limited realism, limited scenarios and sensory experiences, the lack of face-to-face human interaction, a focus on technique or procedure rather than soft skills, and virtual simulation cannot automatically trigger deep, interactive learning. This study aimed to develop a rubric to measure students' deep, interactive learning in the pharmacy virtual simulation environment, aligned with the requirements of future pharmacists. It involved a multi-phase process to ensure the rubric's content and construct validity, integrating the literature review, expert review, students' feedback and statistical analysis using the Rasch measurement model. The participants were second-year pharmacy students from a Chinese medical university. The rubric had 10 dimensions, each with 4 levels. It provided a clear set of task performance criteria and had good usability and administrability for college students' self-assessments. The Rasch analysis showed that the rubric was unidimensional, indicating that the rubric items functioned together to measure a single underlying construct, integrating deep and interactive learning in a pharmacy virtual simulation environment. This rubric empowered students to objectively self-assess complex learning performance and encouraged them to engage deeply with the learning material. It addressed current assessment needs and laid the groundwork for future innovations in educational assessment.

Keywords: *Deep learning, Interactive learning, Pharmacy, Virtual simulation, College students*

CORRESPONDING AUTHOR

Weinie Li, Hainan Medical University, Haikou 571101, Hainan, China

Email: liweinie0928@gmail.com

INTRODUCTION

Experimentation is essential for pharmacy college students. Due to limitations in laboratory space, cost, and other factors, some crucial experiments cannot be conducted or experienced in natural settings (1). As a result, virtual simulation experiments have become a technical means in higher education to simulate the real world and help students acquire skills (2). They help students understand complex concepts more deeply through visualisation and

© Malaysian Association of Education in Medicine and Health Sciences and Penerbit Universiti Sains Malaysia. 2026
This work is licensed under the terms of the Creative Commons Attribution (CC BY)
(http://creativecommons.org/licenses/by/4.0/).

interactive operations (3,4). The immersion and interactivity characteristics of virtual simulation can enhance student engagement (5,6). Overall, they have shown the potential to promote deep and interactive learning. However, the literature has also shown that learning should be regarded as the result of the combined action of scientific teaching design, precise measurement, and dynamic evaluation, rather than solely on the virtual simulation itself (7,8). Some students might not understand the experimental process and only perform mechanical operations in the virtual simulation interface, thus forming shallow cognition (9).

Therefore, it is vital to effectively use virtual simulation experiments to support students' deep, interactive learning and to measure whether students have achieved it. Assessment plays a crucial role in shaping college students' deep and interactive learning behaviour, especially when it is ongoing and formative rather than solely summative (10). Compared with other assessment tools, such as scales and checklists, the most significant advantage of the rubric is that it is not only a consistent and fair measurement tool, but also has multiple dimensions (11). Each dimension is divided into different levels, each with a detailed description. These dimensions empower students to assess their work against explicit criteria, prompting them to reflect on their learning processes and outcomes (12,13). However, validated rubrics are rare for deep and interactive learning in the pharmacy experiment simulation environment. The limited rubrics measured students' performance and procedural skills in specific virtual simulation experiments (14–17).

This study aims to develop a valid and reliable rubric to assess students' deep, interactive learning in pharmacy virtual simulation experiments. It helps teachers provide an operable, clear, and validated rubric to reduce their time spent individually evaluating students' work through descriptive comments. It also provides students with self-assessment tools to foster deep, interactive learning. It addresses the following questions: What is the rubric's content validity and construct validity? How reliable is the rubric? What is the relationship between item difficulty and person ability? What are the results of the rating scale analysis of this rubric? What is the students' feedback after using the rubric?

LITERATURE REVIEW

Deep Learning

Research on deep learning has stemmed from a focus on students' learning processes. Marton and Säljö (18) found that students had two learning styles: superficial and deep. Bloom et al. (19) proposed a hierarchical classification of cognitive processes, ranging from basic recall to higher-order thinking, including application, analysis, synthesis and evaluation. In the revised taxonomy, Anderson and Krathwohl (20) reconceptualised these categories as active cognitive processes (applying, analysing, evaluating and creating). These higher-order cognitive processes are commonly associated with deep learning in contemporary educational research (21).

Besides cognitive aspects, deep learning also focuses on context or situational aspects. Sawyer (22) found that deep learning included establishing new and old knowledge connections, constructing conceptual systems, abstracting and expressing patterns and principles, evaluating and solving problems, engaging in deep dialogue and reflecting. The U.S. National Research Council (23) proposed a framework of student competencies for deeper learning, comprising the cognitive, interpersonal, and intrapersonal domains.

The Alliance for Excellent Education also suggested deep learning, including how deeply students understand the content, how they apply knowledge to solve problems, their ability to communicate using a range of media and their ability to collaborate with peers, self-direct and provide feedback (24). The Programme for International Student Assessment aimed to measure students' deep learning ability—that is, whether they could apply knowledge to a real-world problem and analyse, reason and communicate as they formulate, interpret and solve problems in various contexts (25).

Assessment of Deep Learning

Likert scales are the most popular form of assessment in the literature. However, the dimensions that make up these scales provide valuable material for this study. Panadero et al. (26) developed the deep learning strategies questionnaire, which contains four dimensions: basic learning self-regulation strategies, visual elaboration and summarising strategies, information processing strategies and social learning and self-regulation strategies. The questionnaire shows that interactive learning and deep learning are interrelated, as indicated in other studies (27). Nelson Laird et al. (28) developed the National Survey of Student Engagement with the following sub-scales: sub-scale of higher-order learning, which is based on Bloom's Taxonomy; sub-scale of integrative learning, which emphasises the importance of sourcing and integrating information from various places, encouraging discussions with teachers or peers; and sub-scale of reflective learning, which examines the strengths and weaknesses of one's viewpoints. Compared with the scale of Panadero and Romero (29), this scale also includes interactive learning, deep information processing and reflection, but does not mention visual elaboration and summarising strategies.

Some Chinese studies have also focused on assessing deep learning. Li (30) designed a scale for an intelligent classroom environment that includes four dimensions: knowledge transfer, innovative thinking, interaction, and emotional psychology. Similar to the scale developed by Panadero et al. (26), the knowledge transfer dimension includes connecting and integrating new and old knowledge, applying knowledge to solve new problems and linking knowledge from other disciplines. This scale also involves innovative thinking and emotional psychology. Another Chinese scholar divided emotional psychology into two types: positive emotions manifested by curiosity, joy and pride and negative emotions manifested by anxiety, depression, boredom or confusion (31).

Zhu (32) introduced a scale within Science, Technology, Engineering, Arts, and Mathematics (STEAM) education, delineating four dimensions. First, thematic integration involves synthesising pertinent content from interconnected subject areas. Second, knowledge construction underscores students' ability to apply acquired knowledge to analyse real-world problem scenarios. Third, emotional investment denotes students' capacity to exhibit subjective initiative, actively engage in learning activities and derive pleasure or other positive feelings from the learning process. Finally, thinking diagnosis highlights students' adept use of information technology to address challenges and cultivate advanced forms of thinking. Although Zhu's (32) scale is based on STEAM education, it is similar to the one designed by Li (30).

In addition to the dimensions themselves, scholars have also provided evidence for dimensional measurement. Regarding the dimension of reflection, Masuku et al. (33) viewed deep learning as a reflective learning process rather than merely completing a task. Moreover, they suggested that lectures allow students to reflect on and assess their learning (33). Stupans et al. (34) asserted that reflection should connect performance criteria and

related evidence. Er et al. (35) recognised that reflection reports should be written based on learning experiences, supported by evidence of participation in activities, and include an improvement plan.

In summary, knowledge transfer and integration, interactive learning and emotional psychology were the dimensions that appeared frequently. Summarising and visual elaboration, deep information processing, self-regulation, social learning, knowledge construction, reflection, and high-order thinking (or critical thinking) were also reported in the literature.

Interactive Learning

Interactive learning, grounded in social constructivist theory, emphasises the role of social interaction in cognitive development (36). However, researchers have different views about the elements of interaction. For Simpson and Galbo (37), interaction includes all behaviours in which individuals and groups act upon one another, with the essential characteristic of reciprocity. Wagner (38) defined interaction as reciprocal events that require at least two objects and two actions. Moore (39) identified three types of interaction: learner–content, learner–instructor and learner–learner. Hillman et al. (40) argued that Moore's three types do not account for the interactions between a learner and the technologies that deliver instructions and content, therefore adding a fourth: learner–interface interaction. Sabry and Baldwin (41) used 'information' instead of 'content' in Moore's learner–content interaction, indicating that a broader meaning should include the extensive scope of the course and non-course materials. Holden and Westfall (42) distinguished asymmetrical from symmetrical interaction and asserted that asymmetrical interaction involves one-way communication, whereas symmetrical interaction is equally balanced between the parties. Ashwin (43) believed that interaction is not only face-to-face but also includes online and offline forms, in-class and out-of-class.

Therefore, scholars have regarded interaction as an organic integration of teacher–student, student–student and student–content interactions. It offers greater advantages in increasing learners' reciprocity.

Assessment of Interactive Learning

The scales teachers use to evaluate interactive learning vary by course. A rating scale for small-group interactions in a language course focused on pronunciation, intonation, grammar, vocabulary, the expressed information, ideas and feelings (44). In a marketing course, Pearce and Lee (45) highlighted that oral presentations typically assess a range of skills, including subject knowledge, confidence, conciseness and quality of responses, on-the-spot thinking, communication abilities, application of theory to practice, question handling, body language, professional demeanour, and response clarity. Nevertheless, it is also possible to summarise the dimensions of interactive learning based on the types of interactions. Learner–learner interactions primarily involve effectively expressing opinions and synthesising others' views (44,45) . The learner–content interaction is mainly characterised by using technology tools and learning from mistakes (46). The learner–instructor interaction is a reflection of teachers' feedback (47).

Therefore, deep learning and interactive learning are closely linked in education. The literature on deep learning encompasses various dimensions, such as interactive and social learning. Notably, deep learning is also a focal point in the field of interactive learning. However, only a few rubrics are available for deep learning or interactive learning for college students, especially in virtual simulation environments.

Rubrics as an Appropriate Assessment Tool for Deep and Interactive Learning in Virtual Simulation Environments

Assessment plays a crucial role in shaping college students' learning behaviour, especially when it is ongoing and formative rather than solely summative. Formative assessment provides continuous feedback, helping students identify their strengths and areas for improvement. This ongoing feedback loop encourages learners to reflect on and revise their learning strategies, fostering self-monitoring and self-regulation of learning behaviour (48). Moreover, when assessments are designed to involve students actively through self- or peer assessment, students' intrinsic motivation, responsibility, and autonomy increase, leading to more sustained and self-directed learning behaviours (10).

Rubrics are an appropriate assessment tool for deep and interactive learning in a pharmacy virtual simulation environment. It is well suited for evaluating deep, interactive learning, especially in complex, performance-based environments such as virtual simulations. Virtual simulations are interactive settings in which students must integrate their knowledge and skills to solve realistic and complex scenarios (49,50). In such environments, rubrics provide a structured and transparent assessment framework that aligns with the goals of deep and interactive learning, capturing the process and the results of learning and prompting students to reflect on their learning processes and outcomes (51). It can even be said that rubrics are not just a tool for measuring learning, but also a mechanism to promote deeper engagement and guide students towards meaningful interaction with virtual simulations (52).

Rubrics have been successfully employed to enhance the effectiveness of simulation-based learning. The Implementation Quality Rubric for Simulation has been used in medical education to improve the quality of simulation-based courses (12). Rubrics help identify implementation gaps and monitor progress, ensuring that simulations achieve the desired learning outcomes by providing clear, structured criteria for both instructors and learners. Similarly, the Simulation-Positive Learning Evaluation rubric has demonstrated positive reliability in assessing clinical simulation performance, reinforcing its role in guiding student learning through well-defined expectations and feedback (53). The rubric designed for childhood pneumonia simulation helped nursing students reflect on their clinical skills and caring behaviours, fostering deeper cognitive and affective engagement with the simulation content (51). The scenario-specific rubrics used in deteriorating patient simulations have proven reliable and valid, enhancing learners' self-regulated preparation and performance by clarifying learning goals and providing actionable feedback (13). These examples illustrate that rubrics serve as scaffolds that clarify expectations, promote reflection and support formative feedback, which are essential for meaningful interaction with simulations. By translating complex competencies into transparent, manageable criteria, rubrics empower learners to engage actively, self-assess and internalise learning goals, making them ideally suited to guide student learning behaviour in virtual simulation environments.

Current Challenges in Virtual Simulation Environments and the Rubric Used in Pharmacy Virtual Simulation Environments

Virtual simulation environments have positive features that are beneficial to students' learning. They help students understand complex concepts more deeply through visualisation and interactive operations (54). Instant feedback and a repeatable operating environment provide reflective learning support for students (55). Immersion and interactivity characteristics can enhance student engagement (56). However, the literature also shows that learning outcomes should be regarded as the result of the combined action of multiple factors, rather than the sole reliance on the virtual simulation itself. Enquiry-based learning, blended learning, and collaborative learning, as well as assessment and evaluation, are integrated into the learning process of virtual simulation experiments (57–60).

At present, the challenges in implementing virtual simulation for deep and interactive learning include the following aspects: the first is limited realism in scenarios and sensory experiences (61). Virtual simulation cannot capture all the complexities of the real world, mainly focuses on vision or hearing and lacks other senses. The second aspect is its low-quality immediate feedback, which lacks details, context or relevance (57). The third aspect is that it primarily involves human–computer interactions and lacks face-to-face interaction (62). Finally, virtual simulation may focus on techniques and procedures while neglecting soft skills (63). In summary, a purely technical presentation in virtual simulation cannot automatically trigger deep, interactive learning. Scientific teaching design, precise measurement, and dynamic evaluation are essential elements for meeting current challenges.

To address these challenges, the development of a targeted rubric for pharmacy students in virtual simulation environments should aim to assess higher-order cognitive skills and encourage active engagement and interaction, as well as critical thinking and reflection. The existing rubrics have evaluated students' learning performance, skills or results in the virtual simulation environment (15,16,64,65). However, because the purposes or topics of evaluation vary, the dimensions of the rubrics differ. These rubrics are usually designed according to the operational procedure of simulation experiments; thus, different experiments have different rubric dimensions. Moreover, rubrics are not always validated in studies.

Therefore, rubrics on deep, interactive learning in the pharmacy simulation environment are rare. Although rubrics are applied in various fields of virtual simulation environments, some have not been tested for validity and reliability.

METHODS

Developing a rubric requires a systematic approach to ensure content and construct validity. This section outlines the study setting, participants, data collection and analysis methods.

Study Setting

The pharmacy virtual simulation examined in this study was a desktop virtual reality developed by a Chinese intelligent technology company. Figure 1 shows the students' computer interface of the pharmacy virtual simulation. After logging in on a computer,

users could enter the virtual pharmacy directly and roam its three-dimensional scenes using the keyboard and mouse. The user can create the roles of a pharmacist and a customer. The scenario in this study was a pharmacist serving a customer with hypertension.



Figure 1: Students' computer interface of the pharmacy virtual simulation.

The content of scenarios in virtual simulation experiments is often interdisciplinary due to their realism. Similarly, the scenarios in which pharmacists serve patients with hypertension involve knowledge and skills in physiology, medicine, pharmacy and health management. Therefore, the students' learning process is divided into six topics: physiology, organ damage, blood pressure measurement, drug mechanisms, health management, and patient counselling. The rubric was used six times for self-assessment by students trained by the instructor, who were asked to carefully read and review each level's descriptions of the dimensions. Instructors designed learning experiences that thoughtfully combine virtual simulations with enquiry, exploration, communication, discussion, test, assessment and reflection. Therefore, the students first operated the pharmacy virtual simulation, explored the teacher-raised scenario-related questions, and prepared their experimental reports to answer them. They could use any resources individually. Subsequently, they were divided into small groups to present their slides and discuss their ideas in the group. Finally, each student completed the multiple-choice knowledge test, drew a mind map, and wrote a reflection report.

Process of Developing the Rubric

The development process of the rubric complied with Brookhart's framework (66). First, the learning objective was identified—that is, facilitating the college students' deep and interactive learning in the pharmacy virtual simulation environment. This goal was consistent with the requirements of future pharmacists: professional knowledge and skill, communication and collaboration, critical thinking and interdisciplinary and continuous learning. Second, a literature review was conducted to determine the specific dimensions of the rubric—that is, the skills or behaviours the assessment should measure. Third, the study had been conducted since 2022, and the rubric's content was continually considered and revised. In particular, what kind of content design would make it easier for college students to understand and operate? How do we distinguish between each level of the dimensions? The rubric's content was designed on this basis. Fourth, the student sample was collected to validate the construct of this rubric, and students' feedback was collected after they participated in this study and used the rubric.

Content validity of the rubric

Content validity is defined as the degree to which a measurement instrument adequately represents all facets of the construct it aims to assess (67). First, the rubric for deep and interactive learning was generated from a literature review. Then, the teacher practised the teaching and learning process in the virtual simulation course to decide how to describe the content of each dimension and to make the differences among its levels more significant and easier to understand. Starting in 2022, teachers gradually accumulated evaluation experience, laying a solid foundation for the content design of this rubric.

Second, it was crucial to thoroughly analyse and review the written items by content experts who could provide valuable insights into the instrument items' relevance, clarity and meaningfulness (68). Five experts from the education and pharmacy fields commented on and suggested modifications individually. Third, the pharmacy students used the rubric to conduct self-assessments and to express their feelings and revision suggestions about it. Finally, the five experts reviewed the revised rubric until no amendments were needed. Thus, the modification process continued until the experts and students accepted the rubric.

Construct validity of the rubric

Rasch analysis was used to validate the rubric's construct, which is crucial because it allows for a statistical review of items beyond the scope of expert judgement (69). It can contribute to validating a rubric by assessing unidimensional measurement, item and person reliability, item and person separation, the Wright Map, item fit, and monotonicity. For construct validity, data were collected through the learning process with 73 second-year pharmacy students and included six successive students' self-assessments based on the rubric. Six subjects were extracted from virtual simulation scenarios of hypertension: pathophysiology, organ damage, diagnosis, pharmacological interventions, lifestyle management and patient counselling.

RESULTS

Content Validity and the Dimensions of the Rubric

The literature review and teaching practice were included to form the initial rubric content, and the final content was determined through expert review and student feedback. For the literature review, nine dimensions, including summarising and visual elaboration, knowledge construction, knowledge transfer and integration, higher-order thinking, self-regulation and reflection, learner–content interaction, learner–learner interaction, learner–teacher interaction and emotional psychology, were designed for the first edition of the rubric. Each dimension had five different scales. Five experts in measurement and pharmacy carefully reviewed the rubric and suggested revisions.

The experts considered all the dimensions essential for assessing deep and interactive learning. First, enabling students to conduct self-assessment objectively requires explicit emphasis on supporting evidence. The evidence that helped students conduct self-assessment should therefore be clearly specified in the rubric. Second, the words in the description were relatively abstract rather than specific and understandable to the students. Third, the rubric had so many words for students to review. Complex rubrics may intimidate

and confuse students. Thus, it is recommended that the dimension levels be reduced. Finally, each dimension should be simple and contain only one aspect to be evaluated. The dimension of 'summarising and visual elaboration' included two different aspects: summarising and visual elaboration.

After the amendment, the second edition of the rubric contained 10 dimensions: processing of learning content, production of figures or tables, higher-order thinking, construction of knowledge framework, self-reflection, interaction with learning content, presentation in the group meeting, interaction with group members, interaction with the teacher and psychological emotion. Each dimension had four levels: novice, limited, master and excellent. The rubric was reviewed again by experts. They all agreed that the second edition was more operational and accessible to students.

Finally, 30 students were asked to give their perceptions of the rubric. All students agreed that the rubric provided clear performance criteria for the task, helping them understand what was expected. The rubric served as a guide, allowing them to think critically about their work. However, three students pointed out that the dimension of higher-order thinking was difficult to evaluate.

Why was higher-order thinking abstract for students? Bloom's Taxonomy is a sophisticated educational tool that categorises cognitive processes and distinguishes between lower- and higher-order thinking. However, the abstract concepts were beyond college students' cognitive and perceptual capacities, making it difficult to accurately assess their cognitive processes (70). Therefore, knowing the 'what' and 'why' replaced Bloom's Taxonomy in the rubric, which was much easier for students to measure. Therefore, the literature review, expert review, and student feedback provided strong evidence for the rubric's content validity.

Construct Validity Using Unidimensionality Analysis and Item Fit

The unidimensionality analysis ensured that the rubric measured a single underlying construct of deep and interactive learning (71). The item fit helped assess how well individual dimensions in the rubric conformed to the expectations of the Rasch model. The results of the unidimensionality analysis are presented in Table 1.

Table 1: Results of the unidimensionality analysis

Result		Empirical (%)		Modelled (%)
Total raw variance in observations	24.7	100.0	–	100.0
Raw variance explained by measure	14.7	59.4	–	59.2
Raw variance explained by persons	11.8	48.0	–	47.8
Raw variance explained by items	2.8	11.4	–	11.4
Raw unexplained variance (total)	10.0	40.6	100.0	40.8
Unexplained variance in 1st contrast	1.7	7.0	17.2	–
Unexplained variance in 2nd contrast	1.3	5.3	13.2	–
Unexplained variance in 3rd contrast	1.2	4.9	12.2	–
Unexplained variance in 4th contrast	1.2	4.7	11.7	–
Unexplained variance in 5th contrast	1.1	4.4	10.8	–

In Table 1, the variance explained by the measurement dimension is 59.4%, which is higher than 40% (72,73), indicating that nearly 60% of the data are explained by the primary latent trait. The unexplained variance in the first contrast is 7.0%, which is no more than 15% (72), thus suggesting the poor strength of the secondary latent trait. The unexplained variance in the first contrast is 1.7 eigenvalue units, smaller than 3.0 (72,73). Finally, the ratio of the variance in the measurement dimension to the variance of the first principal component of residuals is 8:1. These results show that the rubric is unidimensional. The items' fitness levels were assessed using the outfit mean-square (MNSQ) and outfit standardised fit statistic (ZSTD). The results are presented in Table 2.

Table 2: Item fit for the rubric

Item	Total score	Total count	Measure	Model S.E.	Infit		Outfit		PT Measure	
					MNSQ	ZSTD	MNSQ	ZSTD	CORR.	EXP.
D9	916	438	0.83	0.11	1.30	4.0	1.31	3.5	0.75	0.77
D3	935	438	0.62	0.11	0.84	−2.5	0.80	−2.6	0.81	0.77
D2	956	438	0.38	0.11	1.32	4.4	1.27	3.2	0.76	0.77
D8	979	438	0.12	0.11	1.02	0.3	0.97	−0.3	0.78	0.77
D4	993	438	−0.03	0.11	1.00	0.3	0.96	−0.4	0.79	0.77
D6	998	438	−0.09	0.11	0.98	−0.3	0.95	−0.6	0.81	0.77
D7	1,000	438	−0.11	0.11	0.87	−2.0	0.85	−2.0	0.74	0.77
D10	1,039	438	−0.55	0.11	0.87	−2.1	0.86	−1.8	0.75	0.77
D5	1,041	438	−0.57	0.11	1.05	0.8	1.01	0.2	0.73	0.77
D1	1,042	438	−0.59	0.11	0.69	−5.1	0.67	−4.8	0.81	0.77

Note: S.E. = Standard error; MNSQ = Mean square; ZSTD = Standardised fit statistic; CORR. = Point-measure correlation; EXP. = Expected point-measure correlation

In Table 2, the outfit MNSQ values for each item range from 0.5 to 1.5 (74). The outfit ZSTD values of D4, D5, D6, D7, D8 and D10 are within the range of −2 to 2 (74), but the outfit ZSTD values of D1, D2, D3 and D9 are beyond this range. However, after a discussion with the experts, the four items were retained. The most significant reason was that these items were crucial for deep and interactive learning assessments. Removing them could lead to gaps in the assessment of the measured construct. Another reason was that the ZSTD values were sensitive to the sample size. If the mean-square values (MNSQ) were within acceptable ranges (typically 0.5–1.5), the ZSTD values could often be ignored (75).

Reliability of the Rubric

The rubric's reliability included person and item reliability. Person reliability and separation are statistics that help assess the quality and effectiveness of a measurement instrument (76). Person reliability indicates the consistency of the person measures. Person separation suggests how well the items separate individuals into distinct ability levels.

In Table 3, the person reliability is 0.91, which is close to 1 (77), indicating that the rubric is reliable in distinguishing between different levels of the trait. The person separation is 3.18, which is greater than 2 (77), indicating that the rubric can differentiate at least three distinct groups of abilities. These results imply that the rubric is reliable for measuring students' deep and interactive learning and can effectively separate students into different ability levels.

Item reliability and separation provide insights into the quality and effectiveness of the items in a measurement instrument. Item reliability indicates how well items can be differentiated along the measured variable. Item separation measures the spread of items along the difficulty continuum. This indicates how well the items can be distinguished from each other in terms of their difficulty levels.

Table 3: Results of person reliability and separation

Statistic	Total	Count	Measure	Model error	Infit		Outfit	
					MNSQ	ZSTD	MNSQ	ZSTD
Mean	22.6	10.0	−1.32	0.78	0.97	−0.2	0.97	−0.2
S.D.	5.5	0.0	2.69	0.19	0.67	1.4	0.70	1.4
Real RMSE		0.81	True S.D.	2.56	Separation	3.18	Person reliability	0.91

In Table 4, the item reliability is 0.95, which is close to 1.0 (78), indicating excellent item reliability. The item separation value of 4.18, which is above the recommended minimum of 3.0 (77), suggests that the items can be divided into at least 4 distinct difficulty levels. These high values of item reliability and separation indicate that the rubric creates a set of items that effectively cover the range of the latent trait and can consistently distinguish between different levels of item difficulty.

Table 4: Results of item reliability and separation

Statistic	Total	Count	Measure	Model error	Infit		Outfit	
					MNSQ	ZSTD	MNSQ	ZSTD
Mean	989.9	438.0	0.00	0.11	0.99	−0.3	0.97	−0.6
S.D.	42.0	0.0	0.47	0.01	0.19	2.7	0.19	2.4
Real RMSE		0.11	True S.D.	0.46	Separation	4.19	Item reliability	0.95

Comparison Between Item Difficulty and Person Ability Using the Wright Map

The Wright Map is a visual representation of a person's abilities and item difficulties on the same linear scale (79). It allows for a direct comparison between a person's abilities and the distribution of item difficulty.

In Figure 2, the distribution of a person's abilities ranges from −7 to +6 logits on the left, with most respondents clustering around −3 to +2, indicating that most had moderate abilities. On the right side, the rubric dimensions range from a logit value of −1 to +1. D9 is the most challenging item, while D1, D10, and D5 are the most accessible, at around 0.5 logits. The item mean is 0.00 logits, while the person mean is −1 logits, indicating that the items are relatively challenging for students. The rubric was used to measure students' deep and interactive learning, and the students needed to make considerable efforts to achieve excellent learning performance.

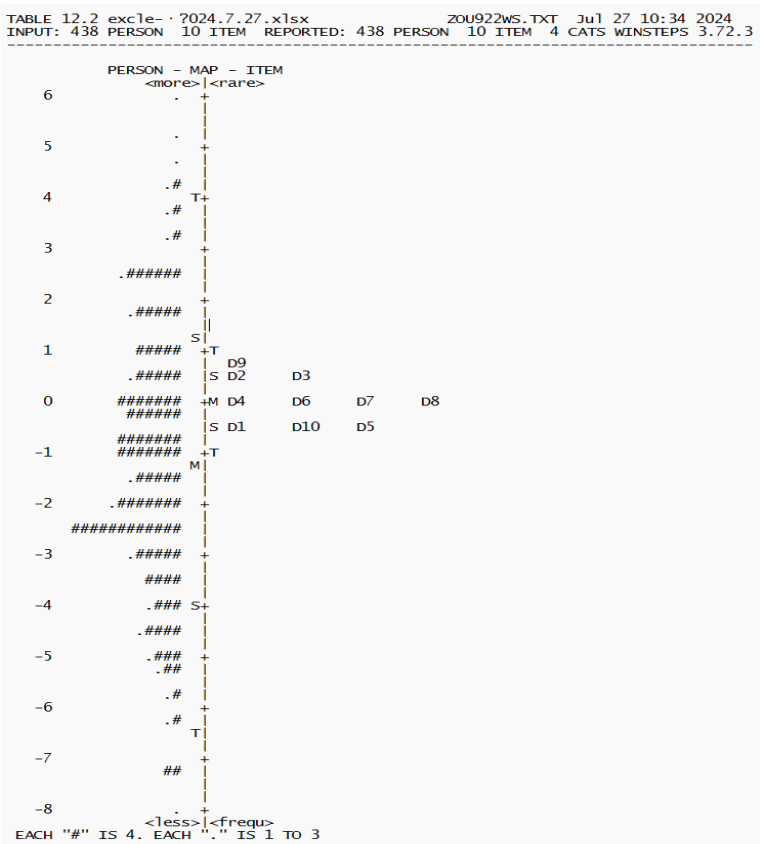


Figure 2: Results of the wright map.

Rating Scale Analysis of the Rubric

Monotonicity is essential for proper category ordering in rating scales. This means that the probability of endorsing an item or giving a correct response will continuously increase as an individual's trait level increases (80). A probability curve can be used to visually assess whether it performs optimally.

In Table 5, each category has at least 100 observations. The observed average values monotonically increase from −4.95 to 3.43. Moreover, the outfit MNSQ is close to 1, which is less than 2.0 (81). The threshold also monotonically increases, and the increase between two adjacent values is in the range of 1.4–5 logits (81), except for the increase between level 3 (master) and level 4 (excellent). The probability of responding to each category, given the difference between a person's ability and an item's difficulty, is illustrated in Figure 3. Each category has a distinct peak, indicating that participants reliably distinguished the response categories (82).

Table 5 shows 1,526 observations for the master level, but only 103 for the excellent level. Level 4 (excellent) represents a rarer performance compared with level 3 (master). The large threshold increase helps differentiate top performers from those who have achieved mastery, but not excellence. There is motivation among students who want to achieve an excellent level. A large gap between mastery and excellence is expected with continued effort and improvement.

Table 5: Summary of the category structure

Category label	Score	Observed		Observed average	Infit MNSQ	Outfit MNSQ	Andrich threshold
		Count	%				
1	1	593	14	−4.95	0.94	0.91	None
2	2	2,158	49	−2.07	0.96	0.92	−4.85
3	3	1,526	35	0.85	1.03	1.03	−0.30
4	4	103	2	3.43	1.15	1.06	5.15

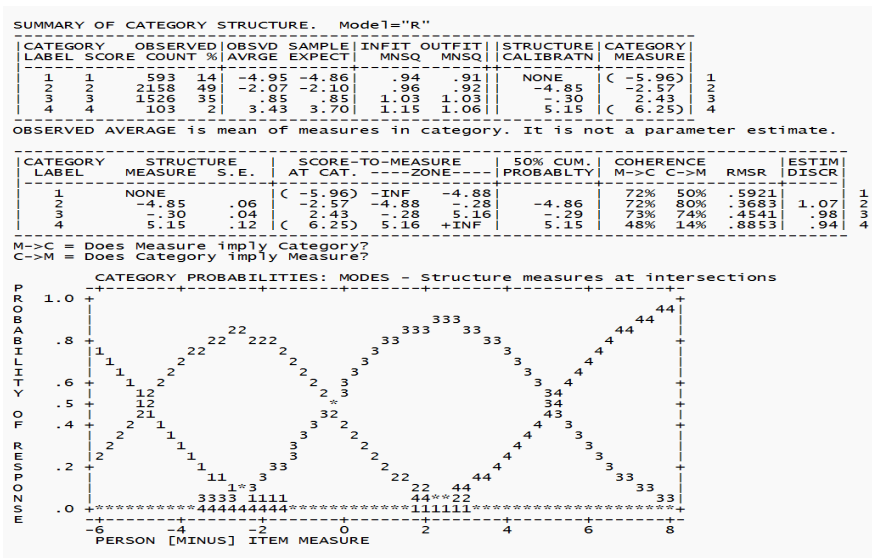


Figure 3: Probability curves for the four categories.

Changes in the Self-Assessment Levels of Each Dimension for Six Times

To compare self-assessment levels across different dimensions at six time points and show their changes, novice was assigned a value of 1, limited was assigned a value of 2, master was assigned a value of 3, and excellent was assigned a value of 4. Then, the average value of each dimension was calculated for comparison in the six self-assessments.

As shown in Figure 4, in the first self-assessment, the mean value of D10 (psychological emotion) is the largest, followed by the mean value of D7 (presentation) in the group meeting. By contrast, D2 (production of figures or tables) has the smallest mean, followed by D8 (interaction with the group members).

As the number of self-assessments increases, the average value of each dimension increases. The largest increase is in D6 (interaction with learning content), as denoted by the green line, and the second largest increase is in D8 (interaction with group members). Conversely, the smallest increase is in D7 (presentation in the group meeting) and D10 (psychological emotion). As mentioned above, D7 and D10 have relatively high average values in the first self-evaluation, which is also the reason for the small increase.

Thus, after six self-assessments, students' levels of the 10 dimensions improved, indicating that their deep and interactive learning levels indeed improved. Especially for the interactive aspect, the dimensions of interaction with learning content and interaction with the group members showed the greatest increase.

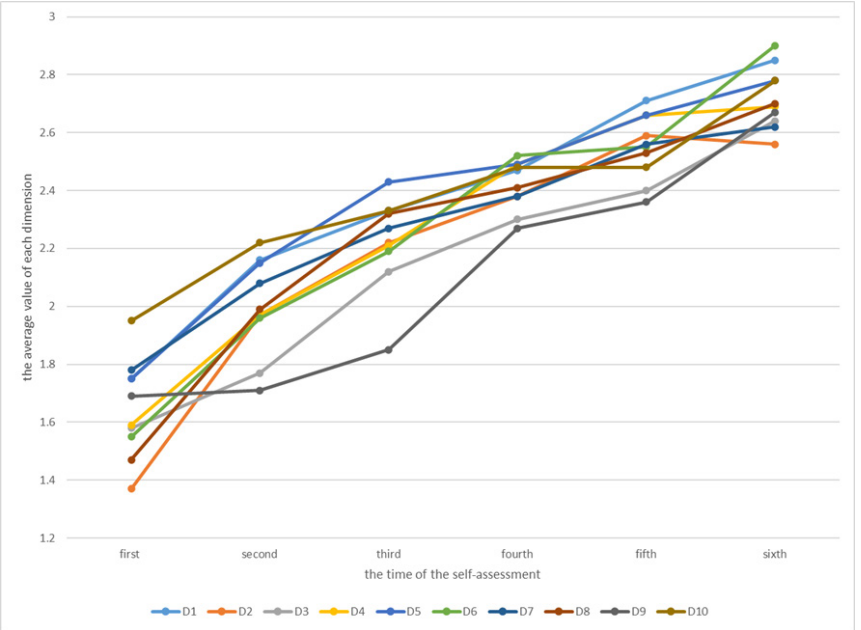


Figure 4: Changes in the mean value of the dimensions in six self-assessments.

Students' Feedback After Using the Rubric

Seventy-three pharmacy students gave feedback on the course and the rubric's usage. They all agreed that their knowledge of hypertension had greatly improved. They also mentioned their progress in creating slides, producing figures and charts, processing learning content, think independently, speaking and presenting in the group, and interacting with group members. However, some participants reported that their progress in interacting with group members was unsatisfactory because they could not take the lead within the group or create a free and positive discussion environment. At the same time, the level of interaction with the teacher needed to be improved, as a two-way interaction between students and teachers had not been spontaneously formed.

How did the rubric facilitate their deep and interactive learning? According to the respondents, they could face their shortcomings and areas for improvement from an objective and diversified perspective. As the dimensions' levels gradually increased, the requirements also increased. The rubric provided a clear set of task performance criteria and was usable and administrable to them.

DISCUSSION

The Rubric is an Appropriate Assessment Tool for Deep and Interactive Learning, Especially for Ongoing or Formative Assessment

The rubric involves a rigorous, multi-phase development process that integrates a literature review, expert review, students' review and Rasch analysis. It is a comprehensive tool for assessing the depth and interactivity of the pharmacy virtual simulation. Compared with the other assessment tools created by Panadero et al. (26), Laird (28), Wang (31), Li (30) and Zhu (32), the rubric encompasses a wide range of dimensions that capture the cognitive

and social aspects of learning and provides a holistic view of the learning process. It can be used as a tool for ongoing or formative assessment, offering actionable feedback to students and helping improve their overall learning approach. In this study, the students conducted six self-assessments and reflections on different topics related to hypertension patients. This process encouraged metacognitive skills, as the students reflected on their learning behaviours and outcomes, deepening their understanding of deep and interactive learning and identifying their strengths and areas for improvement. Feedback from the clear rubric was actionable and specific, allowing students to focus on particular dimensions and make meaningful improvements.

Within the Specific Environment of the Pharmacy Virtual Simulation, Deep Learning and Interactive Learning are Tightly Intertwined, Even as a Unified Construct

Although deep learning and interactive learning are distinct constructs in general educational theory, within the specific environment of the pharmacy virtual simulation, these processes are tightly intertwined, even as a unified construct. In other words, pharmacy virtual simulation tasks require students to engage in deep learning and interactive learning simultaneously.

As described in the literature review, visual and interactive features are the most significant features for virtual simulation (54). Visually presenting learning content and allowing students to manipulate and explore virtual objects can enhance the immersion and engagement in their learning experience. Therefore, learning in pharmacy virtual simulation environments requires students to at least interact with virtual patients (e.g., answering patient questions) to trigger deeper learning content (e.g., medication mechanisms and prescription analysis). Interaction has become a necessary condition for deep learning, leading to their inseparable behaviour.

In addition, the scenarios in pharmaceutical virtual simulation experiments come from the real world. For undergraduate students, the complexity and interdisciplinary nature of learning tasks force them to enter a deep learning state through interaction. Moreover, student interaction should not be limited to objects in virtual simulation experiments (e.g., virtual patients), but should also include peers and instructors. Alternatively, in virtual environments, student knowledge is constructed through interaction, and deep learning is a natural product of interaction rather than an independent process.

This result does not negate traditional theories that hold that deep learning and interactive learning are distinct concepts, but it reflects the particularity of learning goals in virtual simulation environments. The goal of pharmacy virtual simulation is to foster a holistic learning experience in which deep learning and interactive learning are not isolated skills, but components of a single, integrated competency.

The capabilities that future pharmacists should possess can be summarised as follows: professional knowledge and skills, communication and collaboration, critical thinking, and interdisciplinary and continuous learning (83,84). The dimensional content of the rubric is highly consistent with the comprehensive capability system for future pharmacists, which is patient-centred, systematic, and continuously developed.

The Rubric Reflects Marton and Säljö's Deep Versus Surface Learning Theory and Vygotsky's Social Constructivism

Marton and Säljö's deep versus surface learning theory describes two fundamentally different approaches (18). First, the deep learning approach seeks to understand and connect concepts, relates ideas to real-world experience, explores links between evidence and conclusions, critiques arguments and examines rationale and is motivated by genuine interest in the subject. Second, the surface learning approach focuses on rote memorisation without understanding, often resulting in a lack of engagement with the material, and motivation is often extrinsically driven or forced. In this rubric, the novice and limited levels are similar to the characteristics of surface learning, especially the novice level. The master and excellent levels are identical to the characteristics of deep learning, especially the excellent level.

A clear distinction between surface learning and deep learning can be seen in the descriptions of the different levels in the rubric. For example, the descriptions of the novice level include the following: it is a temporary memory of the learning content or just a copy and is easy to forget; new knowledge acquired from the virtual simulation scenario remains disconnected from prior knowledge or experience; the learning process is superficial or confusing, and the report content lacks logic; and there is indifference or even boredom in the learning process, and it is just about completing the task. Conversely, the master or excellent level includes the following: the experimental report is well-processed, well-structured and clearly written, allowing readers to fully absorb it, even if it has been carefully or painstakingly summarised so that readers can absorb it smoothly and quickly; it has the ability to explain the 'what' and 'why' reasonably and accurately while excluding other distracting or erroneous information; deep exploration of knowledge from the virtual simulation scenario leads to the reconstruction of a comprehensive personal knowledge system; and continued interest throughout the learning process is maintained and can even influence other group members to participate actively, thus creating a free and positive learning atmosphere in the group.

Vygotsky's social constructivism is a foundational theory in educational psychology that emphasises the essential role of social interaction in the development of learning (36). Vygotsky argued that cognitive development is not an isolated process within an individual. Instead, knowledge is co-constructed through social interactions with teachers, peers and even the broader community. The rubric dimensions include interaction with the learning content, group members, and the teacher. The excellent level encourages students to engage in proactive and spontaneous communication with group members, elaborating on the 'what' and 'why' of the topics discussed. Not only do students pay attention to the teacher's feedback, they also carefully consider improvement strategies, resulting in a dynamic two-way interaction with the teacher.

The Rubric is Suitable for Students for Conducting Self-Assessment and is Subject to Limitations When Used by Teachers to Evaluate Students

This rubric was developed specifically for student self-assessment, not for teacher evaluation of students. Some of these dimensions are more implicit and subjective, making them more suitable for students' self-cognition and inspiration but difficult for teachers to accurately evaluate students' performance. The most prominent dimension is psychological emotion. Psychological emotion is highly subjective, and it is difficult for teachers to accurately perceive and evaluate students' inner feelings based solely on external behaviour. Potgieter

et al. (85) noted that the emotional aspect of learning is highly individualised, making self-report instruments the preferred method, as teachers cannot objectively assess students' internal emotional states. The dimension of self-reflection also shows a relatively large deviation when applied to students' self-assessments and teachers' evaluations of students. Galdonez (86) found that many students in science majors struggled with producing detailed reports. These challenges were linked to a lack of humanities writing instruction and training, leading to shorter reflections that did not fully reflect their actual thinking or learning processes.

However, most other dimensions not only help teachers evaluate students' performance but also facilitate teachers in promoting students' deep and interactive learning by combining tools such as mind maps or online video conferences. By leveraging these dimensions alongside digital tools, teachers can not only obtain more comprehensive assessment evidence, but also actively guide students to engage in higher-order thinking, knowledge construction or effective participation in interactive learning tasks.

Implications for Education

This study developed a rubric for deep, interactive learning in pharmacy virtual simulation that integrated dimensions reflecting both deep and interactive learning. The rubric is a practical tool that enables pharmacy students to systematically assess and foster their deep, interactive learning in virtual simulation environments. Moreover, it can help teachers understand the content for deep, interactive learning, develop a teaching plan for virtual simulation, and implement learning activities. This rubric is not limited to a virtual simulation scenario and can be extended to other fields of learning assessment.

Limitations and Future Directions

This study has several limitations. First, it relied on students' self-assessments without test-retest or external rater validation, which may have introduced bias, as students could overestimate or underestimate their self-assessment scores due to social desirability, limited self-awareness, or misinterpretation of the assessment criteria. Second, the study population was limited to second-year pharmacy students, which may restrict the generalisability of the findings to students at different academic levels, from other institutions, or in diverse cultural and educational contexts. Third, while the rubric is suitable for self-assessment, its application by teachers for evaluating students may be limited. Future research should expand the rubric's validation across disciplines, compare rubric-based self-assessment with instructor or peer assessment, and examine the long-term impact of rubric-guided feedback on learning outcomes.

CONCLUSION

The proposed rubric was developed using a robust method and supported ongoing efforts to enhance the quality of pharmacy education. It empowered students to self-assess objectively and diversify. It addressed current assessment needs and laid the groundwork for future innovations in educational assessment. The ongoing refinement in other fields will ensure its continued effectiveness, contributing to the goal of promoting deep and interactive learning experiences for all students.

ACKNOWLEDGEMENTS

This work is supported by the Education Department of Hainan Province, project number: Hnjg2025-81.

ETHICAL APPROVAL

All participants were informed of the purpose and procedures of the study and voluntarily consented to participate and provide their data for research purposes. Participation was voluntary, and the confidentiality and privacy of the participants were protected.

REFERENCES

1. Potkonjak V, Gardner M, Callaghan V, Mattila P, Guetl C, Petrović VM, et al. Virtual laboratories for education in science, technology, and engineering: a review. Comput Educ. 2016;95:309–27. https://doi.org/10.1016/j.compedu.2016.02.002

2. Lateef F. Simulation-based learning: just like the real thing. J Emerg Trauma Shock. 2010;3(4):348–52. https://doi.org/10.4103/0974-2700.70743

3. Qorbani S, Dalili S, Arya A, Joslin C. Assessing learning in an immersive virtual reality: a curriculum-based experiment in chemistry education. Educ Sci. 2024;14(5):476. https://doi.org/10.3390/educsci14050476

4. Zheng K, Shen Z, Chen Z, Che C, Zhu H. Application of AI-empowered scenario-based simulation teaching mode in cardiovascular disease education. BMC Med Educ. 2024;24(1):1003. https://doi.org/10.1186/s12909-024-05977-z

5. Boone WJ, Noltemeyer A. Rasch analysis: a primer for school psychology researchers and practitioners. Cogent Educ. 2017;4(1):1416898. https://doi.org/10.1080/2331186X.2017.1416898.

6. Kim J, Kim K, Kim W. Impact of immersive virtual reality content using 360-degree videos in undergraduate education. IEEE Trans Learn Technol. 2022;15(1):137–49. https://doi.org/10.1109/tlt.2022.3157250

7. Casselman MD, Eichler JF, Atit K. Advancing multimedia learning for science: comparing the effect of virtual versus physical models on student learning about stereochemistry. Sci Educ. 2021;105(6):1285–314. https://doi.org/10.1002/sce.21675

8. Serrano-Perez JJ, González-García L, Flacco N, Taberner-Cortés A, García-Arnandis I, Pérez-López G, et al. Traditional vs. virtual laboratories in health sciences education. J Biol Educ. 2023;57(1):36–50. https://doi.org/10.1080/00219266.2021.1877776

9. Bridge P, Trapp JV, Kastanis L, Pack D, Parker JC. A virtual environment for medical radiation collaborative learning. Australas Phys Eng Sci Med. 2015;38(2):369–74. https://doi.org/10.1007/s13246-015-0348-3

10. Zou Y, Yuan M, Mo L, Mustakim SSB. Enhancing teaching and learning through assessment strategies: a practical guide. IIJARBSS. 2024;14(7):1024–36. https://doi.org/10.6007/ijarbss/v14-i7/21928

11. Bolt DM, Adams DJ. Exploring rubric-related multidimensionality in polytomously scored test items. Appl Psychol Meas. 2017;41(3):163–77. https://doi.org/10.1177/0146621616677715

12. Nayahangan LJ, Konge L, Park C, Dubrowski A, Yudkowsky R. Development of a rubric to evaluate implementation quality of simulation-based courses: a consensus study. Simul Healthc. 2023;18(6):351–8. https://doi.org/10.1097/sih.0000000000000684

13. Luctkar-Flude M, Tregunno D, Sears K, Pulling C, Lee K, Egan R. Reliability and validity of scenario-specific versus generic simulation assessment rubrics. J Nurs Educ Pract. 2020;10(8):74–8. https://doi.org/10.5430/jnep.v10n8p74

14. Barnett SG, Porter AL, Allen SM, Bastianelli KMS, Chen JS, Kachlic MD, et al. Expert consensus to finalize a universal evaluator rubric to assess pharmacy students' patient communication skills. Am J Pharm Educ. 2020;84(12):848016. https://doi.org/10.5688/ajpe848016

15. Baumgartner L, Ip EJ, Sasaki-Hill D, Wong T, Israel H, Barnett MJ. Implementation of mock acute care advance pharmacy practice experience simulations and an assessment rubric. Am J Pharm Educ. 2019;83(9):7331. https://doi.org/10.5688/ajpe7331

16. Schroeder MN, Lengel AJ. Evaluation of a rubric to assess pharmacy student performance in a telehealth simulation exercise. Am J Pharm Educ. 2022;86(8):ajpe8778. https://doi.org/10.5688/ajpe8778

17. Garling KA, Wong B. An initial reliability analysis of a patient counseling rubric to objectively measure student pharmacist performance. Heliyon. 2023;9(5):e15768. https://doi.org/10.1016/j.heliyon.2023.e15768

18. Marton F, Säljö R. On qualitative differences in learning: I—Outcome and process. Br J Educ Psychol. 1976;46(1):4–11. https://doi.org/10.1111/j.2044-8279.1976.tb02980.x

19. Bloom BS, Engelhart MD, Furst EJ, Hill WH, Krathwohl DR. Taxonomy of educational objectives: the classification of educational goals. Handbook 1: Cognitive domain. New York: David Mckay Company, Inc.; 1956.

20. Anderson LW, Krathwohl DR, editors. A taxonomy for learning, teaching, and assessing: a revision of Bloom's taxonomy of educational objectives. Complete edition. London: Longman, Inc.; 2001.

21. Thompson AR, Lake LPO. Relationship between learning approach, Bloom's taxonomy, and student performance in an undergraduate human anatomy course. Adv Health Sci Educ Theory Pract. 2023;28(4):1115-30. https://doi.org/10.1007/s10459-023-10208-z.

22. Sawyer RK. The new science of learning. The Cambridge handbook of the learning sciences. Cambridge: Cambridge University Press; 2006. https://doi.org/10.1017/cbo9780511816833.002

23. National Research Council. Education for life and work: developing transferable knowledge and skills in the 21st century. Washington, DC: The National Academies Press; 2012.

24. Alliance for Excellent Education. Demography as destiny: how America can build a better future [Internet]. Washington, DC: Alliance for Excellent Education; 2006 [cited 2026 Sep 20]. Available from: https://www.adlit.org/topics/policy-legislation-initiatives/demography-destiny-how-america-can-build-better-future

25. Hopfenbeck TN, Lenkeit J, El Masri Y, Cantrell K, Ryan J, Baird J-A. Lessons learned from PISA: a systematic review of peer-reviewed articles on the programme for international student assessment. Scand J Educ Res. 2018;62(3):333–53. https://doi.org/10.1080/00313831.2016.1258726

26. Panadero E, Alonso-Tapia J, García-Pérez D, Fraile J, Galán JMS, Pardo R. Deep learning self-regulation strategies: validation of a situational model and its questionnaire. Rev Psicodidáct. 2021;26(1):10–19. https://doi.org/10.1016/j.psicoe.2020.11.003

27. Means B, Toyama Y, Murphy R, Bakia M, Jones K. Evaluation of evidence-based practices in online learning: a meta-analysis and review of online learning studies. Washington DC: U.S. Department of Education, Office of Planning, Evaluation, and Policy Development Policy and Program Studies Service; 2009.

28. Laird TFN, Shoup R, Kuh GD. Measuring deep approaches to learning using the National Survey of Student Engagement. Annual Meeting of the Association for Institutional Research; 2005 May 14–18; Chicago, IL: Association for Institutional Research; 2006.

29. Panadero E, Romero M. To rubric or not to rubric? The effects of self-assessment on self-regulation, performance and self-efficacy. Assess Educ Princ Policy Pract. 2014;21(2):133–48. https://doi.org/10.1080/0969594x.2013.877872

30. Li Z, Li S, Wang Y, Zhang C. Design and verification of college students' deep learning evaluation scale under embodied cognition environment. E-Educ Res. 2020;41(12):92–8. https://doi.org/10.13811/j.cnki.eer.2020.12.013

31. Wang MK, Wang Z, Chen ZZ. Deep learning evaluation: theoretical models, related technologies and practical cases. J Guangxi Norm Univ. 2023;59(06):66–75.

32. Zhu LM. Construction of deep learning evaluation indicator system under the concept of STEM education. J Sichuan Normal Univ. 2022;49(4):125–33.

33. Masuku MM, Jili NN, Sabela PT. Assessment as a pedagogy and measuring tool in promoting deep learning in institutions of higher learning. Int J High Educ. 2021;10(2):274–83. https://doi.org/10.5430/ijhe.v10n2p274

34. Stupans I, March G, Owen SM. Enhancing learning in clinical placements: reflective practice, self-assessment, rubrics and scaffolding. Assess Eval High Educ. 2013;38(5):507–19. https://doi.org/10.1080/02602938.2012.658017

35. Er HM, Ming MKJ, Keng PS, Nadarajah VD. Pharmacy students' perceptions of reflective portfolios and their effect on students' deep information-processing skills. Am J Pharm Educ. 2019;83(6):6851. https://doi.org/10.5688/ajpe6851

36. Vygotsky LS, Cole M. Mind in society: development of higher psychological processes. Harvard University Press; 1978.

37. Simpson RJ, Galbo JJ. Interaction and learning: theorizing on the art of teaching. Interchange. 1986;17(4):37–51. https://doi.org/10.1007/bf01807015

38. Wagner ED. In support of a functional definition of interaction. Am J Distance Educ. 1994;8(2):6–29. https://doi.org/10.1080/08923649409526852

39. Moore MG. Three types of interaction. Am J Distance Educ. 1989;3(2):1–7. https://doi.org/10.1080/08923648909526659

40. Hillman DCA, Willis DJ, Gunawardena CN. Learner-interface interaction in distance education: an extension of contemporary models and strategies for practitioners. Am J Distance Educ. 1994;8(2):30–42. https://doi.org/10.1080/08923649409526853

41. Sabry K, Baldwin L. Web-based learning interaction and learning styles. Br J Educ Technol. 2003;34(4):443–54. https://doi.org/10.1111/1467-8535.00341

42. Holden JT, Westfall PJL. Instructional media selection for distance learning: a learning environment approach. Distance Learn. 2006;3(2):1–12. https://doi.org/10.1108/DL-06-2006-0001

43. Ashwin P. Conceptualising teaching and learning interactions in researching higher education. In: Tight M, Mok KH, Huisman J, Morphew C, editors. The Routledge international handbook of higher education. New York: Routledge; 2009. p. 37–46. https://doi.org/10.4324/9780203882221-4

44. Nunn R. Designing rating scales for small-group interaction. ELT J. 2000;54(2):169–78. https://doi.org/10.1093/elt/54.2.169

45. Pearce G, Lee G. Viva voce (oral examination) as an assessment method: insights from marketing students. J Mark Educ. 2009;31(2):120–30. https://doi.org/10.1177/0273475309334050

46. Tuma F. The use of educational technology for interactive teaching in lectures. Ann Med Surg. 2021;62:231–5. https://doi.org/10.1016/j.amsu.2021.01.051

47. Koster AS, Vermunt JD. Longitudinal changes of deep and surface learning in a constructivist pharmacy curriculum. Pharmacy. 2020;8(4):200. https://doi.org/10.3390/pharmacy8040200

48. Andrade HL, Brookhart SM, Yu EC. Classroom assessment as co-regulated learning: a systematic review. Front Educ. 2021;6:751168. https://doi.org/10.3389/feduc.2021.751168

49. Chang C-L. Effect of immersive virtual reality on post-baccalaureate nursing students' in-dwelling urinary catheter skill and learning satisfaction. Healthcare. 2022;10(8):1473. https://doi.org/10.3390/healthcare10081473

50. Issleib M, Kromer A, Pinnschmidt HO, Süss-Havemann C, Kubitz JC. Virtual reality as a teaching method for resuscitation training in undergraduate first-year medical students: a randomized controlled trial. Scand J Trauma Resusc Emerg Med. 2021;29:27. https://doi.org/10.1186/s13049-021-00836-y

51. Lo K-W, Yang B-H. Development and learning efficacy of a simulation rubric in childhood pneumonia for nursing students: a mixed methods study. Nurse Educ Today. 2022;119:105544. https://doi.org/10.1016/j.nedt.2022.105544

52. Rouleau G, Pelletier J, Côté J, Gagnon M-P, Martel-Laferrière V, Lévesque R, et al. Codeveloping a virtual patient simulation to foster nurses' relational skills consistent with motivational interviewing: a situation of antiretroviral therapy nonadherence. J Med Internet Res. 2020;22(7):e18225. https://doi.org/10.2196/18225

53. Lee E, Mabry JL, Roberts T, Davis KM. Development of the simulation-positive learning evaluation (SIM-PLE) rubric for use in medical-surgical nursing high fidelity simulations. Nurse Educ Pract. 2021;51:102992. https://doi.org/10.1016/j.nepr.2021.102992

54. Ma C-W, Cheng P-S, Chan Y-S, Tipoe GL. Virtual reality: a technology to promote active learning of physiology for students across multiple disciplines. Adv Physiol Educ. 2023;47(3):594–603. https://doi.org/10.1152/advan.00172.2022

55. Karaduman GS, Basak T. Is virtual patient simulation superior to human patient simulation: a randomized controlled study. Comput Inform Nurs. 2023;41(6):467–76. https://doi.org/10.1097/CIN.0000000000000957

56. Helle N, Vikman MD, Dahl-Michelsen T, Lie SS. Health care and social work students' experiences with a virtual reality simulation learning activity: qualitative study. JMIR Med Educ. 2023;9:e49372. https://doi.org/10.2196/49372

57. Wong W-K, Chen K-P, Chang H-M. A comparison of a virtual lab and a microcomputer-based lab for scientific modelling by college students. J Balt Sci Educ. 2020;19(1):157–73. https://doi.org/10.33225/jbse/20.19.157

58. Makarova I, Mustafina J, Boyko A, Fatikhova L, Parsin G, Buyvol P, et al. A virtual reality lab for automotive service specialists: a knowledge transfer system in the digital age. Information. 2023;14(3):163. https://doi.org/10.3390/info14030163

59. De Lorenzis F, Visconti A, Restivo S, Mazzini F, Esposito S, Garofalo SF, et al. Combining virtual reality with asymmetric collaborative learning: a case study in chemistry education. Smart Learn Environ. 2024;11(1):43. https://doi.org/10.1186/s40561-024-00331-8

60. Jallad ST, Natsheh I, Helo LA, Ibdah DM, Salah A, Muhsen R, et al. Nursing student's perceptions, satisfaction, and knowledge toward utilizing immersive virtual reality application in human anatomy course: quasi-experimental. BMC Nurs. 2024;23(1):601. https://doi.org/10.1186/s12912-024-02254-8

61. Tokatlidis C, Rapti S, Tselegkaridis S, Sapounidis T, Papakostas D. Virtual environment in engineering education: the role of guidance, knowledge and skills development in electronic circuits teaching. Educ Sci. 2024;14(12):1336. https://doi.org/10.3390/educsci14121336

62. Chircop A, Cobbett S, Schofield RE, Boudreau C, Egert A, Filice S, Harvey A, et al. Multi-jurisdictional evaluation of Sentinel City virtual simulation for community health nursing clinical education. QANE-AFI. 2022;8(4):2. https://doi.org/10.17483/2368-6669.1352

63. Kolil VK, Achuthan K. Virtual labs in chemistry education: a novel approach for increasing student's laboratory educational consciousness and skills. Educ Inf Technol. 2024;29(18):25307–31. https://doi.org/10.1007/s10639-024-12858-x

64. Barnett SG, Porter AL, Allen SM, Bastianelli KMS, Chen JS, Kachlic MD, et al. Expert consensus to finalize a universal evaluator rubric to assess pharmacy students' patient communication skills. Am J Pharm Educ. 2020;84(12):848016. https://doi.org/10.5688/ajpe848016

65. Garling KA, Wong B. An initial reliability analysis of a patient counselling rubric to objectively measure student pharmacist performance. Heliyon. 2023;9(5):e15768. https://doi.org/10.1016/j.heliyon.2023.e15768

66. Brookhart SM. How to create and use rubrics for formative assessment and grading. ASCD; 2013.

67. Fernández-Gómez E, Martín-Salvador A, Luque-Vara T, Sánchez-Ojeda MA, Navarro-Prado S, Enrique-Mirón C. Content validation through expert judgement of an instrument on the nutritional knowledge, beliefs, and habits of pregnant women. Nutrients. 2020;12(4):1136. https://doi.org/10.3390/nu12041136

68. Zamanzadeh V, Ghahramanian A, Rassouli M, Abbaszadeh A, Alavi-Majd H, Nikanfar A-R. Design and implementation content validity study: development of an instrument for measuring patient-centered communication. J Caring Sci. 2015;4(2):165–78. https://doi.org/10.15171/jcs.2015.017

69. In J. Introduction of a pilot study. Korean J Anesthesiol. 2017;70(6):601–5. https://doi.org/10.4097/kjae.2017.70.6.601

70. Prasad GNR. Evaluating student performance based on bloom's taxonomy levels. J Phys: Conf Ser. 2021;1797(1):012063. https://doi.org/10.1088/1742-6596/1797/1/012063

71. Brentani E, Golia S. Unidimensionality in the Rasch model: how to detect and interpret. Statistica. 2007;67(3):253–61. https://doi.org/10.6092/issn.1973-2201/3508

72. Ishak AH, Osman MR, Mahaiyadin MH, Tumiran MA, Anas N. Examining unidimensionality of psychometric properties via Rasch model. IJCIET. 2018;9(9):1462–7.

73. Yeoh TT, Foo KM, Koh SK, Quek K, Yew KW, Fenwick E, et al. Development of a pharmacy educational environment questionnaire. Curr Pharm Teach Learn. 2023;15(12):993–1005. https://doi.org/10.1016/j.cptl.2023.10.010

74. Smith AB, Rush R, Fallowfield LJ, Velikova G, Sharpe M. Rasch fit statistics and sample size considerations for polytomous data. BMC Med Res Methodol. 2008;8:33. https://doi.org/10.1186/1471-2288-8-33

75. Ravand H, Firoozi T. Examining construct validity of the master's UEE using the Rasch model and the six aspects of the Messick's framework. Int J Lang Test. 2016;6(1):1–23.

76. Arku D, Almatruk Z, Warholak T, Axon DR. Evaluating the reliability and validity of a questionnaire used to measure experiences of teamwork among student pharmacists in a quality improvement course. Curr Pharm Teach Learn. 2022;14(5):552–60. https://doi.org/10.1016/j.cptl.2022.04.015

77. Fisher WP. Rating scale instrument quality criteria. RMT. 2007;21(1):1095.

78. Newsom LC, Augustine J, Momary K. Development of a script concordance test to assess clinical reasoning in a pharmacy curriculum. Curr Pharm Teach Learn. 2022;14(9):1135–42. https://doi.org/10.1016/j.cptl.2022.07.028

79. Boone WJ. Rasch analysis for instrument development: why, when, and how? CBE—Life Sci Educ. 2016;15(4):1–7. https://doi.org/10.1187/cbe.16-04-0148

80. Kean J, Brodke DS, Biber J, Gross P. An introduction to item response theory and Rasch analysis: application using the eating assessment tool (EAT-10). Brain Impair. 2018;19(Spec Iss 1):91–102. https://doi.org/10.1017/brimp.2017.31

81. Yang Y, Peng Y, Li W, Lu S, Wang C, Chen S, et al. Psychometric evaluation of the academic involution scale for college students in China: an application of Rasch analysis. Front Psychol. 2023;14:1135658. https://doi.org/10.3389/fpsyg.2023.1135658

82. Goswami S, Barnard M, Bynum LA, Thompson S, Kang M. Psychometric evaluation of the Health Care Provider HIV/AIDS Stigma Scale (HPASS) using Rasch analysis. Curr Pharm Teach Learn. 2022;14(9):1143–53. https://doi.org/10.1016/j.cptl.2022.07.035

83. Saseen JJ, Ripley TL, Bondi D, Burke JM, Cohen LJ, McBane S, et al. ACCP clinical pharmacist competencies. Pharmacotherapy. 2017;37(5):630–6. https://doi.org/10.1002/phar.1923

84. Kumar S, Bhoi P, Gharat MS, Mohanta GP. Developing a competency assessment framework for pharmacists in primary health care settings in India. PLoS One. 2025;20(3):e0316646. https://doi.org/10.1371/journal.pone.0316646

85. Potgieter ML, Filmalter C, Maree C. Teaching, learning and assessment of the affective domain of undergraduate students: a scoping review. Nurse Educ Pract. 2025;86:104417. https://doi.org/10.1016/j.nepr.2025.104417

86. Galdonez DP. Exploring the experiences of STEM students in writing research: insights, challenges, and strategies. J Basic Educ Res. 2025;6(1):9–16. https://doi.org/10.37251/jber.v6i1.1150