MALAYSIAN JOURNAL OF LEARNING AND INSTRUCTION
https://e-journal.uum.edu.my/index.php/mjli

How to cite this article:
Shahbani, H. I. M. & Rahim, F. A. (2026). Expert insights into criteria for STEM-based authentic assessment in chemistry. *Malaysian Journal of Learning and Instruction*, *23*(2), 1-20. https://doi.org/10.32890/mjli2026.23.2.1

EXPERT INSIGHTS INTO CRITERIA FOR STEM-BASED AUTHENTIC ASSESSMENT IN CHEMISTRY

[1]Husna Izzati Muhamad Shahbani & Fauziah Abdul Rahim
School of Education, Universiti Utara Malaysia, Malaysia

[1]Corresponding author: husna.izzati@ymail.com

Received: 18/9/2025 Revised: 28/1/2026 Accepted: 15/6/2026 Published: 31/7/2026

ABSTRACT

Purpose – Although STEM education and authentic assessment have been widely researched, the criteria for authentic assessments in Chemistry, particularly at the Malaysian pre-university level, remain underexplored. This study explores experts' insights into the criteria for STEM-based authentic assessments in Chemistry that support student learning in the Malaysian pre-university context.

Methodology – A qualitative approach was employed, and a critical review of 50 journal articles was conducted to identify preliminary constructs for developing the criteria for STEM-based authentic assessments in Chemistry. Interviews with nine experts specialising in Chemistry, STEM, and assessment provided the data which were subsequently analysed thematically using Atlas.ti 9 software to establish the criteria for STEM-based authentic assessments in Chemistry.

Findings – Data from the experts revealed that the criteria of STEM-based authentic assessments in Chemistry include real-world applications, interdisciplinary learning, holistic skills development, the use of rubrics, and active learning processes. The findings highlight the importance of aligning assessments with learning objectives that cover cognitive, affective, and psychomotor domains through innovative approaches such as project-oriented problem-based learning and the integration of digital tools.

Novelty – This study contributes to establishing the criteria for STEM-based authentic assessments in Chemistry at the Malaysian pre-university level through expert interviews and critical literature analysis, integrating holistic skills development, interdisciplinary learning, and real-world applications into assessment design.

Significance – This study provides guidelines for educators, curriculum designers, and assessment practitioners in designing STEM-based authentic assessments in Chemistry, equipping students with real-world skills, and aligning national education policies with 21st-century skill demands.

Keywords: authentic assessment criteria, Chemistry education, expert insights, STEM, pre-university, real-world applications

INTRODUCTION

Rapid global developments in science, technology, and education have brought Science, Technology, Engineering, and Mathematics (STEM) to the forefront of national policy frameworks. The primary focus of STEM is to equip students with 21st-century competencies such as problem-solving, creativity, collaboration, and interdisciplinary thinking (Martín-Páez et al., 2019). In response, the Malaysia Education Blueprint 2013–2025 outlines national policy initiatives to integrate STEM approaches, particularly within the science disciplines (Ministry of Education Malaysia, 2013). This commitment is further reinforced in the current Malaysia Higher Education Plan (RPTM) 2026–2035 (Ministry of Higher Education Malaysia, 2026). Although Chemistry has the potential to foster both conceptual understanding and applied scientific reasoning, current assessment practices in Malaysian pre-university settings often emphasise memorisation and structured responses, which do not align with the applied skills promoted in STEM education (Nawawi et al., 2021). As a result, students' ability to integrate knowledge across disciplines and apply concepts to real-world or unfamiliar contexts has not been adequately assessed (Adebusuyi et al., 2022).Consequently, scholars and practitioners have sought to address these shortcomings by exploring authentic assessments that evaluate students' ability to apply knowledge in real-world contexts (Al Sabbah et al., 2022).

Rooted in constructivist theory, authentic assessment encourages students to develop higher-order thinking skills across the cognitive, affective, and psychomotor domains (Villarroel et al., 2018). In Chemistry, this can be achieved through tasks that closely reflect scientific practices, such as conducting experiments, engaging in project-based inquiry, solving interdisciplinary problems, and interpreting real-world data (Giac et al., 2024). Situated within STEM principles, authentic assessment not only fosters students' interdisciplinary reasoning and real-world problem-solving abilities but also prepares them to address complex global challenges (Alt & Raichel, 2022).

In Chemistry, Kriswantoro et al. (2021) and Nahadi et al. (2021) developed authentic assessment instruments and tools covering the cognitive, affective, and psychomotor domains with expert involvement at the secondary school level. In contrast, Wan Yusof et al. (2022) introduced a project-based authentic assessment specifically for pre-university Chemistry. However, their study did not incorporate expert-informed criteria for the project-based approach. Moreover, the assessment focused only on the cognitive and psychomotor domains, overlooking the affective domain. At the tertiary level, Wahyuni et al. (2022) developed authentic assessment rubrics for a research methods course with expert consultation, while Syafdaningsih and Suningsih (2020) developed authentic assessment instruments for early childhood education. Both studies assessed only the cognitive domain. Collectively, these studies did not incorporate STEM integration into their authentic assessment practices. Earlier, however, Widana et al. (2021) developed a validated authentic assessment for project-based tasks that integrated a STEM approach. Nevertheless, the assessment was designed for secondary level Mathematics students and therefore may not be fully applicable to the pre university context.

Although previous research has proposed teaching approaches and assessment instruments for authentic assessment across different educational levels, most studies have not thoroughly examined the criteria required to design STEM-based authentic assessments in Chemistry, particularly for the pre-university context. Existing studies have often focused on task design or student achievement rather than providing expert-informed criteria to guide instructors in designing, implementing, and evaluating the STEM-integrated authentic assessment tasks in Chemistry effectively (Shidiq et al., 2020). This gap has resulted in inconsistent implementation and has limited the alignment of assessment practices with national curriculum objectives and 21st-century educational goals (Gao et al., 2020). Hence, this study contributes new insights by developing an expert-informed set of proposed criteria based on literature analysis and expert perspectives for STEM-based authentic assessments in Chemistry, with a particular focus on the Malaysian pre-university context. The study extends the principles of authentic assessment (Wiggins, 1998) by integrating them with the STEM framework (Kelley & Knowles, 2016). Accordingly, the objective of this study is to investigate experts' perspectives on STEM-based authentic assessments in Chemistry for Malaysian pre-university students. Adopting a qualitative research design, the present study explores the theoretical and practical elements that experts consider essential for STEM-based authentic assessment in Chemistry.

LITERATURE REVIEW

Conventional assessments methods may be insufficient to prepare students for the demands of the modern age (Shafiyeva, 2021) because they primarily focus on knowledge recall, limiting the assessment of students' ability to apply knowledge in real-world contexts. In Malaysia, the limited duration of pre-university education, which typically lasts up to two years (StudyMalaysia, 2017), often results in assessments that prioritise knowledge acquisition and rapid content delivery over critical thinking and problem-solving skills (Al Sabbah et al., 2022). These conventional assessment methods also fail to capture students' understanding of the abstract nature of Chemistry concepts (Sırakaya & Alsancak Sırakaya, 2022) creating a disconnect between classroom learning and the practical skills required for future careers in Chemistry. Authentic assessments seek to address this issue by engaging students in tasks that mirror real-world situations, enabling them to apply both their knowledge and practical skills (Chang et al., 2024; Nahadi et al., 2021; Wiggins, 1998). This approach integrates the cognitive, affective, and psychomotor domains, enabling students to demonstrate their ability to analyse problems and perform practical tasks effectively (Asante, 2023). Authentic contexts also help connect cognitive learning with societal realities (Hussin et al., 2021). The effective implementation of authentic assessment depends on the availability of adequate resources and facilities across institutions (Aziz et al., 2020). In the context of Chemistry, authentic assessment involves students collaborating to solve chemical problems by applying Chemistry knowledge and laboratory skills while developing the competencies needed to address complex challenges beyond the classroom (Sutadji et al., 2021). In addition, authentic assessment tasks promote creativity, critical thinking (Ijirana et al., 2022), and independent and lifelong learning skills (Alt & Raichel, 2022; Li, 2023). While repeated observations are required to assess affective domain (Shankar et al., 2023), scientific communication and hands-on activities are commonly used to evaluate the psychomotor domain (Kallarackal, 2023). In practice, rubrics are essential for ensuring clear and transparent assessment across different evaluation criteria, particularly in STEM-integrated assessment tasks (Fadillah & Ha, 2024; Hadibarata et al., 2024). Moreover, student performance can be enhanced through timely, practical, and constructive feedback (Johnson et al., 2020; Ryan et al., 2024). Accordingly, extensive research has explored various authentic assessment across disciplines at the primary, secondary, and tertiary levels (Karim et al., 2018; Kriswantoro et al., 2021; Nahadi et al., 2021; Syafdaningsih & Suningsih, 2020; Wahyuni et al., 2022; Widana et al., 2021) but there remains limited guidance on its implementation in pre-university Chemistry (Wan Yusof et al., 2022).

The growing trend of incorporating STEM into Chemistry education demands a transformation in assessment methods that reflect interdisciplinary learning (Thibaut et al., 2018). The integration of STEM approaches into authentic assessment has gained significant attention since the early 1990s, leading researchers and educators to develop innovative teaching methods (Du et al., 2018; Reynders et al., 2020). This evolution has given rise to Integrated STEM Education, a multidisciplinary framework that promotes holistic thinking and real-world application by merging STEM concepts with other disciplines, particularly in developing countries (Kelley & Knowles, 2016). The strong emphasis on technology in STEM education, such as the use of virtual labs (Alhashem & Alfailakawi, 2023) can further enhance student engagement (Williams et al., 2022). In Malaysia, the Ministry of Education has introduced the National STEM Action Plan to cultivate STEM awareness, stimulate research initiatives, and expand career opportunities through collaboration between the government and industry (Ministry of Education Malaysia, 2013). However, research on the use of integrated STEM in assessments often focuses more on teaching approaches or assessment outcomes rather than on the practical criteria of authentic assessment (Adebusuyi et al., 2022; Gao et al., 2020; Ijirana et al., 2022; Nadelson & Seifert, 2017; Özkul & Özden, 2020; Roberts et al., 2022) which serve as crucial guidelines for improving institutional delivery and assessment practices. Although some studies (Friench et al., 2023; Widana et al., 2021) have explored authentic assessment within the STEM approach, their findings are not always directly applicable to pre-university settings.

A strong theoretical foundation for designing STEM-based authentic assessments in Chemistry should emphasise real-world application of knowledge. Lev Vygotsky’s Social Constructivism highlights knowledge construction through social and cultural interactions, where learners enhance their understanding by engaging with others. This process is facilitated by scaffolding, which helps learners progress from their current level of ability towards their full potential (Vygotsky, 1978). This perspective aligns well with STEM-based authentic assessment, which promotes socially mediated learning in solving real-world interdisciplinary problems while developing higher mental functions. In other words, there is a need for assessments that enable learners to collaborate and communicate while solving problems to deepen their understanding. Situated Cognition Theory further complements this study by arguing that knowledge is closely tied to context; therefore, learning becomes more effective when tasks reflect professional or everyday practices (Brown et al., 1989). This perspective also corresponds well with authentic assessments in Chemistry, which aim to evaluate students’ ability to apply knowledge in realistic contexts rather than in simulated settings.

The majority of previous research on the approaches and outcomes of authentic assessment has focused on secondary and tertiary-level Chemistry, or on non-Chemistry contexts. Nevertheless, studies on integrating STEM into authentic assessments for pre-university Chemistry students remain underexplored (Nawawi et al., 2021). This significant gap presents a challenge for educators due to the absence of criteria informed by expert perspectives for authentic assessments that meaningfully integrate STEM, leading to inconsistent implementation and limited practical guidance (Gao et al., 2020; Shidiq et al., 2020). To address this issue, the Integrated STEM Education conceptual framework (Kelley & Knowles, 2016), which conceptualises learning as interdisciplinary appears suitable for establishing criteria that emphasise authentic tasks based on Grant Wiggins' authentic assessment principles (Wiggins, 1998). This study contributes to the academic field by proposing a refined set of expert-informed criteria for STEM-based authentic assessments in Chemistry within the Malaysian pre-university context, providing practical guidelines for educators in designing authentic assessments. Ultimately, this study seeks to answer two primary research questions: (1) What are the criteria for STEM-based authentic assessments in Chemistry for the Malaysian pre-university context? (2) How do experts in Chemistry, STEM, and assessment contribute to the understanding of these criteria within the Malaysian pre-university contexts?

METHODOLOGY

Research Design

This qualitative study adopts an interpretive design, positioning the researcher as the primary interpreter of experts' insights on the criteria of STEM-based authentic assessments in Chemistry, specifically tailored to the Malaysian pre-university context (Creswell & Poth, 2018). Rooted in the interpretivist paradigm, this approach emphasises the co-construction of knowledge and its contextual understanding through the interaction and interpretation between researchers and participants regarding effective practices (Bozkurt & Ozturk, 2022). The study comprised two phases. The first phase involved a critical analysis of related literature to develop predetermined categories relevant to the construction of the interview protocol, prior to obtaining empirical data in the second phase (Braun & Clarke, 2012).

Sampling

Purposive sampling was used to recruit experts in Chemistry, STEM education, and authentic assessment, with a strong preference for individuals possessing substantial practical experience in curriculum and assessment. Drawing on Malcolm Gladwell's (2008) notion of expertise, participants with a minimum of ten years teaching, training, or assessment-related experience in their respective fields were selected. This sampling strategy involved recruiting Chemistry, STEM, and assessment experts with at least a master's degree in their respective fields, who had played a role in designing, implementing, training, or evaluating assessment practices at the pre-university level, as shown in Table 1.

Table 1

Distribution of Expert Participants' Profiles by Discipline

Expert Code	Expert Category	Highest Academic Qualification	PhD Specialisation	Years of Professional Experience
C1	Chemistry	PhD	Chemistry	18
C2	Chemistry STEM	PhD	Physics Chemistry	14
C3	Chemistry STEM	PhD	Polymer Chemistry	10
C4	Chemistry	PhD	Chemistry	11

(continued)

Expert Code	Expert Category	Highest Academic Qualification	PhD Specialisation	Years of Professional Experience
S1	STEM	PhD	Physics	12
S2	STEM	PhD	Biosciences	11
A1	Assessment	PhD	Curriculum and Pedagogy	30
A2	Assessment	Master's degree	Chemistry	30
A3	Assessment	Master's degree	Chemistry	30

The Chemistry experts were selected for their advanced knowledge of chemical principles and practical applications, supported by extensive educational and research experience. Building on this expertise, STEM experts were chosen for their ability to integrate science, technology, engineering, and mathematics, as demonstrated through the application of STEM principles in successful project research. Complementing these perspectives, assessment experts were included for their expertise in designing reliable and valid assessments that effectively measure performance, competencies, and learning outcomes within educational settings. In total, nine experts were recruited for interviews to examine the criteria necessary for developing STEM-based authentic assessments in Chemistry within the Malaysian pre-university context, as shown in Table 1. Accordingly, nine interviews were conducted in this study, as previous research suggests that the optimal number of qualitative interviews required to achieve data saturation typically ranges between six and twelve (Guest et al., 2006).

Ethical Considerations

Participants' confidentiality was prioritised by ensuring that all responses and research-related information were kept strictly confidential. The data were securely stored and pseudonyms were used to protect participant anonymity and minimise the risk of potentially harmful consequences arising from unintended disclosure (McNay, 2007). C1, C2, C3, and C4 represented Chemistry experts, with C1 and C4 specialising in core chemical knowledge, while C2 and C3 integrated Chemistry with STEM for interdisciplinary problem-solving. Similarly, S1 and S2 represented STEM experts who applied interdisciplinary approaches in their professional practices, contributing to advancements in STEM fields. Meanwhile, A1, A2, and A3 were assessment specialists who designed and evaluated tools to effectively measure learning outcomes and performance. All participants were informed about the purpose of the study and written consent was obtained prior to the data collection process. Interviews were scheduled at participants' convenience. They were assured that access to interview materials would be restricted to the researchers and supervisors only, and were also given the option to withdraw from the study or request modifications to their interview content if necessary. All interviews were conducted in private settings, and the original recordings, notes, and transcripts were securely stored by the researchers.

Data Collection

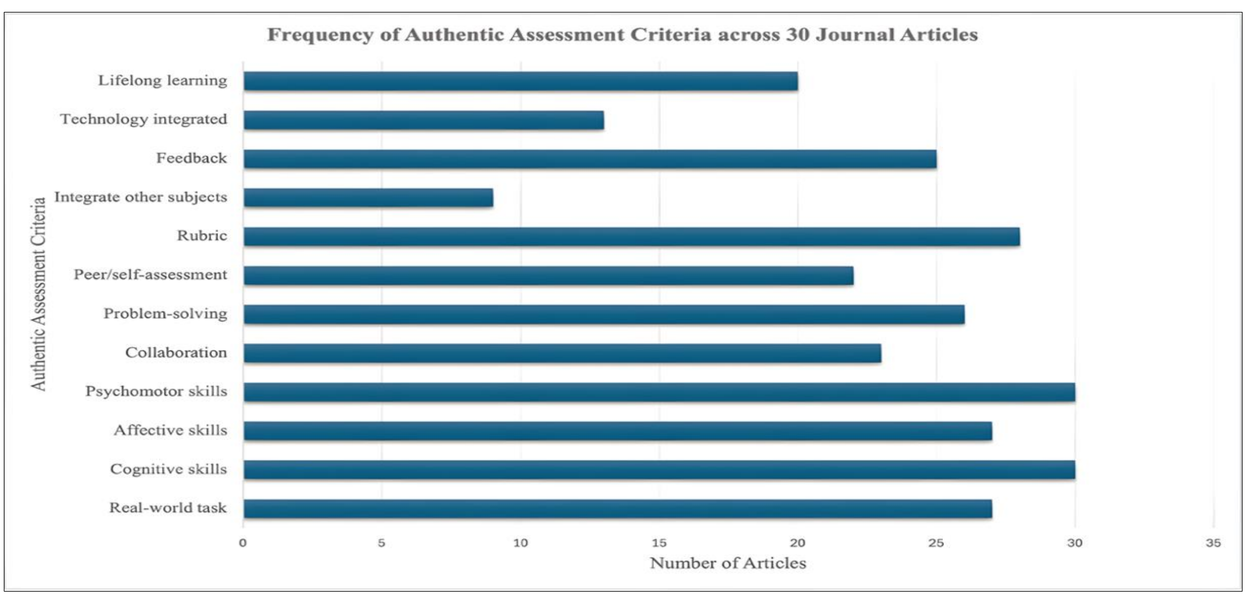
In Phase One, a critical analysis of 50 journal articles on authentic assessment across primary, secondary, and tertiary education published within the last eight years, encompassing both STEM and non-STEM subjects, was conducted to identify relevant criteria. The literature search used the keywords "authentic assessment", "STEM assessment", "STEM Chemistry", "pre-university Chemistry", and "interdisciplinary STEM" across ERIC, Scopus, ScienceDirect, Web of Science and Google Scholar databases. Articles were excluded if, upon review, they were not directly related to authentic assessment practices, STEM, Chemistry relevance, or real-world contexts. This process resulted in 30 of the most relevant studies being selected, as they informed the construction of predetermined categories for STEM-based authentic assessment in Chemistry. The predetermined categories identified in this phase included real-world tasks, cognitive skills, affective skills, psychomotor skills, collaboration, problem-solving, peer assessment, self-assessment, rubrics, interdisciplinary integration, feedback, technology integration, and lifelong learning. These categories were subsequently used to organise the semi-structured interview protocol for Phase Two. The aim of Phase Two was to obtain in-depth perspectives from experts in Chemistry, STEM, and assessment, as well as to refine the criteria through content analysis. A total of four experts participated in semi-structured face-to-face interviews, while five participated via email correspondence based on their personal preference. This asynchronous method allowed participants to respond at their convenience, often resulting in more focused and information-rich responses compared to face-to-face interviews (Braun et al., 2021). The face-to-face interviews lasted approximately 45 to 60 minutes and were conducted in English. The email-based participants were provided with the same semi-structured interview protocol and given a week to complete and return their responses to the researchers.

Data Analysis

Phase One involved a content analysis of the 30 retained articles, specifically using conceptual analysis to identify explicit terms related to authentic assessment in STEM. This study focused only on explicit terms stated in the text of the selected articles, rather than implicit meanings that require researcher interpretation, in order to minimise subjectivity (Carley, 1990). The analysis involved examining the frequency of specific terms related to STEM, Chemistry, and assessment across the past eight years. The results are presented in Figure 1, which highlights the criteria most prominently featured in the literature.

Figure 1

Frequency of Authentic Assessment Criteria across 30 Journal Articles



The text was coded into manageable content categories, which formed the criteria derived from Phase One of this study. This analysis informed the development of semi-structured interview questions to ensure alignment with established research. The criteria identified in Phase One were used as predetermined categories for constructing the semi-structured interview protocol in Phase Two. By referring to these predetermined categories, consistency in the questions posed to each participant was ensured, thereby, enhancing the reliability of the data. The interview questions focused on expert perspectives regarding the relevance and applicability of these criteria in STEM-based authentic assessments in Chemistry. These criteria also guided the thematic analysis, allowing responses to be coded within the established framework and subsequently refined based on expert feedback.

The data from interviews and email correspondence in Phase Two were analysed using thematic analysis, following the six-phase model established by Braun and Clarke (2012), which includes familiarisation with the data through reading of transcripts, generating initial codes across the dataset, searching for themes by grouping related codes, reviewing themes to ensure clear distinctions, defining and naming themes, and producing the final report. Computer Assisted Qualitative Data Analysis Software (CAQDAS), specifically Atlas.ti 9, was used to support the thematic analysis process by systematically identifying, organising, and analysing patterns and themes within the qualitative data, thereby providing deeper insight into concepts and relationships relevant to STEM-based authentic assessment. Emergent categories were derived from the expert responses, revealing new insights that were not initially anticipated within the predetermined categories developed in Phase One. Data collected from both face-to-face interviews and email-based responses were analysed together as complementary qualitative sources using the same coding framework. The integration of both predetermined and emergent categories enabled a more comprehensive understanding of the data.

Acknowledging the interpretive nature of this study, reflexivity was maintained throughout data collection and analysis process to address potential bias. The analysis combined predetermined categories derived from the literature in Phase One with emergent insights from experts in Phase Two, ensuring alignment between theoretical foundations and participant perspectives. Several strategies were employed to enhance trustworthiness, including triangulation and consistency. Triangulation was achieved by comparing the data across both phases, as well as examining responses from Chemistry, STEM, and assessment experts to ensure that the final criteria were grounded in theory and supported by empirical evidence (Denzin, 1989). To ensure the consistency, the same semi-structured interview protocol was used with all experts (Shenton, 2004). In addition, a structured coding procedure was applied using Atlas.ti 9, and the interview protocol was piloted in advance to assess its clarity and relevance (van Teijlingen & Hundley, 2002).

RESULTS AND DISCUSSION

The analysis of data from both phases resulted in 72 distinct codes. In Phase One, authentic assessment criteria were developed through the literature review, confirming several predetermined categories, including real-world tasks, cognitive skills, affective skills, psychomotor skills, collaboration, problem-solving, peer assessments, self-assessments, rubrics, interdisciplinary integration, feedback, technology integration, and lifelong learning. The analysis of interview data in Phase Two further revealed several emergent categories, namely project-based assessment, ethical and societal driven considerations, and interpersonal skills. The findings identified five main themes and 15 sub-themes that outline the criteria for STEM-based authentic assessment in Chemistry, as presented in Table 2.

Table 2

Criteria of STEM-based Authentic Assessment in Chemistry Identified by Experts

Theme	Sub Theme	Description
Real-World Applications	Tasks integrating real-world applications	Tasks that enable students to apply knowledge to solve real-world problems
	Project-based assessment	Tasks involving projects that address real-world issues
	Ethical and societal driven	Assessment that emphasises ethical reasoning and encourages responsibility towards the community
STEM Integration	Knowledge and skills	Activities that apply scientific inquiry, mathematical analysis, engineering design, and technological tools to solve complex problems
	Digital technologies	The use of digital tools to evaluate students' understanding and application of Chemistry concepts
Holistic Skills	Cognitive skills	Evaluates intellectual abilities such as understanding, applying, analysing, synthesising, and evaluating Chemistry concepts
	Affective skills	Measures attitudes, interests, and values related to personal and societal themes.
	Psychomotor skills	Evaluates students' ability to perform hands-on tasks such as experiments or laboratory techniques
Rubrics	Accommodating various targeted skills	Assesses different skills to cater to specific student abilities
	Clear criteria	Specific, fair, and consistent criteria for evaluation
	Alignment with learning objectives	Rubrics that measure intended learning outcomes
Learning Process	Collaboration	Evaluates teamwork, communication, and contribution within groups
	Interpersonal skills	Measures the ability to interact effectively and respectfully with others
	Independent learning	Assesses the ability to solve problems independently
	Effective and constructive feedback	Immediate, specific, and practical feedback to guide student improvement

Findings across the themes revealed that real-world applications and STEM integration formed the core of STEM-based authentic assessment practices in Chemistry, whereas rubrics, learning processes, and holistic skills functioned as facilitating elements for effective implementation.

Real-world Applications

The theme of real-world applications comprises the sub-themes of tasks integrating real-world applications, project-based assessments, and ethical and societal driven projects. This theme emphasises the need for assessments to integrate fundamental Chemistry concepts with real-world applications. Multiple experts highlighted that authentic assessments should reflect real-world challenges and focus on practical skills, although their perspectives differed in how authenticity should be operationalised within assessment tasks. To illustrate this, C4 and S2 emphasised the importance of focusing on real-life scenarios, stating that "*authentic assessments should move beyond rote memorisation or standardised testing.*" Similarly, A1 and A2 stressed the importance of assessments that "*connect classroom learning with real-world professional challenges.*" These perspectives align with the current study, which highlights authentic tasks as crucial for fostering meaningful learning rooted in real-world contexts (Nahadi et al., 2021).

Authentic contexts that integrate real world scenarios in assessment can be implemented through project-based tasks (Wan Yusof et al., 2022). As suggested by C3, such assessments should "*foster higher-order thinking and problem-solving skills*" by enabling students to apply scientific knowledge "*to solve problems related to daily lives.*" Similarly, C1 also provided an example of a project connecting Chemistry concepts to practical applications, such as the use of photochemistry in daily scenarios which research has identified as project-based tasks with authentic contexts that enhance student engagement (Chang et al., 2024). In this regard, A1 also revealed that effective pedagogy for authentic assessment may consist of problem-based learning, project-based learning, or a combination of both integrated as project-oriented problem-based learning, which has been recognised as an effective pedagogical approach with established methods. These approaches may be considered authentic assessments, as they enable instructors to measure 21st-century skills through such pedagogies.

The importance of designing assessments centred on complex, open-ended problems was also emphasised by the participants. C4 believed that these types of assessments "*enhance creativity and critical thinking through project-based learning,*" while C2 highlighted "*their role in stimulating higher-order cognitive abilities."* This perspective is supported by evidence showing that project-based authentic assessment nurtures critical thinking, creativity, and fosters deeper learning (Kriswantoro et al., 2021). The effectiveness of these instructional approaches in supporting the acquisition of both theoretical knowledge and hands-on skills depend on adequate instructor support, sufficient resources, and appropriate scaffolding (Widana et al., 2021).

The participants also underscored the importance of incorporating social and ethical considerations into authentic STEM-based Chemistry assessments. S2 noted that students are able to understand how their "*learning is relevant to society and develop a sense of responsibility when they are involved in community-based projects.*" To prepare students for real-world challenges whilst encouraging them to recognise the broader societal role of Chemistry, A1 further stressed that "*ethical dilemmas are common in practical settings*" and recommended assessments that prioritise ethical reasoning and responsibility, which is consistent with situated the Situated Cognition theory (Brown et al., 1989). Similarly, A2 highlighted the *"integration of values such as respect and moral behaviour to guide students to analyse their decisions and responsibilities as responsible citizens,*" suggesting assessments that involve community service. Such practices can stimulate a stronger sense of responsibility among students by exposing them to the realities of community life and helping them understand the consequences of their actions (Hussin et al., 2021). However, the impact of these approaches may vary depending on institutional support and student motivation.

STEM Integration

STEM integration is a major theme comprising sub-themes of knowledge and skills, as well as digital technologies. This theme highlights the importance of integrating Chemistry with scientific inquiry, mathematical analysis, engineering design, and technological tools to address real-world problems. Differences in perspectives on STEM integration emerged among the experts, with some emphasising disciplinary connections involving Chemistry, while others highlighted design-based and problem-solving approaches. S2 and C4 proposed that tasks should "*integrate concepts from disciplines beyond Chemistry*". The experts further stated that interdisciplinary connections *"can be incorporated by designing assessments that require the application of Chemistry concepts to solve problems in other STEM fields,*" such as linking Chemistry content with engineering, technology and mathematics. A1 shared an example of a project in which students from diverse STEM fields collaborated to design a robot for medical purposes, demonstrating how Chemistry intersects with engineering, technology, and medicine. A2 also supported this perspective by recommending *the "integration of mathematical calculations, engineering frameworks, and technological tools for data analysis and presentation"* to demonstrate the relevance of STEM in Chemistry projects situated in real-world contexts. These suggestions support the view that STEM-based authentic assessments in Chemistry should emphasise interdisciplinary problem-solving rather than focusing solely on Chemistry-based problem-solving (Kelley & Knowles, 2016; Shidiq et al., 2020). Although STEM interdisciplinary practices are increasingly recognised, their effective implementation requires careful curriculum design to ensure meaningful integration of subjects while preserving essential Chemistry competencies.

All experts agreed on the importance of incorporating technology into Chemistry assessments to meet the needs of 21st-century learners. C1 emphasised the use of digital tools for problem analysis and concept application, while S2 and A3 highlighted the role of "*industry-relevant software and molecular modelling to enhance understanding.*" Virtual platforms such as PhET simulations and virtual labs were also noted by C4 and A1 as cost-effective tools that build students' confidence for physical lab work. This view is supported by the literature, with Alhashem and Alfailakawi (2023) demonstrating that virtual labs significantly enhance student learning and engagement. Meanwhile, C3 and A2 recommended interactive tools such as Quizizz and Mentimeter to increase student engagement, aligning with Williams et al. (2022), who found that visual reality promotes positive learning experiences. Nevertheless, dependence on digital platforms may limit accessibility for some students and AI tools such as ChatGPT were recognised by A1 as raising regarding plagiarism. A1 further suggested the careful design of written assignments to minimise plagiarism risks, which is consistent with the concerns highlighted by Sysoyev (2024). Hence, as identified in this study, authentic assessment demands careful attention to both equity and academic integrity.

Holistic Skills

The broader theme of holistic skills comprises three distinct sub-themes: cognitive skills, affective skills, and psychomotor skills. This theme highlights the need for authentic assessments to evaluate students' cognitive, affective and psychomotor domains in a comprehensive manner. Although holistic evaluation was acknowledged by the experts, cognitive evidence was most strongly emphasised, as the affective and psychomotor domains were perceived as more challenging to assess. The experts pointed out that project-based assessments and problem-solving tasks are effective in measuring cognitive skills, particularly higher-order thinking and problem-solving abilities (Ijirana et al., 2022), as demonstrated by C3 through polymer Chemistry applications. C3 cited an example of a task involving the design of water-based paint using polymer Chemistry principles, which involved "*assessing students' ability to justify their approaches using core Chemistry principles.*" Both quantitative and qualitative justifications provided by students in such tasks can demonstrate critical thinking and creativity. This point was further reinforced by S2, who stated that "*critical-thinking and creativity should be focused on early in tertiary education because these skills require time to develop.*" The experts further argued that STEM-based authentic assessments in Chemistry should measure not only cognitive skills but also affective and psychomotor domains (Sutadji et al., 2021; Wiggins, 1998). A1 expanded on this by recommending "*competition-based learning for students to innovate and be creative in proposing solutions to challenges.*" The implementation of competition-based learning through authentic assessments can increase student motivation while fostering metacognition and knowledge sharing opportunities (Arantes do Amaral et al., 2023).

Meanwhile, the experts also suggested assessing affective skills through surveys, observation, and reflective practices. For instance, C3 recommended "*entrance and exit surveys to measure students' interest in Chemistry throughout their studies*" as a means of comparing whether students' interest in the subject had increased over time (Özkul & Özden, 2020). A2 recommended observing students' engagement and collaboration during lab tasks, while S1 proposed "*analysing student intelligence types through video recordings to understand their attitudes and interests.*" This approach resonates with current findings that affective assessment, which evaluates feelings and internal processes of the mind and heart, requires multiple observations of specific behaviours (Shankar et al., 2023). Reflective writing, as suggested by C2, was also identified as "*a valuable tool for assessing affective skills by encouraging students to explore their experiences with Chemistry,*" express their feelings on specific topics, and relate these to "*their personal goals and values.*" Despite this, self-reported data obtained from surveys and reflective writing may be susceptible to bias, as students may not always express their true feelings or attitudes.

Shifting the focus to psychomotor skills, participants concurred on the importance of psychomotor assessment, particularly through hands-on activities and practical task evaluation. C1 stated that psychomotor skills "*can be measured by conducting a lab test involving accurate apparatus handling and following procedures,*" while C3 and C4 added that such tests "*can assess students' technical competences and application of theoretical knowledge.*" These approaches support the development of a strong knowledge base, research skills, and scientific communication abilities (Kallarackal, 2023). S1 and A2 also supported these approaches, emphasising observation as an effective method for measuring psychomotor skills. The implementation of the CDIO (Conceive, Design, Implement, Operate) framework as suggested by A1, was identified as "*an effective approach for developing psychomotor skills because each phase involves active application of these skills*" during experimentation. However, this approach may present challenges for institutions with limited resources and facilities, highlighting the need for effective implementation that takes into account both evaluative and practical constraints across all three domains (Aziz et al., 2020).

Rubrics

Rubrics generally comprise three sub-themes, namely accommodating various targeted skills, clear criteria, and alignment with learning objectives. Notably, the experts described rubrics as instruments that ensure transparency, fairness, and alignment, rather than as indicators of authenticity. According to C2 and S1, rubrics "*should be divided into specific components to assess different skill areas,*" such as cognitive, psychomotor, and affective skills. In STEM-based authentic assessments in Chemistry, multiple skills need to be assessed; therefore, this targeted approach helps to avoid generalised assessment. As mentioned by A2 and C1, such an approach ensures that each skill is evaluated accurately (Hadibarata et al., 2024).

This, in turn, contributes to the development of clear rubrics that enable students to understand their learning goals and enhance their learning, as noted by A2. The interdisciplinary nature of STEM-based assessment requires clear, concise, and fair evaluation of student work, which can be supported through rubrics with well-defined criteria that specify expectations at each level of achievement (Karim et al., 2018). Rather than relying on general or vague criteria, each rubric should outline the specific abilities and competencies being evaluated in STEM. This specificity helps students understand precisely what is expected of them and how their work will be assessed (Wahyuni et al., 2022).

C4 asserted that rubrics for assessments "*must align with learning objectives.*" As such, it is critical that students undergoing STEM-based authentic assessments in Chemistry are evaluated appropriately, as this supports their performance in both content knowledge and practical skills (Reynders et al., 2020). A1 further stressed that "*involving students in the development of rubrics enhances student understanding of assessment requirements and also serves as a learning tool."* It should be noted that although students' academic achievement can be improved through their participation in rubric development (Fadillah & Ha, 2024), the rigid application of rubrics may limit creativity and overlook individual student progress. Therefore, it is imperative that instructors strike a balance between providing clear, structured criteria and allowing students to apply their knowledge in diverse ways.

Learning Process

The learning process within authentic assessment emphasises active engagement by integrating collaboration, interpersonal skills, independent learning, and constructive feedback. Effective collaboration involves shared responsibility within groups, exposure to diverse perspectives, and peer assessment (Du et al., 2018). However, concerns were raised that competitive approaches and intensive peer assessment may influence collaboration patterns, requiring appropriate scaffolding to maintain fairness. Setting clear expectations for equal participation during group tasks, as pointed by C1, ensures "*fair participation and students [being] responsible for their part of the project."* A2 and A3 further suggested that a "*sense of ownership can be promoted in students through group tasks because successful collaboration requires individuals who are mindful of their responsibilities.*" S1 also highlighted the value of diverse groups comprising students from different academic backgrounds, as this can enrich discussions and problem-solving processes. Moreover, introducing "*competition-based learning in project-based tasks,*" as suggested by A1, can encourage teamwork while supporting knowledge construction, which aligns with Lev Vygotsky's Social Constructivism (Vygotsky, 1978). Various approaches can be used to assess collaboration, such as observing group work, evaluating the quality of discussions and presentations, and monitoring individual contributions to group projects, as noted by C4 and supported by research by Asante (2023). In addition, C3 and A1 highlighted that "*peer assessment allows students to assess their members' contribution and attitude towards the completion of the project.*" However, Alves et al. (2024) argued that student personality can influence the peer assessment process. In response to this, A1 shared that she used "*personality questionnaire tests to allocate students into teams to ensure that students with strong leadership traits"* were not placed in the same team. Similarly, competition-based learning in project-based tasks, as suggested by A1, can encourage teamwork through social constructivist principles (Vygotsky, 1978). However, an excessive focus on rewards may undermine the collaborative value of teamwork in the learning process.

Similarly, C3 highlighted the importance of the "*ability to communicate with peers and lecturers effectively and work independently,*" while C2 and C4 stressed the importance of scientific communication during laboratory work and research projects. These views resonate with the findings of Thibaut et al. (2018) which highlighted the effectiveness of project-based learning in STEM for enhancing students' communication skills. Additionally, C1 explained how interpersonal skills can be strengthened when "*students engage with the community, which is consistent with research*" indicating that science communication skills, science identity, and science self-efficacy are crucial in motivating students' involvement and behaviours in STEM community engagement (Gao et al., 2020). The experts further raised the concern that students should develop the ability to build and maintain both personal and professional relationships, as younger generations nowadays often "*lack the ability to foster respectful relationships, especially with elders.*" In this regard, C1 referred to the "*generational gap in interpersonal skills," particularly* in *"building respectful relationships,*" highlighting the need for focused soft skills development alongside technical education.

In relation to this, A1 and C2 stressed the important role of independent learning strategies, such as "*reflective journals to promote lifelong learning.*" Research evidence demonstrates that nurturing students' ability to learn independently can positively impact their academic performance and cultivate essential skills for lifelong learning (Li, 2023), although this approach may pose challenges for students who require greater scaffolding. According to A3 and C4, reflective journals, portfolios, and self-assessment tools "*can be effective in supporting independent learning."* A1 and A3 further added that capstone projects based on real-world applications challenge students to integrate interdisciplinary knowledge and skills, thereby supporting lifelong learning. The implementation of reflective journals and problem-based learning approaches can empower students to become independent learners, fostering a lifelong learning mindset, and encourage self-assessment (Alt & Raichel, 2022).

Most experts shared the view that effective feedback should be consistent, specific, and constructive. A2 stated that "*face-to-face feedback allows for clear communication,*" minimising the risk of misunderstanding often associated with written feedback, and contributing to significant improvements in task performance (Johnson et al., 2020). According to A1, "*immediate feedback can facilitate students to learn from mistakes and improve continuously.*" In contrast, delayed feedback may hinder students from applying corrections in real-time. Moreover, C1, C4, and A3 echoed the view that "*specific, clear, and actionable feedback should be given*" and should be interpreted accurately to support student progress. This was further reinforced by A1, who made the pertinent point that "*feedback should empower students to build on their strengths,*" as the role of assessment is not only to grade students but also to support learning. Collectively, the experts were in agreement on the role of feedback as a tool for improving student learning, rather than merely serving as a grading mechanism (Ryan et al., 2024).

By highlighting the importance of the criteria for STEM-based authentic assessments in Chemistry, this study supports the relevance of Social Constructivism (Vygotsky, 1978) and Situated Cognition theory (Brown et al., 1989) in authentic assessment. Practical strategies such as project-based and problem-based learning, active engagement, and the use of digital tools align with these theories by promoting holistic skills, ethical awareness, and lifelong learning. Although the criteria were developed for the Malaysian pre-university context, they may also inform assessment practices in other educational settings that adopt authentic and integrated STEM approaches, with variation in implementation depending on institutional resources, policies, and curriculum constraints.

Hence, this study emphasises the importance of STEM-based authentic assessments in Chemistry by illustrating their ability to connect real-world applications, facilitate interdisciplinary learning, encourage the development of holistic skills, highlight the significance of rubrics, and foster an active learning process. Authentic assessment becomes an effective tool when the materials used in lessons are meaningfully connected to students' daily lives. It also promotes critical thinking, higher-order thinking, problem-solving, collaboration, and lifelong learning. Therefore, educators are strongly encouraged to design assessments that integrate interdisciplinary and real-world scenarios.

CONCLUSION

The present study has generated an expert-informed set of proposed criteria for STEM-based authentic assessments in Chemistry within the Malaysian pre-university context. The five themes identified were real-world applications, STEM integration, holistic skills encompassing cognitive, affective, and psychomotor domains, rubric use, and the learning process. Conceptually, these themes establish criteria for evaluating STEM-based authentic Chemistry tasks by clarifying both their fundamental features and the supportive conditions required for effective implementation. This study provides guidelines for educators in designing authentic assessments in Chemistry by integrating STEM practices, while also offering recommendations for Chemistry administrators on improving assessment practices through professional development. At the same time, the findings may support the Ministry of Education Malaysia in equipping students with the competencies needed to meet the demands of the modern workforce.

Although these criteria were established for the Malaysian pre-university system, they offer global relevant guidance by outlining the essential features of STEM-based authentic Chemistry assessment, including real-world tasks, interdisciplinary integration, clear evidence of reasoning and practice, and fair evaluation through rubrics and feedback. Internationally, educators and researchers may use these criteria to review and design assessments that align with learning outcomes and professional standards. However, their implementation should be adapted to suit local curricula, available resources, and assessment policies.

Although expert opinions provide valuable guidance for students, educators and policymakers, this study does not fully examine the challenges educators may encounter when implementing STEM-based authentic assessments, including insufficient time and lack of resources, which may limit the practicality of applying the findings in broader educational contexts. To better understand these limitations, future studies should consider broader populations by including experts from other fields and by incorporating quantitative methods to further validate the proposed criteria. The use of quantitative methods may provide empirical evidence supporting the development of cognitive, affective, and psychomotor skills, and may also demonstrate measurable impacts on learning outcomes through the integration of STEM in the teaching and learning of Chemistry. Furthermore, future studies should explore the practical challenges faced by educators in implementing STEM-based authentic assessments in Chemistry, as well as students' perspectives, which could offer valuable insights for improving their effectiveness.

ACKNOWLEDGMENT

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

REFERENCES

Adebusuyi, O. F., Bamidele, E. F., & Adebusuyi, A. S. (2022). The role of knowledge and epistemological beliefs in chemistry teachers STEM professional development and instructional practices: Examination of STEM-integrated classrooms. *European Journal of Science and Mathematics Education*, *10*(2), 243–255. https://doi.org/10.30935/scimath/11799

Al Sabbah, S., Almomani, J., Drawish, A., & Fares, N. (2022). Traditional versus authentic assessments in higher education. *Pegem Journal of Education and Instruction*, *12*(1), 283–291. https://doi.org/10.47750/pegegog.12.01.29

Alhashem, F., & Alfailakawi, A. (2023). Technology-enhanced learning through virtual laboratories in chemistry education. *Contemporary Educational Technology*, *15*(4), ep474. https://doi.org/10.30935/cedtech/13739

Alt, D., & Raichel, N. (2022). Problem-based learning, self- and peer assessment in higher education: Towards advancing lifelong learning skills. *Research Papers in Education*, *37*(3), 370–394. https://doi.org/10.1080/02671522.2020.1849371

Alves, T., Sousa, F., Gama, S., Jorge, J., & Gonçalves, D. (2024). How personality traits affect peer assessment in distance learning. *Technology, Knowledge and Learning*, *29*(1), 371–396. https://doi.org/10.1007/s10758-023-09694-2

Arantes do Amaral, J. A., Meister, I. P., Sperduti Lima, V., & Grinevicius Garbe, G. (2023). Using competition to improve students' learning in a project-based learning course. *Journal of Problem Based Learning in Higher Education*, *11*(3), 1–24. https://doi.org/10.54337/ojs.jpblhe.v11i3.7514

Asante, G. (2023). Knowledge and practices of authentic assessment in Junior High Schools. *Open Access Library Journal*, *10*(05), 1–14. https://doi.org/10.4236/oalib.1110123

Aziz, M. N. A., Yusoff, N. M., & Yaakob, M. F. M. (2020). Challenges in using authentic assessment in 21st century ESL classrooms. *International Journal of Evaluation and Research in Education*, *9*(3), 759–768. https://doi.org/10.11591/ijere.v9i3.20546

Bozkurt, M., & Ozturk, F. (2022). Qualitative research in social sciences: A research profiling study. *Educational Policy Analysis and Strategic Research*, *17*(1), 247–277. https://doi.org/10.29329/epasr.2022.248.13

Braun, V., & Clarke, V. (2012). Thematic analysis. In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher (Eds.), *APA handbook of research methods in psychology, Vol. 2. Research designs: Quantitative, qualitative, neuropsychological, and biological* (pp. 57–71). American Psychological Association. https://doi.org/10.1037/13620-004

Braun, V., Clarke, V., Boulton, E., Davey, L., & McEvoy, C. (2021). The online survey as a qualitative research tool. *International Journal of Social Research Methodology, 24*(6), 641–654. https://doi.org/10.1080/13645579.2020.1805550

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, *18*(1), 32–42. https://doi.org/10.3102/0013189X018001032

Carley, K. (1990). Content analysis. *The Encyclopaedia of Language and Linguistics*, *2*, 725–730.

Chang, Y., Choi, J., & Şen-Akbulut, M. (2024). Undergraduate students' engagement in project-based learning with an authentic context. *Education Sciences*, *14*(2), 168. https://doi.org/10.3390/educsci14020168

Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). SAGE.

Denzin, N. K. (1989). *The research act: A theoretical introduction to sociological methods* (3rd ed.). Prentice Hall.

Du, W., Liu, D., Johnson, C. C., Sondergeld, T. A., Bolshakova, V. L. J., & Moore, T. J. (2018). The impact of integrated STEM professional development on teacher quality. *School Science and Mathematics*, *119*(2), 105–114. https://doi.org/10.1111/ssm.12318

Fadillah, S. M., & Ha, M. (2024). Using rubrics to enhance accuracy of peer assessment and self-judgment in biology learning. *Asia-Pacific Science Education*, *10*(2), 265–288. https://doi.org/10.1163/23641177-bja10080

Friench, D., Rhew, B., Hauze, S., & Dorsey, S. (2023). Using ePortfolios as an authentic assessment tool for integrated STEM professional development. *Contemporary Issues in Technology and Teacher Education*, *23*(1), 84–102.

Gao, X., Li, P., Shen, J., & Sun, H. (2020). Reviewing assessment of student learning in interdisciplinary STEM education. *International Journal of STEM Education*, *7*, 24. https://doi.org/10.1186/s40594-020-00225-4

Giac, C. C., An, D. T. T., Hiep, L. T. T., Hoang, L. H., & Duc, N. M. (2024). Organizing activities for students of chemistry pedagogy to research according to the CDIO approach in Vietnam. *Journal of Education and E-Learning Research*, *11*(2), 253–262. https://doi.org/10.20448/jeelr.v11i2.5451

Gladwell, M. (2008). *Outliers: The story of success*. Little, Brown and Company.

Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? *Field Methods*, *18*(1), 59–82. https://doi.org/10.1177/1525822X05279903

Hadibarata, T., Hidayat, T., & Kwabena, J. (2024). The use of scoring rubrics in university. *Acta Pedagogia Asiana*, *3*(1), 1–12. https://doi.org/10.53623/apga.v3i1.312

Hussin, M., Said, M. S., Mohd Norowi, N., Husin, N. A., & Mustaffa, M. R. (2021). Authentic assessment for affective domain through student participant in community services. *Asia-Pacific Journal of Information Technology and Multimedia*, *10*(01), 52–62. https://doi.org/10.17576/apjitm-2021-1001-05

Ijirana, I., Aminah, S., Supriadi, S., & Magfirah, M. (2022). Critical thinking skills of chemistry education students in team project-based STEM-metacognitive skills learning during the Covid19 pandemic. *Journal of Technology and Science Education*, *12*(2), 397–409. https://doi.org/10.3926/jotse.1697

Johnson, C. E., Weerasuria, M. P., & Keating, J. L. (2020). Effect of face-to-face verbal feedback compared with no or alternative feedback on the objective workplace task performance of health professionals: A systematic review and meta-analysis. *BMJ Open*, *10*(3), e030672. https://doi.org/10.1136/bmjopen-2019-030672

Kallarackal, A. J. (2023). Benefits of using the invertebrate model organism *C. elegans* in an inquiry-based laboratory activity. *Teaching of Psychology*, *50*(1), 26–31. https://doi.org/10.1177/00986283211029649

Karim, A. A., Abduh, A., Manda, D., & Yunus, M. (2018). The effectivity of authentic assessment based character education evaluation model. *TEM Journal*, *7*(3), 495–500. https://doi.org/10.18421/TEM73-04

Kelley, T. R., & Knowles, J. G. (2016). A conceptual framework for integrated STEM education. *International Journal of STEM Education*, *3*(1), 11. https://doi.org/10.1186/s40594-016-0046-z

Kriswantoro, Kartowagiran, B., & Rohaeti, E. (2021). A critical thinking assessment model integrated with science process skills on chemistry for Senior High School. *European Journal of Educational Research*, *10*(1), 285–298. https://doi.org/10.12973/eu-jer.10.1.285

Li, N. (2023). Research on the cultivation of students' independent learning ability. *Adult and Higher Education*, *5*(19), 49–54. https://doi.org/10.23977/aduhe.2023.051906

Martín-Páez, T., Aguilera, D., Perales-Palacios, F. J., & Vílchez-González, J. M. (2019). What are we talking about when we talk about STEM education? A review of literature. *Science Education*, *103*(4), 799–822. https://doi.org/10.1002/sce.21522

McNay, I. (2007). Values, principles and integrity: Academic and professional standards in higher education. *Higher Education Management and Policy*, *19*(3), 34–57.

Ministry of Education Malaysia. (2013). *Malaysia Education Blueprint 2013-2025*. Ministry of Education Malaysia.

Ministry of Higher Education Malaysia. (2026). *Malaysia Higher Education Blueprint 2026-2035*.

Nadelson, L. S., & Seifert, A. L. (2017). Integrated STEM defined: Contexts, challenges, and the future. *The Journal of Educational Research*, *110*(3), 221–223. https://doi.org/10.1080/00220671.2017.1289775

Nahadi, Purnawarman, P., & Siswaningsih, W. (2021). Development of an electronic portfolio assessment model in learning chemistry to develop the habits of mind and reasoning of Indonesian students. *Journal of Physics: Conference Series*, *1788*(1). https://doi.org/10.1088/1742-6596/1788/1/012034

Nawawi, N. M., Sout, N. M., Hassan, K. B., Samah, N. N. A., Kamaruddin, H. H., Khalid, R. M., & Azman, H. H. (2021). The perception of pre-university students on STEM. *Journal of Physics: Conference Series*, *1882*(1), 012155. https://doi.org/10.1088/1742-6596/1882/1/012155

Özkul, H., & Özden, M. (2020). Investigation of the effects of engineering-oriented STEM integration activities on scientific process skills and STEM career interests: A mixed methods study. *Education and Science*, *45*(204), 41-63. https://doi.org/10.15390/EB.2020.8870

Reynders, G., Lantz, J., Ruder, S. M., Stanford, C. L., & Cole, R. S. (2020). Rubrics to assess critical thinking and information processing in undergraduate STEM courses. *International Journal of STEM Education*, *7*(1), 9. https://doi.org/10.1186/s40594-020-00208-5

Roberts, T., Maiorca, C., Jackson, C., & Mohr-Schroeder, M. (2022). Integrated STEM as problem-solving practices. *Investigations in Mathematics Learning*, *14*(1), 1–13. https://doi.org/10.1080/19477503.2021.2024721

Ryan, T., Henderson, M., Ryan, K., & Kennedy, G. (2024). Feedback in higher education: Aligning academic intent and student sensemaking. *Teaching in Higher Education*, *29*(4), 860–875. https://doi.org/10.1080/13562517.2022.2029394

Shafiyeva, U. (2021). Assessing students' minds: Developing critical thinking or fitting into procrustean bed. *European Journal of Education*, *4*(2), 78–91. https://doi.org/10.26417/452bxv17s

Shankar, N., Fathima, F. N., & Lakshmi, T. A. (2023). Assessment of attitudes, ethics and communication skills in a competency-based undergraduate medical curriculum. *South-East Asian Journal of Medical Education*, *17*(1), 4–11. https://doi.org/10.4038/seajme.v17i1.544

Shenton, A. K. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Education for Information*, *22*(2), 63–75. https://doi.org/10.3233/EFI-2004-22201

Shidiq, A. S., Permanasari, A., & Hernani. (2020). *Chemistry teacher's perception toward STEM learning.* Proceedings of the 2020 International Conference on Education Development and Studies, 40–43. https://doi.org/10.1145/3392305.3396901

Sırakaya, M., & Alsancak Sırakaya, D. (2022). Augmented reality in STEM education: A systematic review. *Interactive Learning Environments*, *30*(8), 1556–1569. https://doi.org/10.1080/10494820.2020.1722713

StudyMalaysia. (2017). *Study in Malaysia handbook* (10th ed.). Challenger Concept (M) Sdn Bhd.

Sutadji, E., Susilo, H., Wibawa, A. P., Jabari, N. A. M., & Rohmad, S. N. (2021). Authentic assessment implementation in natural and social science. *Education Sciences*, *11*(9). https://doi.org/10.3390/educsci11090534

Syafdaningsih, & Suningsih, T. (2020). Development of authentic assessment model in the field of mathematic for early childhood in Palembang City. *Journal of Physics: Conference Series*, *1480*(1). https://doi.org/10.1088/1742-6596/1480/1/012077

Sysoyev, P. V. (2024). Ethics and AI-plagiarism in an academic environment: Students' understanding of compliance with author's ethics and the problem of plagiarism in the process of interaction with generative artificial intelligence. *Vysshee Obrazovanie v Rossii = Higher Education in Russia*, *33*(2), 31–53. https://doi.org/10.31992/0869-3617-2024-33-2-31-53

Thibaut, L., Ceuppens, S., De Loof, H., De Meester, J., Goovaerts, L., Struyf, A., Boeve-de Pauw, J., Dehaene, W., Deprez, J., De Cock, M., Hellinckx, L., Knipprath, H., Langie, G., Struyven, K., Van de Velde, D., Van Petegem, P., & Depaepe, F. (2018). Integrated STEM education: A systematic review of instructional practices in secondary education. *European Journal of STEM Education*, *3*(1), 1–12. https://doi.org/10.20897/ejsteme/85525

van Teijlingen, E., & Hundley, V. (2002). The importance of pilot studies. *Nursing Standard*, *16*(40), 33–36. https://doi.org/10.7748/ns2002.06.16.40.33.c3214

Villarroel, V., Bloxham, S., Bruna, D., Bruna, C., & Herrera-Seda, C. (2018). Authentic assessment: Creating a blueprint for course design. *Assessment & Evaluation in Higher Education*, *43*(5), 840–854. https://doi.org/10.1080/02602938.2017.1412396

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Harvard University Press.

Wahyuni, S., Ikashaum, F., Wulantina, E., Mustika, J., & Putri, L. M. (2022). *Development of authentic assessment models in research methods courses*. Proceedings of the Eighth Southeast Asia Design Research (SEA-DR) & the Second Science, Technology, Education, Arts, Culture, and Humanity (STEACH) International Conference (SEADR-STEACH 2021), 98–102. https://doi.org/10.2991/assehr.k.211229.016

Wan Yusof, W. R., Kimi, M., Wan Zullkiplee, W. S. H., Zailani, M. A., Shahabudin, M., Ismail, A. A.-H., Abd. Aziz Abdullah, S. M., Ismail, I. N. A., & Raja Gopal, D. J. (2022). Assessment of practical and scientific writing skills for pre-university students through project-based learning. *Journal of Chemical Education*, *99*(2), 715–722. https://doi.org/10.1021/acs.jchemed.1c00843

Widana, I. W., Sopandi, A. T., & Suwardika, G. (2021). Development of an authentic assessment model in mathematics learning: A science, technology, engineering, and mathematics (STEM) approach. *Indonesian Research Journal in Education |IRJE|*, *5*(1), 192–209. https://doi.org/10.22437/irje.v5i1.12992

Wiggins, G. (1998). *Educative assessment: Designing assessments to inform and improve student performance*. Jossey-Bass Publishers.

Williams, N. D., Gallardo-Williams, M. T., Griffith, E. H., & Bretz, S. L. (2022). Investigating meaningful learning in virtual reality organic chemistry laboratories. *Journal of Chemical Education*, *99*(2), 1100–1105. https://doi.org/10.1021/acs.jchemed.1c00476