Jurnal Ilmiah Pendidikan Fisika
https://ppjp.ulm.ac.id/journals/index.php/jipf/index
e-ISSN: 2549-9963 p-ISSN: 2549-9955
Vol 9 No 2 2025
DOI: https://doi.org/10.20527/jipf.v9i2.13259

Validity of Assessment Instruments for Evaluating the Feasibility of Developing an Integrated STEM Inquiry Model in 21st-Century Education

Isnaini*, Atmazaki, Yuni Ahda, and Suci Fajrina
Education Science Doctoral Program, Universitas Negeri Padang, Padang, Indonesia
*isnaini010190@gmail.com

Article Info

Article history:
Received: May 9, 2025
Accepted: June 16, 2025
Published: July 2, 2025

Keywords:
21st century skills
Assessment validity
Expert validation
Instrument validity
Integrated STEM inquiry

This article is licensed under a creative commons attribution-share alike 4.0 international license

© 2025 Jurnal ilmiah Pendidikan Fisika

Abstract

The validity of assessment instruments is a crucial element in educational research, particularly in evaluating the success of a learning model. In developing an integrated STEM inquiry learning model, it is necessary first to validate the assessment instruments to ensure they can effectively measure the desired learning outcomes, including both academic knowledge and the 21st-century skills developed by students. This study evaluates the validity of the instruments used to assess the feasibility of the integrated STEM inquiry model, which includes the model book, teacher's book, student's book, and a 21st-century skills assessment instrument. The instruments were validated through expert judgment involving three validators: an educational expert in evaluation, a STEM curriculum developer, and a language and communication practitioner. The instruments were assessed using a five-point Likert scale, and the data were analyzed using descriptive statistics to determine content, construct, and criterion validity. The results show that the 21st-century skills instrument scored 0.89, the model book scored 0.84, the student book scored 0.87, and the teacher book scored 0.86, all falling into the very high validity category. These findings indicate that the instruments are appropriate and reliable for measuring both academic competencies and essential 21st-century skills such as problem-solving, creativity, and collaboration. The study concludes that employing valid assessment instruments is essential for the effective implementation of the integrated STEM inquiry model and for ensuring its success in fostering the holistic development of students. The findings also highlight the importance of instrument validation in educational settings to ensure that learning models achieve their intended outcomes and prepare students for future challenges.

To cite this article: Isnaini, I., Atmazaki, A., Ahda, Y., & Fajrina, S. (2025). Validity of assessment instruments for evaluating the feasibility of developing an integrated STEM inquiry model in 21st-century education. *Jurnal Ilmiah Pendidikan Fisika, 9*(2), 251-262.

INTRODUCTION

The validity of the assessment instrument is a crucial element in educational research (Agusta et al., 2024), especially when assessing the effectiveness of a learning model. In developing an integrated STEM inquiry learning model, validating the instrument ensures it can appropriately measure both the academic and non-academic outcomes targeted by the model (Gunada et al., 2023). These outcomes include mastery of subject matter and the cultivation of essential 21st-century skills such as critical thinking, creativity, collaboration, and communication. Without proper validation, instruments risk producing misleading results that compromise the integrity of educational assessments (Kaewmanee et al., 2024).

The integrated STEM inquiry learning model aims to provide students with a more

comprehensive understanding of how the fields of science, technology, engineering, and mathematics (STEM) are interconnected in solving real-world problems (Capraro et al., 2013). This approach not only teaches theoretical content in each discipline but also teaches students to think critically and apply their knowledge to practical situations (Martín-Páez et al., 2019; Pratama et al., 2020). Inquiry-based instruction encourages active student participation through questions, experiments, and investigation (Novitra et al., 2021), which helps foster meaningful engagement in science and mathematics learning.

Despite the growing attention to STEM education, research on validating integrated assessment tools, especially those that assess both cognitive and non-cognitive dimensions, is still limited. Most existing studies assess content knowledge but fail to capture soft skills such as creativity, resilience, and teamwork, which are critical for success in modern contexts (Anggraini & Huzaifah, 2017). Furthermore, many tools rely on traditional assessments (e.g., multiple-choice tests) that do not adequately reflect the multifaceted competencies emphasized in STEM learning (Setyowati et al., 2024).

Integrated STEM inquiry emphasizes active student participation in the learning process (Ješková et al., 2022; Kaewmanee et al., 2024; Kotsis, 2024; Lafifa et al., 2023). Instead of being passive receivers of information, students are encouraged to explore knowledge through experiments, observations, and deep reflection. Inquiry as a learning method allows students to engage in the actual scientific process, where they can ask questions, design experiments, and solve problems based on the data they obtain (Novitra et al., 2021).

To address this gap, this study proposes the validation of a comprehensive set of instruments designed to assess the feasibility, practicality, and effectiveness of the integrated STEM inquiry learning model. These instruments include validation tools for the model book, teacher's book, and student's book, as well as specific tools to evaluate students' 21st-century skills. The study aligns with prior findings that underscore the importance of developing instruments that can assess both content knowledge and critical soft skills.

The importance of assessment instrument validity in education has been widely discussed in the literature, but research focusing on instrument validation for measuring 21st-century skills in STEM contexts remains limited. Previous studies, such as those conducted by Anggraini and Huzaifah (2017) and Redhana (2019), emphasize the need for the development of assessment instruments that not only assess academic knowledge but also non-cognitive skills, including creativity, collaboration, and problem-solving. These skills have become increasingly relevant as the demand for these abilities grows in today's complex work environments.

Earlier research also indicates that most assessment instruments in STEM education continue to focus heavily on content knowledge, employing traditional methods such as written exams and multiple-choice tests. However, these instruments often fail to adequately measure the essential non-cognitive skills, which are crucial for real-world problem-solving (Fajrina et al., 2020). Hence, there is a pressing need for innovative assessment tools that can evaluate 21st-century skills such as critical thinking, creativity, and collaboration, which are vital for preparing students for the future challenges they will face.

Several studies have also highlighted the importance of validating assessment instruments in STEM contexts. For example, research by Kaewmanee et al. (2024) emphasizes that invalid instruments cannot yield reliable results for assessing the feasibility of using textbooks in STEM learning. This highlights the need for a thorough validation process for assessment instruments to ensure they accurately measure both academic knowledge and 21st-century skills developed through integrated STEM inquiry learning.

Through this study, the authors aim to address the gap identified in previous research

by validating assessment instruments that measure both academic content and 21st-century skills. This research aims to make a significant contribution to the development of assessment instruments that can effectively evaluate both of these aspects, ultimately improving the quality of STEM education and preparing students for future success.

As the world becomes more interconnected through technological advancements, 21st-century skills have become essential for student success. These skills include critical thinking, creativity, collaboration, and communication, all of which are necessary for students to thrive in the workplace. Current STEM education assessment tools largely focus on academic knowledge but fail to adequately measure these important non-cognitive skills (Fajrina et al., 2020).

The validation of the assessment instrument is a very important process in STEM education, especially when used in an integrated inquiry learning model. Without a valid instrument, the evaluation of this learning model will be inaccurate, which in turn can affect the implementation of the model in the classroom. Invalid instruments can give results that do not reflect the ability of students to develop 21st-century skills, which ultimately affects the quality of the learning being implemented. Therefore, the assessment instrument used in this study needs to be tested first to ensure its feasibility in measuring the results to be achieved, both in terms of knowledge and skills (Abdullah et al., 2022).

Most existing assessment instruments in STEM education primarily assess academic knowledge through traditional methods, such as multiple-choice tests and written exams. These methods often fail to measure non-cognitive skills, such as problem-solving, collaboration, and creativity, which are equally important for students' future success (Redhana, 2019). Thus, there is a need for innovative assessment tools that incorporate real-world scenarios to evaluate students' ability to think critically, work in teams, and solve complex problems. By addressing these gaps, the development of new multidimensional assessment instruments will provide a more comprehensive understanding of student achievement and skill development.

The practicality of the assessment instruments is crucial to their success in diverse classroom settings. In STEM education, classrooms vary significantly in terms of resources, student abilities, and teaching environments. Therefore, assessment tools must be easy to use and adaptable to these varying conditions (Afriana, Permanasari, & Fitriani, 2016). The instruments used to assess the model book, teacher's book, and student's book should not only measure student learning outcomes but also be simple enough for educators to implement without requiring excessive time, training, or resources. This ensures that practicality is a key feature of the assessment instruments, making them suitable for use in real-world classroom settings.

Many assessment instruments currently in use focus on academic content but neglect the measurement of critical non-cognitive skills such as problem-solving, creativity, and collaboration. These skills are crucial in STEM education, especially when students are tasked with solving complex, real-world problems. This study addresses this limitation by developing assessment tools that focus on real-world problem-solving scenarios, where students apply their STEM knowledge in practical situations. These innovative tools will capture the true level of students' ability to think critically and creatively, providing a more accurate measure of their overall skill development.

One of the greatest challenges in developing assessment instruments for STEM education is the lack of tools that can measure all dimensions of integrated learning. Most existing assessment instruments focus more on measuring students' academic achievement, while non-cognitive skills such as critical thinking, creativity, and collaboration are not well measured through traditional assessment methods (Anggraini & Huzaifah, 2017; Nuraini, 2017; Nurhaifa et al., 2020). Therefore, it is important to develop assessment

instruments that not only measure academic content mastery but also evaluate the development of 21st-century skills in students.

Furthermore, the assessment instrument used must be practical and easily applicable in diverse classrooms. STEM learning requires educators to apply models that actively engage students and allow them to develop relevant skills (Anita et al., 2021). The instruments used to assess the success of this model must be easy to use by educators without requiring excessive time or resources (Anwar et al., 2014; Ardiansyah et al., 2020; Bhakti et al., 2018; Dewi et al., 2021). Therefore, a valid instrument should not only accurately measure learning outcomes but also be designed to be easily applied in diverse classroom settings with varying student abilities (Anita et al., 2021; Bhakti et al., 2018; Geschwind et al., 2024; Kurnia et al., 2021).

21st-century skills, including critical thinking, creativity, collaboration, and communication, are essential for students' success in a knowledge- and technology-based workplace (AlAli, 2024; Redhana, 2019; Utami, 2018; Yokhebed, 2019). Therefore, the assessment instrument used in this study not only aims to measure students' academic achievement but also to evaluate how the integrated STEM learning model can develop these skills in students. The assessment of 21st-century skills is crucial because it ensures that the learning model applied not only teaches scientific knowledge but also prepares students to face future challenges.

The instruments include a validation questionnaire and a validation sheet, both of which employ a 5-point Likert scale. These tools are assessed by three expert validators: a curriculum developer, an evaluation specialist, and a practitioner experienced in STEM education. Their feedback is analyzed using descriptive statistics and categorized based on content validity index (CVI) interpretations. Scores ranging from 0.81 to 1.00 are categorized as very high, reflecting strong validity. Hence, the instrument used to evaluate this learning model must be closely considered so that it can provide a clearer picture of the effectiveness of the model in developing relevant skills for students' future (Papagiannis & Pallaris, 2024; Xiaoming & Wiebe, 2023; Yang et al., 2025).

The practicality of the assessment instrument is crucial to ensure it can be effectively used in various educational contexts. STEM learning requires a model that can be applied to various types of classrooms with different levels of student abilities and limited resources (Anita et al., 2021; Ardiansyah et al., 2020; Asyshifa Astri et al., 2024; Lestari et al., 2025). Therefore, the instruments used to assess the model book, teacher's book, and student's book must be designed to be easily applied by educators without disrupting the learning process. A valid and practical assessment instrument will ensure that teachers can use it effectively to evaluate how well this learning model works in improving 21st-century skills in students. The practicality of this assessment instrument greatly depends on how easily it can be applied in diverse classrooms and how quickly the results from the assessment can be used to make decisions about improving the quality of STEM learning.

This study offers a methodological and practical contribution by providing tools that can be used by educators, researchers, and curriculum designers to evaluate both the structural and pedagogical aspects of STEM-based learning. By validating these tools, the study supports efforts to improve educational quality, student preparedness, and alignment between assessment and instructional goals.

In conclusion, valid and practical assessment instruments are necessary for the success of integrated STEM inquiry models. They ensure that both academic and 21st-century skills are adequately measured and supported. As education evolves to meet future demands, the tools used to evaluate learning must also evolve, moving beyond content mastery toward a richer, multidimensional understanding of student growth.

METHODS

This study adopts a quantitative approach with a descriptive design to measure the validity of the assessment instruments used for evaluating the feasibility of the integrated STEM inquiry-based learning model. This design was selected to gain a clear picture of how well the instruments can measure the feasibility aspects of the learning model within the context of STEM education and their influence on students' 21st-century skills.

The subjects of this study include three validators, consisting of educational experts, curriculum developers, and practitioners proficient in STEM education and assessment instruments. These validators were selected based on their expertise in evaluation, subject matter, and language, ensuring they are capable of providing objective evaluations of the instruments' feasibility. The validators included (1) One educational expert with expertise in evaluation techniques, (2) One curriculum developer specializing in STEM education, and (3) One practitioner with proficiency in language and communication within the educational context. Each validator's extensive knowledge and experience in their respective fields contribute to ensuring a comprehensive and objective assessment of the instruments.

This study employs two main instruments: validation of the questionnaire and validation sheet. The validation of the questionnaire is designed to gather evaluations from validators regarding the feasibility of the assessment instruments in developing the integrated STEM inquiry-based learning model. The questionnaire consists of a set of questions focused on the overall effectiveness and relevance of the instruments. Validators are asked to assess the appropriateness, clarity, and alignment of the assessment tools with the learning objectives of the integrated STEM inquiry-based model. Rating Scale: The questionnaire uses a 5-point Likert scale, where 1 = Strongly Disagree; 2 = Disagree; 3 = Neutral; 4 = Agree; and 5 = Strongly Agree. This scale allows for a clearer understanding of the validators' perspectives on the instruments' feasibility.

The validation sheet contains a more detailed set of questions or specific indicators that the validators must assess regarding the feasibility of the instruments. It focuses on critical aspects, such as the relevance of the questions, clarity of instructions, and the ability of the instruments to measure the desired academic and non-cognitive skills (including critical thinking, problem-solving, and creativity) in students. Rating Scale: Similar to the Questionnaire Validation, the 5-point Likert scale is used for the validation sheet to rate the instruments on specific indicators. Validators rate each statement based on their judgment of how well the instruments meet the required criteria.

After calculating the average score for each indicator or statement on the Validation Questionnaire and Validation Sheet, the raw scores obtained using a 5-point Likert-type scale are divided by the maximum possible score (i.e., 5) to obtain a normalized value between 0 and 1. These normalized values represent the degree of agreement among validators and are then categorized based on the interpretation of thresholds adapted from Aiken's V and the Content Validity Index (CVI), as presented in Table 1.

Table 1 Category based on thresholds interpretation

Score Range	Category	Interpretation
0.00 – 0.20	Very Low	The instrument is irrelevant and not suitable for use in the integrated STEM inquiry-based learning model.
0.21 – 0.40	Low	The instrument has low relevance and requires major revisions.
0.41 – 0.60	Moderate	The instrument is fairly relevant but needs improvement to align with the learning objectives.
0.61 – 0.80	High	The instrument is relevant and suitable for use, with only minor adjustments needed.
0.81 – 1.00	Very High	The instrument is highly relevant and fully appropriate for use in the intended learning model.

Based on the above categories, the researchers concluded the feasibility of the validated instruments. High average scores indicate that the instrument is appropriate for achieving the learning objectives, while low average scores suggest that the instrument needs significant revision. The data obtained from the validation questionnaires were analyzed using descriptive analysis. The average scores from each validator to assess the appropriateness and relevance of the instruments in the context of STEM education were calculated.

The validation process involved several systematic steps to ensure the feasibility and content validity of the assessment instruments used in this study. The instruments were developed based on a synthesis of relevant literature on integrated STEM inquiry learning and 21st-century skills (e.g., P21 Framework, OECD Learning Compass) and were refined through expert consultation.

1. Validation Questionnaire, used to assess the model book, teacher's book, and student's book. Each version of this instrument includes eight items across domains such as instructional clarity, alignment with learning objectives, structural completeness, and language appropriateness.
   Example item (teacher's book): "The instructional structure and teaching guidance are presented and aligned with STEM inquiry principles."
2. Validation Sheet for 21st-Century Skills Assessment Instrument – This instrument includes seven indicators that evaluate clarity, relevance, item structure, assessment coverage, and language usage.
   Example item: “The instrument covers essential 21st-century skills such as critical thinking, collaboration, and creativity.”
3. Practicality Questionnaires – Used to assess how practical the books and learning model are when applied by students and teachers. Each instrument contains seven items focusing on user-friendliness, clarity of instructions, ease of implementation, and readability. These were administered to both students and teachers.
   Example item: “The student book can be used independently and is easy to understand.”

Data were collected by administering the validation instruments to three expert validators with expertise in curriculum, evaluation, and STEM-based pedagogy. Content validity was analyzed descriptively by calculating average scores and interpreting them in relation to the CVI scale. In addition, Aiken's V index was used to assess inter-rater agreement. Values between 0.81 and 1.00 were interpreted as indicating very high content validity.

RESULTS AND DISCUSSION

The results of the instrument validation in this study cover various aspects relevant to assessing the feasibility of the learning model, student books, teacher books, and the 21st-century skills assessment instruments. The assessment was conducted by validators who have expertise in their respective fields. The instruments tested achieved average scores that indicate high quality in almost all of the aspects evaluated. Table 2 presents the results of the instrument validation, which include the average scores for each evaluated aspect.

Table 2 Summary of validator ratings for content validity of assessment instruments

Validation Category	Average Score	Rating Category	Evaluation Aspects
21st Century Skills Assessment Instrument	0.89	Very High	Clarity of instructions, Alignment with the blueprint, Consistency of terms, Completeness of book components, Correct use of language
Model Book Validation Instrument	0.84	Very High	Clarity of instructions, Alignment with the blueprint, Consistency of terms, Completeness of book components, Correct use of language
Student Book Validation Instrument	0.87	Very High	Clarity of instructions, Alignment with the blueprint, Consistency of terms, Completeness of book components, Correct use of language
Teacher Book Validation Instrument	0.86	Very High	Clarity of instructions, Alignment with the blueprint, Consistency of terms, Completeness of book components, Correct use of language

Table 3 provides the practicality validation results, showing how practical the instruments are for use in real-world classroom settings based on the validators' assessments.

Table 3 Practicality ratings of the integrated STEM inquiry learning model by teachers and students

Validation Category	Average Score	Rating Category	Evaluation Aspects
Model Practicality Validation According to Students	0.81	Very High	Clarity of instructions, Alignment with the blueprint, Consistency of terms, Completeness of book components, Correct use of language
Model Practicality Validation According to Teachers	0.84	Very High	Clarity of instructions, Alignment with the blueprint, Consistency of terms, Completeness of book components, Correct use of language
Student Book Practicality Validation According to Students	0.85	Very High	Clarity of instructions, Alignment with the blueprint, Consistency of terms, Completeness of book components, Correct use of language
Student Book Practicality Validation According to Teachers	0.86	Very High	Clarity of instructions, Alignment with the blueprint, Consistency of terms, Completeness of book components, Correct use of language
Teacher Book Practicality Validation According to Teachers	0.87	Very High	Clarity of instructions, Alignment with the blueprint, Consistency of terms, Completeness of book components, Correct use of language

Instrument Feasibility Analysis

Based on the validation results obtained, the instruments used in this study demonstrate excellent feasibility in assessing the learning model and the books being evaluated. All instruments achieved high average scores, indicating that they are valid and reliable for evaluating the various components involved in the learning model, student books, and teacher books.

The 21st Century Skills Assessment Instrument, with an average score of 0.89, falls into the High category, indicating that this instrument is effective in measuring skills relevant to the 21st century. The validator provided positive feedback regarding the clarity of instructions, alignment with the established blueprint, and the completeness of the book components.

The Model Book Instrument received an average score of 0.84, which also falls into the High category. The aspects evaluated include clarity of instructions, consistency of terms, and completeness of book components. This score indicates that the instrument can be used to assess the feasibility of model books with satisfactory results, although there were recommendations for refining some technical terms.

The Student Book and Teacher Book instruments received average scores of 0.87 and 0.86, respectively, both of which show very good results in the high category. The validators paid particular attention to the completeness of the book components, as this instrument was able to cover all the necessary aspects for comprehensively assessing student and teacher books.

The Model Practicality Instrument, according to students and teachers, received average scores of 0.81 and 0.84, respectively, which also fall within the high category. Although there were some suggestions related to simplifying the instructions, this instrument effectively assesses the practicality of the learning model.

The instrument validation in this study indicates that the instruments used are highly reliable and valid for assessing the feasibility of the learning model and the books evaluated. Most aspects evaluated, whether for model books, student books, teacher books, or 21st-century skills assessments, received high scores, indicating the effectiveness and feasibility of the instruments. This supports the construct validity of the instruments, which is crucial for ensuring that they measure what they are intended to assess. Recent studies also affirm that a well-designed instrument with high validity plays a key role in educational assessments (Cohen et al., 2017).

However, there were some suggestions from the validators, particularly concerning the refinement of technical terms and the simplification of instructions. These improvements aim to make the instruments more understandable for an audience with different backgrounds. This aligns with cognitive load theory, which emphasizes the importance of minimizing extraneous cognitive load by making instructions clear and accessible, ensuring that learners focus on the essential content. Simplifying the language and refining technical terms can significantly enhance the usability of the instruments for diverse audiences.

For example, the assessment of Model Practicality According to Students received a lower score (0.81) compared to other instruments, indicating that there is room for improvement in the clarity of the instructions to make them more understandable for students. This is in line with instructional design theory, which stresses that clear and well-structured instructions are fundamental to successful learning outcomes. It also highlights the importance of considering learners' cognitive capacity and prior knowledge when developing educational tools.

Overall, these instruments are highly reliable and effective in evaluating various aspects of the feasibility of the learning model and related books. The validation results provide a strong foundation for the further development of the instruments used in this study, which, in turn, enhance the effectiveness of implementing the learning model in practice. This process of iterative improvement aligns with feedback loop theory in educational research, where continuous evaluation and adjustments lead to better alignment with learning goals (Sinharay, 2017).

The validators also provided several important suggestions, such as improving the consistency of terms and separating aspects more specifically in the statements. These

improvements make it easier for users of the instruments to provide more accurate and objective evaluations. This is consistent with principles of clarity and consistency in assessment design, which emphasize that well-defined criteria and language reduce ambiguity and enhance the reliability of evaluations.

Furthermore, the availability of validated assessment instruments has practical implications for classroom implementation. Teachers can adopt the integrated STEM inquiry model with greater confidence, knowing that the evaluation tools are aligned with learning objectives and sensitive to both academic and 21st-century skill development. These instruments serve as a bridge between instructional design and assessment, ensuring that learning activities are meaningfully measured. As the instruments were developed using generalizable frameworks such as the P21 Framework and the OECD Learning Compass, they hold the potential to be scaled and adapted for use in various educational contexts beyond the original research setting.

Despite the strong validity results, this study is not without limitations. One notable limitation is the absence of field testing or pilot implementation in real classroom environments. Future research should consider piloting these tools in authentic educational settings to examine their practicality and usability further. Field testing could also provide insight into predictive validity and user response patterns under actual instructional conditions.

CONCLUSION

This study aimed to validate assessment instruments for evaluating the feasibility of an integrated STEM inquiry-based learning model. Results show that the instruments are highly valid and reliable across various aspects, including clarity of instructions, alignment with learning objectives, and appropriate use of language.

These validated instruments are practical and applicable for supporting curriculum development, classroom implementation, and assessing 21st-century skills. Suggestions from validators, such as simplifying instructions and refining terminology, will improve their usability across diverse educational settings.

The findings suggest that policymakers and curriculum developers in education can adopt these tools as part of STEM evaluation frameworks. Likewise, teacher training programs are encouraged to integrate the instruments into professional development to enhance instructional quality in STEM classrooms.

CONFLICTS OF INTEREST

The authors declare that there is no conflict of interest.

AUTHOR CONTRIBUTIONS STATEMENT

The authors would like to express their sincere gratitude to the instrument validators for their valuable feedback and contributions in ensuring the validity of the research instruments used in this study.

REFERENCES

Abdullah, M., M. Hasan, A., & Ahmad, J. (2022). Uji validitas pengembangan perangkat pembelajaran terintegrasi model inkuiri terbimbing materi pertumbuhan dan perkembangan tumbuhan untuk meningkatkan hasil belajar kognitif peserta didik. *BIOEDUKASI (Jurnal Pendidikan Biologi)*, *13*(2), 179. https://doi.org/10.24127/bioedukasi.v13i2.6346

Afriana, J., Permanasari, A., Fitriani, A. (2016). Penerapan project based learning terintegrasi stem untuk meningkatkan literasi sains peserta didik ditinjau dari gender. *Jurnal Inovasi Teknologi Pendidikan IPA*, *2*(2), 202–212.

Agusta, F., Rahmawati, A. Z., & Nurohman, S. (2024). Validation of interactive e-worksheet based on a scientific approach assisted by solar system scope. *Jurnal Pendidikan Fisika Dan Teknologi*, *10*(2), 312–319. https://doi.org/10.29303/JPFT.V10I2.7280

AlAli, R. (2024). Enhancing 21st century skills through integrated stem education using project-oriented problem-based learning. *Geojournal of Tourism and Geosites* , *53*(2), 421–430. https://doi.org/10.30892/GTG.53205-1217

Anggraini, F. I., & Huzaifah, S. (2017). Implementasi STEM dalam pembelajaran ipa di sekolah menengah pertama. *Program Studi Pendidikan Biologi Fakultas Keguruan Dan Ilmu Pendidikan Universitas Sriwijaya*, *1*(1), 722–731. https://conference.unsri.ac.id/Index.Php/Semnasipa/Article/View/738

Anita, Y., Thahir, A., & Dwi, R. N. (2021). Buku saku digital berbasis STEM: Pengembangan media pembelajaran terhadap kemampuan pemecahan masalah. *Mosharafa: Jurnal Pendidikan Matematika*, *10*(3), 401–412. https://doi.org/10.31980/MOSHARAFA.V10I3.672

Anwar, M., Fitrani, E., Beers, B., Bennett, R. E., Bransford, J., Brown, A. L., Cocking, R. R., Bybee, R. W., Darling-Hammond, L., Adamson, F., Fadel, C., Prakash, N., Freeman, S., others, Hattie, J., Timperley, H., Council, N. R., Perkins, D., Saavedra, A. R., … Willett, T. (2014). A framework for k-12 science education: Practices, crosscutting concepts, and core ideas. *Technology and Engineering Teacher*, *70*(1), 15–25.

Ardiansyah, R., Diella, D., Herni, D., & Suhendi, Y. (2020). Pelatihan pengembangan perangkat pembelajaran abad 21 dengan model pembelajaran project based learning berbasis stem bagi guru ipa. *Publikasi Pendidikan*, *10*(1).

Asyshifa, A. N., Hufri, H., Gusnedi, G., & Yulia Sari, S. (2024). Validasi e-modul fisika berbasis model pembelajaran inkuiri terbimbing pada materi gelombang kelas xi sma/ma. *Jurnal Pendidikan Tambusai*, *8*(1), 5320–5330. https://doi.org/10.31004/jptam.v8i1.13214

Bhakti, C. P., Ghiffari, M. A. N., & Regita, S. M. (2018). Strategy of core curriculum to improving student's critical thinking skill. *TERAPUTIK: Jurnal Bimbingan Dan Konseling*, *1*(3), 176. https://doi.org/10.26539/1374

Capraro, R. M., Capraro, M. M., & Morgan, J. R. (2013).*STEM project-based learning: An integrated science, technology, engineering, and mathematics (stem) approach, second edition*. Sense Publishers. https://doi.org/10.1007/978-94-6209-143-6

Cohen, L., Manion, L., & Morrison, K. (2017). *Research Methods in Education*. https://doi.org/10.4324/9781315456539/research-methods-education-lawrence-manion-louis-cohen-keith-morrison

Dewi, D. K., Ardhana, W., Irtadji, Chusniyah, T., & Sulianti, A. (2021). Inquiry-based learning implementation to improve critical thinking of prospective teachers. *International Journal of Information and Education Technology*, *11*(12), 638–645. https://doi.org/10.18178/IJIET.2021.11.12.1575

Fajrina, S., Lufri, L., & Ahda, Y. (2020). Science, technology, engineering, and mathematics (STEM) as a learning approach to improve 21st century skills: A review. *International Journal of Online and Biomedical Engineering*, *16*(7), 95–104. https://doi.org/10.3991/ijoe.v16i07.14101

Geschwind, G., Vignal, M., Caballero, M. D., & Lewandowski, H. J. (2024). Evidence for validity and reliability of a research-based assessment instrument on measurement uncertainty. *Physical Review Physics Education Research*, *20*(2), 020125. https://doi.org/10.1103/PHYSREVPHYSEDUCRES.20.020125/FIGURES/5/MEDIUM

Gunada, I. W., Wahyudi, W., Ayub, S., Taufik, M., & Busyairi, A. (2023). Validitas

perangkat model project based learning berbasis stem pada pokok bahasan perubahan energi untuk meningkatkan sikap ilmiah. *Empiricism Journal*, *4*(1), 134–144. https://doi.org/10.36312/ej.v4i1.1287

Ješková, Z., Lukáč, S., Šnajder, Ľ., Guniš, J., … D. K.-E., & 2022, undefined. (2022). Active learning in STEM education with regard to the development of inquiry skills. *Education Sciences*.

Kaewmanee, T., Srisenpila, A., & Worapun, W. (2024). Integrating 5E inquiry-based learning and STEM education to enhance grade 5 students' science process skills and achievement in friction. *International Journal of Advanced and Applied Sciences*, *11*(12), 207–214. https://doi.org/10.21833/ijaas.2024.12.023

Kotsis, K. T. (2024). Integrating ChatGPT into the inquiry-based science curriculum for primary education. *European Journal of Education and Pedagogy*, *5*(6), 28–34. https://doi.org/10.24018/EJEDU.2024.5.6.891

Kurnia, M., Yulianti, E., & Pratiwi, N. (2021). Pengembangan instrumen penilaian bermuatan steam untuk mengukur kemampuan berpikir kreatif dan problem solving pada tema getaran dan gelombang. *Jurnal MIPA Dan Pembelajarannya*, *1*(4), 297–303. https://doi.org/10.17977/um067v1i4p297-303

Lafifa, F., Rosana, D., Suyanta, S., Nurohman, S., & Dwi Astuti, S. R. (2023). Integrated STEM approach to improve 21st century skills in indonesia: A systematic review. *International Journal of STEM Education for Sustainability*, *3*(2), 252–267. https://doi.org/10.53889/ijses.v3i2.219

Lestari, S., Uswatun, D. A., Sutisnawati, A. (2025). Pengembangan instrumen penilaian kognitif siswa berbasis stem edp pada mata pelajaran ipa di sekolah dasar. *Didaktika: Jurnal Kependidikan*, *14*(1).

Martín-Páez, T., Aguilera, D., Perales-Palacios, F. J., & Vílchez-González, J. M. (2019). Peran pendekatan stem (science, technology, engineering, and mathematic) dalam pembelajaran. *Journal.Mahesacenter.Org*, *103*(4), 799–822. https://doi.org/10.1002/SCE.21522

Novitra, F., Festiyed, Yohandri, & Asrizal. (2021). Development of online-based inquiry learning model to improve 21st-century skills of physics students in senior high school. *Eurasia Journal of Mathematics, Science and Technology Education*, *17*(9), em2004. https://doi.org/10.29333/EJMSTE/11152

Nuraini, N. (2017). Profil keterampilan berpikir kritis mahasiswa calon guru biologi sebagai upaya mempersiapkan generasi abad 21. *Jurnal Penelitian Pendidikan Biologi*, *1*(2), 89–96.

Nurhaifa, I., Hamdu, G., & Suryana, Y. (2020). Rubrik penilaian kinerja pada pembelajaran stem berbasis keterampilan 4C. *Indonesian Journal of Primary Education*, *4*(1), 101–110. https://doi.org/10.17509/IJPE.V4I1.24742

Papagiannis, P., & Pallaris, G. (2024). *Evaluating 21st century skills development through makerspace workshops in computer science education*. https://arxiv.org/pdf/2411.05012

Pratama, A. T., Limiansi, K., & Anazifa, R. D. (2020). Penggunaan stem (science, technology, engineering, and mathematics) terintegrasi pembelajaran berbasis proyek untuk mahasiswa. *Biosel: Biology Science and Education*, *9*(2), 115. https://doi.org/10.33477/bs.v9i2.1627

Redhana, I. W. (2019). Mengembangkan keterampilan abad ke-21 dalam pembelajaran kimia. *Jurnal Inovasi Pendidikan Kimia*, *13*(1).

Setyowati, Y., Aldila Aribowo, L., Kunci, K., Abad, K., Berpikir Kritis, K., & Berpikir Kreatif, K. (2024). Literature review: Efektifitas PjBL-STEM meningkatkan keterampilan abad 21 dalam pembelajaran IPA di SMP. *Science Education and*

Development Journal, *2*(2), 2024. https://doi.org/10.59923/sendja.v2i2.225
Sinharay, S. (2017). How to compare parametric and nonparametric person-fit statistics using real data. *Journal of Educational Measurement*, *54*(4), 420–439. https://doi.org/10.1111/JEDM.12155;CTYPE:STRING:JOURNAL
Utami, I. I. S. (2018). *Kontribusi model cooperative learning dalam pembelajaran ips terhadap pengembangan 4c 21st century skills peserta didik*. Doctoral dissertation, Universitas Indonesia.
Xiaoming, & Wiebe. (2023). *Classroom-based stem assessment: Contemporary Issues and Perspectives.* https://cadrek12.org/
Yang, K. L., Fang, S. C., & Fan, S. C. (2025). Development and validation of an instrument for assessing secondary students' transdisciplinary STEM practices. *International Journal of STEM Education*, *12*(1), 1–16. https://doi.org/10.1186/s40594-025-00529-3
Yokhebed, Y. (2019). Profil kompetensi abad 21: komunikasi, kreativitas, kolaborasi, berpikir kritis pada calon guru biologi profile of 21st century competency: communication, creativity, collaboration, critical thinking at prospective biology teachers. *Bio-Pedagogi*, *8*(2), 94. https://doi.org/10.20961/bio-pedagogi.v8i2.36154

AUTHOR PROFILES

Isnaini
Isnaini010190@gmail.com
Expertise in Education Science
Affiliation: Universitas Negeri Padang

Atmazaki
atmazaki@fbs.unp.ac.id
Expertise in language learning
Affiliation: Universitas Negeri Padang

Yuni Ahda
ahdayuni@fmipa.unp.ac.id
Expertise in biology learning
Affiliation: Universitas Negeri Padang

Suci Fajrina
sucifajrina@fmipa.unp.ac.id
Expertise in biology learning
Affiliation: Universitas Negeri Padang