International Journal of Evaluation and Research in Education (IJERE)
Vol. 15, No. 6, December 2026, pp. 5625~5638
ISSN: 2252-8822, DOI: 10.11591/ijere.v15i6.40680

Developing a citizenship literacy instrument for biology education students through the Rasch model

Yan Piter Basman Ziraluo[1,2], Sulisetijono[1], Susriyati Mahanal[1]
[1]Department of Biology, Faculty of Mathematics and Natural Sciences, Universitas Negeri Malang, Malang, Indonesia
[2]Department of Biology Education, Faculty of Teacher Training and Education, Universitas Nias Raya, Nias Selatan, Indonesia

Article Info

Article history:

Received May 27, 2026
Revised Sep 30, 2026
Accepted Oct 24, 2026

Keywords:

Biology education
Citizenship literacy
Instrument development
Rasch model
Sustainability education

ABSTRACT

Citizenship literacy (CL) is an important 21st-century competency that integrates scientific understanding, social responsibility, and environmental awareness in biology education. The research method used was research and development (R&D) with stages of construct formulation, item development, expert validation, and testing of 626 students from various universities in Indonesia. Data analysis was performed using Winsteps software version 3.75 to test the validity, reliability, and unidimensionality of the instrument. The results of the analysis showed that the instrument met the criteria for unidimensionality with a raw variance explained by measure of 37.6% and unexplained variance below 15%, indicating that all items measured a single construct consistently. The person reliability value of 0.89 and item reliability of 0.99 indicate high consistency, while the separation index of 2.82 (person) and 8.11 (item) indicates good discrimination. All items met the fit statistics criteria (outfit mean square (MNSQ) 0.5–1.5; Z-standard outfit (ZSTD) −2.0 to +2.0; point measure correlation (PMC) 0.4–0.85), so the instrument was declared valid without revision. These findings confirm that the Rasch model is effective in producing valid and reliable instruments. These instruments can be used to assess and improve the CL of biology education students, while also supporting the development of a curriculum oriented towards sustainability and social responsibility.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Yan Piter Basman Ziraluo
Department of Biology, Faculty of Mathematics and Natural Sciences, Universitas Negeri Malang
Cakrawala Street No. 5, 65145, Malang, East Java, Indonesia
Email: ypbziraluo@uniraya.ac.id

1. INTRODUCTION

Citizenship literacy (CL) has become an increasingly important competency in higher education as societies confront complex socio-scientific issues such as climate change, biodiversity loss, public health crises, and sustainable resource management. These challenges require citizens not only to possess scientific knowledge but also to make informed decisions, participate in citizenship processes, and take responsible actions that contribute to societal well-being. For students in biology education programs, CL is particularly important because they are expected to become future educators who can connect biological knowledge with environmental responsibility, ethical reasoning, and active citizenship. Without adequate CL competencies, prospective biology teachers may struggle to critically evaluate socio-scientific issues, facilitate informed public discourse, and foster sustainability-oriented behaviors among future generations [1].

The urgency of strengthening CL is supported by empirical evidence indicating that university students often demonstrate limited engagement with citizenship issues despite possessing adequate disciplinary knowledge. Studies have reported gaps between scientific understanding and citizenship

participation, particularly in addressing environmental and sustainability challenges. Such limitations are concerning in biology education because graduates are expected to serve as agents of change who promote evidence-based decision-making and environmental stewardship within their communities. Consequently, higher education institutions require valid assessment tools capable of identifying students' strengths and weaknesses in CL so that appropriate educational interventions can be designed and evaluated.

Despite growing scholarly attention to citizenship education, sustainability education, and scientific literacy, several important research gaps remain. First, existing studies generally examine citizenship competencies, environmental attitudes, or scientific literacy as separate constructs, providing limited understanding of how these dimensions interact within the context of biology education [2]. Second, most available instruments emphasize either citizenship knowledge or citizenship engagement, while neglecting the integration of knowledge, attitudes, and practical citizenship skills needed to address contemporary socio-scientific issues [3]. Third, empirical investigations focusing specifically on biology education students remain scarce, resulting in insufficient evidence regarding how prospective biology teachers develop and demonstrate CL competencies. These limitations restrict both theoretical understanding and practical efforts to promote citizenship-oriented science education.

A further gap exists at the methodological level. Many existing assessment instruments rely predominantly on classical test theory approaches, which provide limited information regarding item functioning, measurement precision, and construct calibration [4], [5]. As a result, concerns regarding validity, reliability, and generalizability remain evident across different educational settings. Modern psychometric approaches, particularly the Rasch measurement model, offer substantial advantages by enabling item-level analysis, interval-scale measurement, and rigorous evaluation of construct structure. However, the application of Rasch analysis in developing and validating CL instruments within biology education remains largely unexplored.

Addressing these gaps, the novelty of the present study lies in three key contributions. First, it conceptualizes CL as an integrated construct encompassing scientific citizenship, sustainability values, and leadership dimensions within the specific context of biology education. Second, it develops a context-specific assessment instrument designed explicitly for prospective biology teachers, a population that has received limited attention in previous studies. Third, it employs the Rasch model to provide comprehensive psychometric validation, thereby offering stronger evidence regarding measurement quality and construct representation than many existing instruments [6], [7].

Based on these identified gaps and novel contributions, this study aims to develop and validate a CL assessment instrument for biology education students and to examine its psychometric properties using the Rasch model [8]. Specifically, the study seeks to answer the following research questions:

- What dimensions constitute CL in the context of biology education?
- To what extent does the developed instrument demonstrate validity and reliability according to Rasch measurement criteria?
- How can a validated CL instrument support the development of citizenship education and sustainability-oriented leadership among prospective biology teachers?

2. THEORETICAL FRAMEWORK

The CL refers to an individual's capacity to understand citizenship concepts, critically evaluate public issues, and participate responsibly in social and environmental decision-making processes. Contemporary perspectives conceptualize CL as a multidimensional construct that extends beyond factual knowledge of citizenship to include the ability to apply knowledge, exercise critical judgment, and demonstrate citizenship responsibility in addressing societal challenges. Within higher education, CL is increasingly viewed as an essential competency for preparing students to navigate complex socio-scientific and sustainability-related issues that require informed and ethical decision-making [9]. This perspective aligns with constructivist and democratic education theories, which emphasize the development of reflective, responsible, and participatory citizens capable of contributing to the public good.

In the context of biology education, CL is closely associated with students' ability to connect scientific understanding with social and environmental responsibilities. Biology-related issues such as biodiversity conservation, climate change, public health, and sustainable resource management require citizens who can interpret scientific information, evaluate competing claims, and make decisions based on both evidence and ethical considerations [10]. Consequently, biology education provides a relevant context for fostering CL because it exposes students to authentic socio-scientific issues that demand scientific reasoning alongside citizenship engagement. Through such experiences, students develop an awareness of how biological knowledge can be applied to address challenges affecting communities and ecosystems [11].

The present study conceptualizes CL through three interrelated dimensions: scientific citizenship, sustainability values, and leadership skills. Scientific citizenship refers to an individual's ability to understand and utilize scientific knowledge when evaluating public issues and participating in citizenship decision-making processes [12]. This dimension emphasizes evidence-based reasoning, critical evaluation of information, and informed participation in discussions concerning socio-scientific issues. Sustainability values represent the ethical and attitudinal orientation that guides individuals toward environmental stewardship, social responsibility, and long-term societal well-being [13]. These values influence how students perceive their obligations toward sustainable development and collective welfare. Leadership skills encompass the capacity to influence, collaborate with, and mobilize others in addressing community and environmental challenges [14]. Within citizenship contexts, leadership reflects the practical application of citizenship knowledge and values through active participation and collective action.

The integration of these three dimensions reflects the holistic nature of CL in biology education. Scientific citizenship provides the cognitive foundation for understanding and analyzing issues, sustainability values shape the ethical orientation underlying citizenship decisions, and leadership skills facilitate the translation of knowledge and values into meaningful citizenship action. Together, these dimensions represent a comprehensive framework for assessing CL among biology education students, capturing not only what students know but also how they evaluate issues, internalize citizenship responsibility, and contribute to solving societal and environmental problems.

3. METHOD

This study employed a quantitative approach using an instrument development design based on the Rasch measurement model. The Rasch model was selected because it enables the development of psychometrically sound instruments through the transformation of ordinal data into interval measures, allowing simultaneous estimation of respondent ability and item difficulty on a common linear scale. This approach provides more robust evidence of validity and reliability compared to classical test theory and is widely recommended for educational measurement and instrument development.

The instrument development process was conducted through several systematic stages. First, the construct domain of CL was identified through a systematic review of relevant literature, policy documents, and previous empirical studies [15]. Based on the review findings, key dimensions and indicators of CL were formulated. These indicators were subsequently organized into a construct map describing the continuum of CL competencies expected among biology education students, ranging from basic citizenship awareness to higher-order citizenship reasoning and participation competencies.

Second, an initial item pool was developed according to the identified indicators. Each item was designed to represent specific aspects of the construct and was formulated using a 5-point Likert scale ranging from strongly disagree to strongly agree. Particular attention was given to ensuring content relevance, conceptual clarity, language appropriateness, and variation in item difficulty. The preliminary instrument consisted of items distributed proportionally across all CL dimensions identified during the construct development stage.

Third, content validation was conducted through expert judgment involving three experts with expertise in citizenship education, biology education, educational measurement, and instrument development. The experts evaluated each item based on relevance, representativeness, clarity of wording, and suitability with the intended construct [16]. Suggestions and recommendations provided by the experts were used to revise, refine, or eliminate items that were considered ambiguous, redundant, or insufficiently representative of the construct domain. The content validation process ensured that the instrument adequately reflected the theoretical dimensions of CL before empirical testing in Table 1.

Following expert validation, a pilot study was conducted to evaluate the preliminary psychometric properties of the instrument. The pilot participants consisted of biology education students who met the inclusion criteria of having completed courses related to citizenship education, environmental education, or socio-scientific issues. The pilot data were analyzed using Winsteps software to examine item functioning, item polarity, rating scale effectiveness, item fit statistics, and reliability indices. Items exhibiting misfit characteristics based on outfit mean square (MNSQ), infit MNSQ, Z-standard outfit (ZSTD), and point measure correlation (PMC) were reviewed and revised prior to the main data collection stage.

The main study involved undergraduate students enrolled in biology education programs at Indonesian universities. Participants were selected using purposive sampling. The sampling criteria included: i) being actively enrolled as a biology education student; ii) having completed courses related to citizenship, environmental issues, or socio-scientific decision-making; and iii) voluntarily agreeing to participate in the study. These criteria were established to ensure that respondents possessed sufficient academic experience relevant to the CL construct being measured. Demographic information collected included gender, age, semester level, and institutional affiliation to provide a comprehensive description of respondent characteristics.

Data collection was conducted online using Google Forms to facilitate broad participation and ensure data completeness. Before completing the questionnaire, participants were provided with information regarding the purpose of the study, confidentiality procedures, voluntary participation, and their right to withdraw at any stage without penalty. Only participants who provided informed consent were allowed to proceed to the questionnaire.

The study adhered to established ethical principles for educational research. Participant anonymity was maintained by removing personally identifiable information from the dataset. All collected data were stored securely and used exclusively for research purposes. Participation was entirely voluntary, and respondents were informed that their responses would not affect their academic standing. Ethical approval was obtained from the relevant institutional review board or ethics committee prior to data collection (approval number to be inserted by the authors).

Data analysis was performed using the Rasch model through Winsteps software. Item validity was evaluated using infit and outfit MNSQ statistics, standardized residuals, and PMC to determine the degree of fit between observed responses and model expectations. Reliability analysis included person reliability, item reliability, separation indices, and Cronbach's alpha. Item difficulty estimates were calculated to examine the distribution of items across the CL continuum. Construct validity was further assessed using a Wright Map, which visualized the alignment between respondent ability and item difficulty levels. Items demonstrating substantial misfit or poor measurement performance were considered for revision or removal. Through these procedures, the study aimed to produce a valid, reliable, and theoretically grounded instrument for measuring CL among biology education students.

Table 1. Rasch model testing criteria

Aspect	Indicator/criteria	Acceptable value/range	Interpretation
Unidimensionality	Raw variance explained by measures (principal component analysis residual)	≥20% (minimal), >40% (good), >60% (excellent)	An instrument is considered unidimensional if it meets this minimum value.
	Unexplained variance (by 1st contrast)	<15%	No other dimensions are dominant
Item validity	Outfit MNSQ	0.5<MNSQ<1.5	Items are considered fit within this range.
	ZSTD	-2.0<ZSTD<+2.0	A value close to 0 is ideal; values outside this range are considered unfit.
	PMC	0.4<PMC<0.85	Items are valid/contribute well if they fall within this range.
	Item reliability	≥0.67 (adequate), ≥0.80 (good)	Inter-item consistency: the higher the score, the more consistent the item.
Instrument reliability	Person reliability	0.67 (fair), ≥0.80 (good)	Overall measurement consistency
	Item separation	index≥2.0	The differentiation power is good if it is greater than 2.

4. RESULTS

The development of this CL instrument aims to produce a valid CL measurement instrument. This research was conducted using the research and development (R&D) method through the Rasch model framework. The development of this student CL measurement instrument refers to the theoretical basis of the instrument dimensions determined through a systematic literature review. The research population consisted of biology education students at several universities in Indonesia who had taken or were currently taking courses in animal ecology and/or environmental education.

4.1. Formulation of constructs and derivation of indicators

The formulation of CL constructs was carried out by simplifying the concept into indicators that can be measured clearly and practically. CL is built on three main dimensions, namely knowledge, attitudes and values, and skills. Each dimension is broken down into several operational indicators, which are then developed into statements in the form of a Likert scale so that they are easy for respondents to understand and respond to. This approach aims to ensure that the concept of CL is not only theoretical but can also be measured empirically and systematically in the context of citizenship education research.

The development of the CL instrument identifies these three main dimensions as the foundation of CL. The knowledge dimension includes an understanding of the system, principles, and history of citizenship. In this dimension, the indicators focus on the ability to explain the structure and mechanisms of government, examine the relationship between the rights and obligations of citizens, and assess the process of public policy formulation involving the role of society. In addition, the indicators also emphasize an understanding of the principles of the rule of law, human rights, democracy, and the social and political dynamics that shape the concept of citizenship in Indonesia in Table 2.

Table 2. Indicator formulation based on CL dimensions

Dimension	Sub-dimension	Indicator
Knowledge	System	Explains the structure, working mechanisms, and roles of institutions within the government system. (1)-C2
		Examine the relationship between the rights and obligations of citizens in the state system. (6)-C4
		Evaluate the public policy-making process with the role of society. (7)-C5
	Principles	Connecting the principles of the rule of law, human rights, and democracy as pillars of national life. (3)-C3
		Linking moral principles, ethics, and social responsibility in citizenship. (4)-C4
	History	Provide examples of important figures and their roles in each. (2)-C2
		Analyze the impact of social and political dynamics that influence the evolution of the concept of citizenship in Indonesia. (5)-C4
Attitude and values	Commitment	Prioritizing common interests over personal or group interests (10)
		Demonstrating responsibility in exercising rights and obligations as a citizen. (20)
		Affirming active involvement in maintaining national integrity and unity. (19)
	Tolerance	Respecting differences in views, cultures, ethnicities, and religions in social life. (15)
		Demonstrating an open attitude toward diversity of ideas and social practices. (16)
	Critical	Analyzes citizenship issues with logical and data-based thinking. (17)
		Demonstrates reflective ability regarding social and political practices in their environment. (18)
		Expresses opinions with rational arguments. (8)
Skills	Interaction	Demonstrate the ability to engage in dialogue and negotiation to reach mutual agreements. (11)
		Integrate ethical values and moral responsibility into social interactions. (14)
	Participation	Actively participate in social, political, and environmental activities that strengthen democracy. (9)
		Demonstrate the ability to make decisions based on democratic principles and deliberation. (12)
		Take concrete action in solving social problems as a form of civic responsibility. (13)

The attitude and values dimension are directed at shaping responsible and inclusive citizenship. Indicators in this dimension reflect a commitment to prioritizing the common good, exercising rights and obligations responsibly, and maintaining national unity and integrity. Tolerance is manifested through mutual respect for differences in views, culture, ethnicity, and religion, as well as openness to diversity of ideas. Meanwhile, critical thinking is demonstrated through the ability to analyze citizenship issues logically, reflectively, and based on data, as well as to express opinions with rational arguments.

The skills dimension emphasizes the practical abilities of citizens in social and democratic life. Indicators in this dimension include the skills of interacting through dialogue and negotiation, as well as integrating ethical values and moral responsibility in social relationships. In addition, participation skills are manifested through active involvement in social, political, and environmental activities, the ability to make decisions democratically, and concrete actions in solving social problems. With this structure, the CL instrument is expected to be able to describe CL as a whole, covering cognitive, affective, and psychomotor aspects in a balanced manner.

4.2. Questionnaire item development and review

In the initial stage, the instrument was compiled in the form of 20 statements derived directly from the indicators formulated in the construct development process. Each item was designed to represent the cognitive, affective, and participatory dimensions of citizenship education focused on the environment and wildlife conservation. Before use, the draft instrument first underwent content and logic validation by experts in the fields of education and instrument measurement. This validation process aimed to ensure the suitability of each item with the construct being measured as well as the accuracy, maturity, and clarity of the wording of the statements in Table 3. As a result, only items that were conceptually relevant and free from ambiguity were retained.

The cognitive dimension in this instrument emphasizes students' understanding, analysis, and evaluation of citizenship concepts related to governance, law, public policy, and sustainability. A number of items measure the ability to explain the role of government agencies in wildlife conservation efforts, analyze the relationship between the rights and obligations of citizens, and evaluate public policies, such as the relocation of animals due to infrastructure development. Other items assess the ability to relate the principles of the rule of law, human rights, and democracy to efforts to prevent illegal wildlife trade, as well as analyze the influence of international agreements, such as the convention on biological diversity (CBD), on conservation practices in Indonesia. Overall, this dimension reflects higher-order thinking skills, including analysis, evaluation, and logical reasoning, which are at the core of modern citizenship education.

The affective dimension reflects values, attitudes, and ethical orientations toward citizenship and environmental responsibility. Items in this dimension examine students' willingness to prioritize the common good over personal or group interests, demonstrate responsibility in exercising their rights and obligations as citizens, and respect differences in views, cultures, and social backgrounds. Other statements explore openness to diversity of ideas, reflective attitudes toward social and political practices, and moral

responsibility in supporting conservation initiatives. These affective indicators are important for understanding the extent to which citizenship values are internalized and manifested in pro-environmental attitudes.

Table 3. Questionnaire item formulation

No	Indicator	Questionnaire item	Item number
1	Explain the structure, working mechanisms, and roles of institutions in the government system. (1)-C2	You know that the Ministry of Environment and Forestry has established a new conservation area to protect the orangutan habitat. You are asked to explain the role of this institution in protecting wildlife populations. I can explain the role of the Ministry of Environment and Forestry in preserving the orangutan habitat.	1
2	Examining the relationship between the rights and obligations of citizens in the state system. (6)-C4	You can see changes in the local culture, which now supports the conservation of the Nias parrot and Nias stone magpie habitats. I cannot describe the positive impact of cultural changes on wildlife conservation.	6
3	Evaluating the public policy-making process with the role of the community. (7)-C5	You are asked to evaluate policies on the relocation of wildlife due to infrastructure development in certain areas. I am able to evaluate wildlife relocation policies based on the principle of sustainability.	7
4	Connecting the principles of the rule of law, human rights, and democracy as pillars of national life. (3)-C3	In class discussion, you are asked to explain how the rule of law can help reduce illegal wildlife trade. I am able to explain the relationship between the rule of law and illegal wildlife trade.	3
5	Linking moral principles, ethics, and social responsibility in citizenship. (4)-C4	A local community has successfully preserved sea turtle species with the support of the community and government. I am unable to see how the principle of sustainability can be applied through collaboration between the community and government.	4
6	Provide examples of important figures and their roles in each. (2)-C2	Several global conservation figures have proposed policies to stop deforestation in conservation areas. I am unable to provide examples of global conservation figures.	2
7	Analyzing the impact of social and political dynamics that influence the evolution of the concept of citizenship in Indonesia. (5)-C4	You are asked to describe the impact of signing the Convention on Biological Diversity (CBD) on wildlife protection in Indonesia. I am able to diagnose the impact of international events such as the convention on biological diversity on wildlife conservation in Indonesia.	5
8	Prioritizing common interests over personal or group interests. (10)	In the student council election, one of the candidates proposed an educational program on animal conservation. I voted for the candidate who did not prioritize environmental conservation.	10
9	Demonstrate responsibility in exercising the rights and obligations as a citizen. (20)	You envision a career as a biology teacher who can educate students about the importance of protecting biodiversity. However, I think that it is unnecessary to educate the public about the importance of wildlife conservation.	20
10	Affirming active involvement in maintaining national integrity and unity. (19)	A student organization is holding a campaign about animal protection on your campus. As a student, I want to get involved in this campaign to raise awareness among my friends.	19
11	Respecting differences in views, cultures, ethnicities, and religions in social life. (15)	In group work, your team members have different opinions on how to protect animal habitats. I respect differences of opinion within the team and seek solutions together.	15
12	Demonstrate openness to diverse ideas and social practices. (16)	You attended a seminar discussing the role of women in wildlife conservation. In my opinion, the role of women in environmental preservation and wildlife conservation is not very important.	16
13	Analyzing citizenship issues with logical and data-driven thinking. (17)	You read a social media article that mentions a certain species has become extinct. You are asked to verify the accuracy of this information. I check the validity of information before sharing it with others.	17
14	Demonstrates reflective ability regarding social and political practices in their environment. (18)	You see someone damaging public facilities in a wildlife conservation area. What I do is let the person carry on with the damage because I feel it's none of my business.	18
15	Expressing opinions with rational arguments (8)	You hear about conflicts between local communities and wildlife that often enter residential areas. I am unable to explain the causes of these conflicts rationally.	8
16	Demonstrate the ability to engage in dialogue and negotiation to reach mutual agreements. (11)	You want to propose an environmental education program to the local Forestry Service. I feel capable of communicating with the Forestry Service to support environmental conservation programs.	11
17	Integrating ethical values and moral responsibility into social interactions. (14)	At the student council meeting, you were asked to voice your opinion about the campus conservation project. I am not willing to express my opinion in a discussion that supports environmental preservation.	14
18	Actively participate in social, political, and environmental activities that strengthen democracy. (9)	You have joined a community of students involved in a tree planting project to expand wildlife habitats in urban forests. I am willing to actively participate in activities that support wildlife habitat conservation.	9
19	Demonstrates the ability to make decisions based on democratic principles and deliberation. (12)	A non-governmental organization offers training in wildlife protection advocacy to your team. I am not interested in participating in wildlife protection advocacy training, so I do not need to ask my colleagues for their opinions.	12
20	Implementing concrete actions in solving social problems as a form of civic responsibility. (13)	You discover a local community threatened with the loss of wildlife habitat due to environmental destruction. I feel responsible for helping to protect endangered wildlife habitats.	13

The participatory and skill dimensions emphasize communication skills, decision-making, and concrete citizenship actions. Items in this dimension assess the ability to express opinions rationally, engage in dialogue and negotiation with various stakeholders, and make decisions based on the principles of democracy and deliberation. In addition, several statements measure active participation in social, political, and environmental activities, such as conservation campaigns, advocacy training, and community-based habitat protection efforts. The final items emphasize the implementation of concrete actions to solve social and environmental problems as a form of citizenship responsibility. Overall, this instrument is coherently structured and theory-based to measure citizenship learning that integrates knowledge, values, and actions within a framework of sustainability.

4.3. Instrument trial (pilot test)

The instrument trial was conducted to determine the feasibility, clarity, and preliminary psychometric characteristics of the developed questionnaire items. Before conducting the trial, the questionnaire was first validated through expert judgment involving three experts in biology education and educational assessment. The results of the expert validation indicated that several instrument items required revision to improve their clarity, relevance, and alignment with the measured construct in Table 4.

Table 4. Validator suggestions and input

Suggestions/input	Improvements
Pay attention to the use of operational verbs.	Review and improve operational verbs.
Use statements/questions that are easy to understand	Improve sentence structure and replace words that are difficult to understand.
Complete the questionnaire instructions.	Complete the instrument instructions.

After revision based on expert feedback, the instrument was tested on 626 biology education students from several universities in Indonesia. The participants were recruited using a convenience sampling approach, in which respondents were selected based on accessibility and willingness to participate through online distribution using Google Forms. Although the participants were obtained from multiple universities representing different institutional contexts, the sample was not designed as a nationally representative sample through probability sampling procedures.

Therefore, the findings from this instrument trial primarily provide evidence regarding the feasibility, clarity, and initial validity of the instrument among biology education students within the accessible sample. The use of convenience sampling may limit the extent to which the results can be generalized to all biology education students in Indonesia. Future studies involving larger and more diverse samples using probability-based sampling strategies are recommended to further confirm the reliability, validity, and broader applicability of the developed instrument.

4.3.1. Rasch analysis: item and respondent statistical evaluation

Data analysis was performed using Winsteps software version 3.75 following the principles of the Rasch model. This analysis aimed to evaluate the validity of the items, reliability, and distribution of respondent ability levels on the developed instrument. The results of the unidimensional test in Table 5 show that the observed raw variance measure is 37.6% (≥20%) of the expected value of 39.7%, so it can be concluded that it meets the minimum standard/category in Table 5. Meanwhile, for the unknown variance/unexplained variance, none had a value greater than 15%. This means that the instrument is unidimensional, meaning that there is only one latent attribute underlying the respondents' answers to the test items, or it can be said that the instrument is capable of measuring what it is supposed to measure.

From the data analysis, the results are presented in Table 6. Before conducting the Rasch analysis, the dataset was screened to identify missing responses. In this study, no missing data were found in the response matrix; therefore, all participants' responses were included in the analysis without imputation. This procedure was applied to maintain the original response patterns and avoid potential bias caused by estimating or replacing missing values.

Table 5. Dimensionality test results

Raw variance	Eigenvalue	Observed	Expected	Description
Raw variance explained by measure	12.0	37.6	39.7	Good
Unexplained variance in 1st contrast	2.8	9.8	14.0	Good
Unexplained variance in 2nd contrast	2.0	5.1	10.0	Good
Unexplained variance in 3rd contrast	1.4	4.4	7.1	Good
Unexplained variance in 4th contrast	1.2	3.9	6.2	Good
Unexplained variance in 5th contrast	1.2	3.9	6.2	Good

The Rasch model analysis showed that the person reliability value was 0.89, indicating good consistency of respondents in responding to the items, while the item reliability value was 0.99, indicating excellent stability and reproducibility of the item hierarchy. The item fit analysis showed an outfit MNSQ value of 1.02 and an outfit ZSTD value of 0.00, indicating that all items met the acceptable fit criteria. The selection of MNSQ and ZSTD as fit criteria followed Rasch measurement guidelines, where MNSQ values close to 1.00 indicate that the observed data fit the expectations of the Rasch model, while ZSTD values around 0 represent an acceptable deviation between observed and expected responses. Therefore, these criteria were used to evaluate whether the items functioned consistently with the measurement model. Furthermore, the PMC values for all items were within the acceptable range (0.4<PMC<0.85), indicating that the items appropriately discriminated between respondents with different ability levels. This suggests that respondents with higher ability levels tended to have a greater probability of answering the items correctly, as seen in Table 6.

Table 6. Data analysis results with Winsteps

Criteria	Parameter (N)	
	Person	Item
Mean measure	1.91	0.00
Reliability	0.89	0.99
Separation index	2.82	8.11
Outfit MNSQ (mean)	1.02	1.02
Outfit ZSTD (mean)	-0.50	0.00

Statistical fit analysis: the infit and outfit MNSQ values for each item and person were confirmed to be within the range of 0.5-1.5, ensuring the accuracy of the model. All values in this study met these criteria in Table 7. Ability distribution and difficulty level: the average measure of ability is 1.91 logit (SD=1.89), while the mean item is 0.00 logit (SD=0.65), indicating a balanced distribution of items and abilities. High reliability: the reliability of the person is 0.89-very good for distinguishing the ability levels of students, and the reliability of the item is 0.99, indicating exceptional stability. Separation index: the separation test shows that the respondent separation index is 2.82, while the item separation is 8.11, meaning that the quality of the respondents is quite good, while the quality of the items is excellent.

Table 7. Results of item-by-item fit analysis

Total score	Item measure (logit)	Infit MNSQ	Outfit MNSQ	Status fit
2428	0.35	0.92	0.99	Fit
2236	1.19	1.13	1.30	Fit
2439	0.30	0.88	0.83	Fit
2450	0.24	0.72	0.66	Fit
2370	0.64	0.66	0.68	Fit
2457	0.20	0.78	0.75	Fit
2353	0.72	0.78	0.83	Fit
2494	-0.01	0.86	0.79	Fit
2551	-0.37	1.11	1.04	Fit
2575	-0.52	1.17	1.11	Fit
2392	0.53	0.87	0.86	Fit
2271	1.06	1.50	2.07	Fit
2532	-0.24	1.02	0.93	Fit
2515	-0.14	0.90	0.75	Fit
2668	-1.17	1.11	1.10	Fit
2555	-0.39	1.09	1.04	Fit
2571	-0.50	1.01	0.95	Fit
2480	0.07	1.48	1.48	Fit
2582	-0.57	1.18	0.96	Fit
2696	-1.38	1.45	1.37	Fit

4.3.2. Wright Map instrument

Based on the Rasch analysis results, no items with problematic statistical fits were found, meaning that all items met the psychometric standards. Therefore, no instrument revisions were necessary. The measurement results were visualized using Wright Map to assess the fit between the distribution of respondents' abilities and the difficulty level of the items. The interpretation of the results was based on the logit distribution and the ability categories that were formed.

This Wright Map simultaneously displays the distribution of respondents' abilities and the level of difficulty of items in the CL measurement instrument based on the Rasch model. On the map, the vertical axis represents the logit scale, which serves as an interval measure for ability and item difficulty. This logit scale allows for valid and standardized comparisons between individual abilities and the difficulty of the items tested.

The distribution of respondents' abilities is shown on the left side of the map with blue dots that reflect the variation in students' CL abilities. The range of respondents' abilities can be seen from the logit values, which vary from low to high, indicating the diversity of literacy levels in the research sample. On the right side of the map, the distribution of item difficulty levels is marked with red symbols accompanied by number labels that refer to the identity of each item. The vertical position of each item point indicates its relative difficulty level on the logit scale, with high logit values indicating more difficult items and low logit values indicating items that are relatively easy for respondents to answer.

Figure 1 shows that the distribution of respondents' abilities (person measures) has a fairly wide range, indicating variation in CL abilities among participants. This means that the research sample includes respondents with varying levels of ability, from low to high. The questionnaire items also show varying levels of difficulty, ranging from relatively easy to quite difficult. This variation in difficulty is important for comprehensively measuring the spectrum of respondents' abilities.

The majority of respondents' abilities are within a range that is balanced with the difficulty level of the items, indicating that this CL questionnaire instrument is appropriate for use with this group of respondents and is capable of accurately measuring their abilities. The balance between respondents' abilities and item difficulty supports the construct validity of the instrument for CL research. If extreme differences are found between the respondents' ability levels and the difficulty of the items, for example, the majority of respondents have very high abilities while the items are relatively easy, or vice versa, this indicates the need to revise the measurement instrument. However, based on this Wright Map, this indication is not dominant. Thus, this Wright Map indicates that the CL questionnaire used is effective in measuring respondents' abilities at various levels of item difficulty, thereby supporting the validity and reliability of the instrument in the context of this study.



Figure 1. Wright Map from the CL questionnaire

5. DISCUSSION

This study employed a quantitative instrument-development design using the Rasch model as the primary psychometric framework. The Rasch approach was selected because it enables the transformation of ordinal responses into interval-level measurements while simultaneously estimating item difficulty and respondent ability on a common linear scale [17], [18]. This capability is particularly relevant for measuring CL, a complex construct that encompasses cognitive, affective, and behavioral dimensions. Previous studies on CL have predominantly focused on citizenship knowledge or political participation and have often relied on classical test theory, which provides limited information regarding item functioning and measurement

precision across different respondent ability levels [19]. Consequently, a significant gap remains in the availability of psychometrically robust instruments capable of capturing the multidimensional nature of CL among biology education students.

The instrument development process began with a systematic review of the literature to identify the theoretical dimensions of CL relevant to biology education. Three major dimensions were established: knowledge (citizenship systems, democratic principles, and citizenship history), attitudes and values (commitment, tolerance, responsibility, and critical thinking), and skills (social interaction, communication, and active participation) [20]. These dimensions were translated into Likert-scale statements and subsequently evaluated by experts in educational assessment and psychometrics to ensure content relevance, construct representation, and linguistic clarity. This stage is particularly important because previous CL instruments have frequently been criticized for emphasizing cognitive aspects while inadequately representing value-oriented and participatory dimensions, thereby limiting their usefulness in assessing holistic citizenship competencies [19], [20].

The psychometric evaluation demonstrated that the instrument fulfilled the assumptions required by the Rasch model. The explained raw variance of 37.6% exceeded the minimum criterion of 20%, indicating satisfactory unidimensionality and confirming that the instrument measures a coherent latent construct [17], [21]. Although CL is conceptually multidimensional, the Rasch results suggest that the dimensions function collectively as an integrated competency. This finding supports theoretical perspectives that view CL as a unified educational outcome rather than a collection of isolated attributes. Furthermore, the person reliability coefficient of 0.89 indicates strong consistency in respondents' answer patterns, whereas the item reliability coefficient of 0.99 demonstrates excellent item hierarchy stability [17]. Compared with previous CL studies reporting moderate reliability levels and limited item discrimination, these findings indicate that the present instrument possesses superior measurement precision and greater capacity to differentiate respondents across varying competency levels.

The item fit statistics further reinforce the quality of the instrument. All items met the acceptable Rasch fit criteria for outfit MNSQ and standardized residuals (ZSTD), indicating that each statement contributed meaningfully to measuring the intended construct [17]. Earlier studies on citizenship-related assessments have frequently reported misfitting items caused by ambiguous wording, overlapping indicators, or inadequate construct representation. The absence of problematic items in the current study suggests that the theoretical framework and expert validation process successfully translated the CL construct into empirically functioning indicators. This finding contributes to addressing methodological concerns identified in previous research regarding the alignment between theoretical definitions of CL and their operational measurement [20], [21].

The Wright Map analysis provides additional evidence regarding the instrument's effectiveness. The broad distribution of respondent abilities reveals substantial variation in CL among biology education students, suggesting that CL development remains uneven across higher education contexts. At the same time, the distribution of item difficulties covers a wide range from relatively easy to highly challenging items, enabling the instrument to discriminate effectively among students with different competency levels [17], [21]. This finding is particularly important because several previous instruments exhibited ceiling effects or limited item difficulty variation, reducing their sensitivity to differences in respondent performance. The alignment between item difficulty and respondent ability observed in this study indicates that the instrument is appropriately targeted to the intended population and capable of generating meaningful diagnostic information for educational improvement.

The variation in item difficulty also provides insight into the specific challenges faced by biology education students in developing CL. Based on expert discussions during the validation stage, items related to active participation, environmental decision-making, and the application of citizenship principles in real-life contexts tended to represent higher-level competencies because they require students to integrate scientific understanding with social responsibility and ethical reasoning. In contrast, items associated with basic citizenship concepts and awareness were relatively easier because these aspects are more frequently encountered through formal education and general social experiences. This pattern suggests that biology education students may possess foundational citizenship knowledge but require further learning opportunities to strengthen practical engagement and action-oriented citizenship competencies, particularly in addressing environmental and socio-scientific issues.

The results also provide important educational implications. First, the study addresses a critical gap in biology education research by providing a validated instrument that measures CL beyond factual knowledge and incorporates values, attitudes, and participatory competencies associated with sustainability-oriented citizenship [20], [22]. This contribution is significant because contemporary biology education increasingly emphasizes socio-scientific issues, environmental responsibility, and democratic participation, all of which require broader citizenship competencies [22]. Second, the instrument offers educators and

curriculum developers a practical tool for evaluating the effectiveness of instructional interventions designed to foster citizenship engagement and responsible citizenship [23]. Rather than merely assessing content mastery, the instrument enables the identification of strengths and weaknesses across multiple dimensions of CL, thereby supporting evidence-based curriculum refinement [21].

Practically, the Wright Map can serve as a diagnostic reference for biology education lecturers in designing instructional priorities. For example, if students demonstrate high proficiency in citizenship knowledge but lower proficiency in environmental responsibility or participatory skills, learning activities can be adjusted by emphasizing socio-scientific issues, environmental problem-solving projects, community-based activities, and reflective discussions. Therefore, the instrument does not only provide an overall CL score but also helps educators identify which competency dimensions require greater instructional attention.

Third, the findings contribute to the growing body of literature advocating the integration of citizenship education within disciplinary learning contexts. Previous studies have highlighted the importance of contextualized CL assessment but have provided limited empirical evidence regarding valid measurement models in science education settings. The present study demonstrates that CL can be measured reliably within biology education using a Rasch-based framework, thereby extending existing theoretical and methodological discussions [20], [24]. In this regard, the study not only validates a measurement instrument but also strengthens the conceptual argument that citizenship development should be considered an integral outcome of biology education [22].

Despite these contributions, several limitations should be acknowledged. The instrument was validated primarily with biology education students in Indonesia; therefore, the generalizability of the findings to other academic disciplines, educational levels, or cultural contexts remains uncertain. Previous cross-cultural assessment studies have shown that item functioning may vary according to sociocultural and institutional factors, highlighting the need for further validation studies and differential item functioning analyses before broader implementation [25]. Another limitation concerns the use of self-reported Likert-scale responses to measure CL, particularly the skill and action-oriented dimensions. Although self-reported instruments are valuable for capturing students' perceptions, attitudes, and intended behaviors, they may not fully represent actual citizenship practices in real-life situations. Therefore, future studies should consider combining this instrument with complementary approaches, such as behavioral observations, performance-based assessments, or participation records, to obtain a more comprehensive understanding of students' citizenship competencies. Future research should therefore examine the instrument's applicability across diverse populations and investigate its predictive value for educational outcomes such as citizenship engagement, environmental citizenship, and participation in sustainability initiatives.

Overall, the findings demonstrate that a comprehensive, valid, and reliable CL instrument can be developed through the Rasch model [17], [18], [21]. More importantly, the study fills a notable gap in previous research by providing a multidimensional assessment framework that integrates knowledge, values, attitudes, and participatory skills within biology education. Consequently, the instrument offers both a theoretical contribution to the conceptualization of CL and a practical contribution to the assessment of citizenship competencies in higher education, particularly within the context of sustainability-oriented learning and character development [20], [22]. Beyond its role as an assessment instrument, this tool serves as a bridge connecting environmental learning, scientific literacy, and broader citizenship competencies. By integrating knowledge, values, attitudes, and participatory dimensions, the instrument provides a more comprehensive framework for evaluating students' readiness to engage with sustainability challenges and societal issues.

6. CONCLUSION

This study concludes that the Rasch model-based CL instrument developed for biology education students demonstrates satisfactory psychometric properties and can serve as a valid and reliable tool for measuring multidimensional CL competencies. Evidence from the Rasch analysis indicates adequate unidimensionality, strong person and item reliability, and an appropriate distribution of item difficulty levels. These findings suggest that the instrument is capable of distinguishing variations in students' CL abilities and generating accurate measurement outcomes. Consequently, the instrument provides a robust empirical framework for assessing key dimensions of CL within the context of higher education, particularly in teacher education programs.

The main contribution of this study lies in the provision of a rigorously validated assessment instrument that can support evidence-based educational decision-making. From a pedagogical perspective, the instrument can be utilized to identify students' strengths and weaknesses in CL and to inform the design of learning activities that promote critical thinking, ethical reasoning, social responsibility, and active citizenship. At the curriculum level, the instrument offers a practical means for monitoring the achievement of citizenship learning outcomes and evaluating the effectiveness of curriculum implementation.

Furthermore, in the context of educational evaluation, the instrument can support formative and summative assessment processes by providing objective and psychometrically sound data for measuring CL development among university students.

Despite these contributions, several limitations should be acknowledged. First, the validation process was conducted within a relatively limited sample of biology education students, which may restrict the generalizability of the findings to other academic disciplines and institutional contexts. Second, the study focused primarily on internal psychometric validation through Rasch analysis and did not examine external validity indicators, such as predictive validity or relationships with other established measures of citizenship competence. Future research is therefore recommended to conduct cross-institutional and cross-cultural validation studies involving larger and more diverse samples. Further studies may also integrate confirmatory factor analysis, investigate measurement invariance across demographic groups, and explore the relationship between CL and educational outcomes such as citizenship engagement, environmental citizenship, critical thinking, and democratic participation. Such efforts will contribute to strengthening the applicability and generalizability of the instrument while advancing the assessment of CL in higher education.

ACKNOWLEDGMENTS

The author would like to express profound gratitude to all biology education lecturers across Indonesia for their support, permission, and encouragement that made this research possible. Special thanks are also extended to all biology education students from various universities throughout Indonesia who voluntarily participated as respondents.

FUNDING INFORMATION

We thank Lembaga Pengelola Dana Pendidikan (LPDP), Ministry of Finance of the Republic of Indonesia, for supporting our research under Grant Number 20230621198726.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Yan Piter Basman Ziraluo	✓	✓	✓	✓	✓	✓		✓	✓	✓			✓	
Sulisetijono		✓				✓		✓	✓	✓	✓	✓		
Susriyati Mahanal	✓		✓	✓			✓			✓	✓		✓	✓

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

The author declares that there is no conflict of interest regarding the conduct of this research, the authorship, or the publication of this article.

INFORMED CONSENT

We have obtained informed consent from all individuals included in this study.

ETHICAL APPROVAL

The research related to human use has been complied with all the relevant national regulations and institutional policies in accordance with the tenets of the Helsinki Declaration and has been approved by the authors' institutional review board or equivalent committee.

DATA AVAILABILITY

Derived data supporting the findings of this study are available from the corresponding author [YPBZ] on request.

REFERENCES

[1] S. Jorgenson, "De-centering and re-visioning global citizenship education abroad programs," *International Journal of Development Education and Global Learning*, vol. 3, no. 1, pp. 23–38, 2010, doi: 10.18546/ijdegl.03.1.03.

[2] I. A. Pratiwi, T. Lutfiyah, and M. A. Fardani, "Integrating the local wisdom of Bulusan Kudus into A P5 E-module: a feasibility and practicality study for elementary school students," *Jurnal Kependidikan: Jurnal Hasil Penelitian dan Kajian Kepustakaan di Bidang Pendidikan, Pengajaran dan Pembelajaran*, vol. 11, no. 2, pp. 621–632, Jun. 2025, doi: 10.33394/jk.v11i2.15256.

[3] C. Nicklin and J. P. Vitta, "Assessing Rasch measurement estimation methods across R packages with yes/no vocabulary test data," *Language Testing*, vol. 39, no. 4, pp. 513–540, Oct. 2022, doi: 10.1177/02655322211066822.

[4] A. Rasmussen, "Imagining and building the nation through citizenship education: an interculturalist perspective on the case of Denmark," *Language and Intercultural Communication*, vol. 23, no. 3, pp. 268–279, 2023, doi: 10.1080/14708477.2023.2187058.

[5] S. A. Fleary and M. Millington, "Validation of the assessments of adult health literacy: a Rasch measurement model approach," *Health Promotion International*, vol. 40, no. 3, pp. 1–12, 2025, doi: 10.1093/heapro/daaf090.

[6] P. Susongko, Y. Arfiani, and M. Kusuma, "Determination of gender differential item functioning in Tegal-students' scientific literacy skills with integrated science (SLiSIS) test using Rasch model," *Jurnal Pendidikan IPA Indonesia*, vol. 10, no. 2, pp. 270–281, Jun. 2021, doi: 10.15294/jpii.v10i2.26775.

[7] M. Özden, "Science education for citizenship: a case study," *Anadolu Journal of Educational Sciences International*, vol. 10, no. 1, pp. 150–188, Jan. 2020, doi: 10.18039/ajesi.682020.

[8] A. E. Poropat and L. Jones, "Development and validation of a unifactorial measure of citizenship performance," *Journal of Occupational and Organizational Psychology*, vol. 82, no. 4, pp. 851–869, Sep. 2009, doi: 10.1348/096317908X374139.

[9] M. Sánchez-Mendiola *et al.*, "Large-scale diagnostic assessment in first-year university students: pre- and transpandemic comparison," *Educational Assessment, Evaluation and Accountability*, vol. 35, no. 4, pp. 503–523, Nov. 2023, doi: 10.1007/s11092-023-09410-9.

[10] A. W. Clemmons, D. A. Donovan, E. J. Theobald, and A. J. Crowe, "Using the intended–enacted–experienced curriculum model to map the vision and change core competencies in undergraduate biology programs and courses," *CBE Life Sciences Education*, vol. 21, no. 1, pp. 1–18, Mar. 2022, doi: 10.1187/CBE.21-02-0054.

[11] M. Ayar and D. Özalp, "Analyzing a human and environment unit at the 5th grade science curriculum within the environmentally lliterate citizenship context," *Academy Journal of Educational Sciences*, vol. 5, no. 1, pp. 1–14, Jun. 2021, doi: 10.31805/acjes.839228.

[12] J. Weinberg and M. Flinders, "Learning for democracy: the politics and practice of citizenship education," *British Educational Research Journal*, vol. 44, no. 4, pp. 573–592, Aug. 2018, doi: 10.1002/berj.3446.

[13] E. Bechraki, E. Mavrikaki, V. Gialamas, and E. Galanaki, "Development and validation of an instrument for the health literacy assessment of secondary school students (HeLiASeSS)," *Health Education*, vol. 122, no. 6, pp. 678–699, Nov. 2022, doi: 10.1108/HE-08-2021-0111.

[14] T. L. Cary, C. J. Wienhold, and J. Branchaw, "A biology core concept instrument (BCCI) to teach and assess student conceptual understanding," *CBE Life Sciences Education*, vol. 18, no. 3, pp. 1–17, Sep. 2019, doi: 10.1187/cbe.18-09-0192.

[15] M. B. Atoy, F. R. O. Garcia, R. R. Cadungog, J. D. O. Cua, S. C. Mangunay, and A. B. de Guzman, "Linking digital literacy and online information searching strategies of Philippine university students: the moderating role of mindfulness," *Journal of Librarianship and Information Science*, vol. 52, no. 4, pp. 1015–1027, Dec. 2020, doi: 10.1177/0961000619898213.

[16] O. Paakkari, M. Torppa, L. Kannas, and L. Paakkari, "Subjective health literacy: development of a brief instrument for school-aged children," *Scandinavian Journal of Public Health*, vol. 44, no. 8, pp. 751–757, Dec. 2016, doi: 10.1177/1403494816669639.

[17] M. V. Palka, B. Clarkston, and B. Steinwand, "The state of introductory biology following 16 years of pedagogical reform since vision and change for biology undergraduate education," *Journal of Microbiology & Biology Education*, vol. 26, no. 3, pp. 1–11, Dec. 2025, doi: 10.1128/jmbe.00160-25.

[18] I. R. Lukman, M. Mellyzar, S. Alvina, and N. Saa'dah, "Development of a chemical literacy assessment on Colloid (CLAC) instrument to measure chemical literacy," in *Proceedings of Malikussaleh International Conference on Multidisciplinary Studies (MICoMS)*, pp. 88–95, Dec. 2022, doi: 10.29103/micoms.v3i.50.

[19] P. Apriliyani, E. Susantini, and Y. Yuliani, "Validity of science literacy on the respiratory system in Indonesia's Merdeka Curriculum," *IJORER: International Journal of Recent Educational Research*, vol. 4, no. 2, pp. 163–175, Mar. 2023, doi: 10.46245/ijorer.v4i2.297.

[20] W. J. Boone, "Rasch analysis for instrument development: why, when, and how?" *CBE Life Sciences Education*, vol. 15, no. 4, pp. 1–7, Dec. 2016, doi: 10.1187/cbe.16-04-0148.

[21] V. Aryadoust, H. A. H. Tan, and L. Y. Ng, "A scientometric review of Rasch measurement: the rise and progress of a specialty," *Frontiers in Psychology*, vol. 10, p. 2197, Oct. 2019, doi: 10.3389/fpsyg.2019.02197.

[22] D. P. Ratri, S. Rachmajanti, M. Anugerahwati, E. D. Laksmi, and A. Gozali, "Fostering cultural competence: developing an English syllabus for young learners in the Indonesian EFL context with emphasis on local culture to maintain students' identity," *Cogent Education*, vol. 12, no. 1, pp. 1–21, Dec. 2025, doi: 10.1080/2331186X.2024.2440177.

[23] R. R. Isdendi, I. Nanda, and M. W. Suganda, "The role of Indonesia's digital citizenship generation in the realization of ethics in cyberspace," in *4th Annual Civic Education Conference (ACEC 2022)*, 2023, pp. 514–525, doi: 10.2991/978-2-38476-096-1_57.

[24] S. M. Habibah and Sapriyah, "Global citizenship education; a framework of citizenship literacy in Indonesia," in *Proceedings of the International Joint Conference on Arts and Humanities 2021 (IJCAH 2021)*, 2021, pp. 625–628, doi: 10.2991/assehr.k.211223.106.

[25] P. A. Edelsbrunner and F. Dablander, "The psychometric modeling of scientific reasoning: a review and recommendations for future avenues," *Educational Psychology Review*, vol. 31, no. 1, pp. 1–34, Mar. 2019, doi: 10.1007/s10648-018-9455-5.

BIOGRAPHIES OF AUTHORS



Yan Piter Basman Ziraluo is a lecturer and researcher in the field of Biology Education at Nias Raya University. He completed his undergraduate studies at IKIP Gunungsitoli and pursued graduate studies in Educational Management and Biology Education at HKBP Nommensen University and Padang State University. His academic interests focus on biology education, environmental education, literacy, and innovative learning models in science education. He has contributed to various scientific publications and educational books related to biology learning and instructional development. He can be contacted at email: ypbziraluo@uniraya.ac.id.



Sulisetijono is a senior lecturer in the Department of Biology at the Faculty of Mathematics and Natural Sciences, State University of Malang. He earned his bachelor's degree from IKIP Malang and went on to pursue a master's degree in Biology at the Bandung Institute of Technology. His areas of expertise include plant development, biological statistics, experimental design, and biology education research. He is active in research and scientific publications related to botany, biology education, research methodology, and the development of critical thinking skills in higher education. He can be contacted at email: sulisetijono.fmipa@um.ac.id.



Susriyati Mahanal is a professor of Biology Education at Malang State University. She completed her bachelor's and master's degrees in Biology Education at IKIP Malang and earned her Ph.D. in Biology Education from Malang State University. She was appointed as a professor in 2019. Her research interests include biology education, science literacy, inquiry-based learning, higher-order thinking skills, metacognition, and research-based learning. She can be contacted at email: susriyati.mahanal.fmipa@um.ac.id.