ISEJ Indonesian Science Education Journal

ISSN Online: 2716-3350
https://siducat.org/index.php/isej/
Published by Yayasan Darussalam Bengkulu

OPEN ACCESS ARTICLE

Validity of an Assessment Instrument for Measuring Students' Eco-Awareness of Sidoarjo's Traditional Aquaculture Pond Ecosystem

Wigati, Endy Sri[1*], Wisanti[2], Raharjo[3].
Biology Education, Universitas Negeri Surabaya, Jawa Timur, Indonesia[1,2,3]
E-mail Corresponding: endywigati@gmail.com

Received: March 17, 2028	Revised: May 13, 2026	Accepted: July 12, 2026

Abstract
This study aimed to develop and obtain preliminary validation evidence for an essay-based instrument assessing senior high school students' eco-awareness ecological knowledge, attitudes, and behavior through the local context of Sidoarjo's traditional aquatic pond ecosystem. Using a research and development design adapted from the MEASURE approach, the study included blueprint preparation, item construction, expert validation, revision, pilot testing, and psychometric analysis. The initial instrument comprised seven essay items reviewed by four validators: two university lecturers and two senior high school biology teachers. Pilot testing involved 30 Grade 11 students who had studied ecosystem concepts and were excluded from the main sample. Content validity was evaluated through expert-rating percentages, empirical validity through Pearson item–total correlations, and internal consistency through Cronbach's alpha. Expert-validation scores ranged from 97.92% to 100%, indicating very high content relevance, construction quality, and linguistic clarity. Six items showed significant item–total correlations ranging from .484 to .774, while one item failed to meet the empirical-validity criterion (r = .092, p = .629) and required revision or removal. The six retained items yielded a Cronbach's alpha of .656, indicating acceptable preliminary consistency for formative use. These findings suggest that the instrument can support contextual assessment of students' environmental reasoning and assist biology teachers in diagnosing eco-awareness through locally grounded ecological learning. However, the small pilot sample and limited item pool restrict generalizability; further testing with larger, more diverse samples and broader validity evidence is required before high-stakes use.

Keywords: Eco-Awareness; Instrument Development; Reliability; Traditional Aquatic Pond Ecosystem; Validity

INTRODUCTION

The increasing frequency and severity of ecological disturbances have intensified the need to develop environmentally responsible citizens who can understand ecological systems, evaluate the consequences of human activities, and participate in sustainable environmental practices. Contemporary environmental challenges are not solely biophysical problems but also reflect limitations in people's knowledge, attitudes, values, and behaviors toward the environment. Consequently, environmental education has become essential for strengthening learners' capacity to interpret human–environment interactions and make informed ecological decisions. Within science education, biology has a particularly strategic role because it enables students to examine ecological relationships, biodiversity, environmental degradation, and sustainability through both scientific concepts and authentic environmental phenomena (Jeronen & Jeronen, 2026; Kang & Teig, 2025; Zebua et al., 2025). Integrating environmental issues into biology instruction can therefore support the development of students who are not only knowledgeable about ecosystems but also willing and able to demonstrate responsible environmental behavior (AlAli et al., 2025; Sarwi et al., 2025; Zebua et al., 2025).

One important educational construct related to environmental responsibility is eco-

awareness. Eco-awareness refers to an individual's capacity to recognize ecological conditions, understand reciprocal relationships between humans and the environment, evaluate the environmental consequences of personal and collective actions, and respond through responsible decisions and practices. It is a multidimensional construct that cannot be adequately represented by factual knowledge alone. Instead, it encompasses ecological knowledge, ecological attitudes, and ecological behavior as interconnected dimensions of environmental competence (Begum et al., 2025; Burgos-Espinoza et al., 2025; Lin et al., 2025). Ecological knowledge provides the conceptual basis for understanding ecosystem processes and environmental problems, whereas ecological attitudes reflect concern, responsibility, and value orientations toward environmental sustainability. Ecological behavior represents the enactment of such knowledge and attitudes through observable choices and actions. These dimensions should be examined together because students may possess favorable environmental attitudes without demonstrating adequate ecological understanding or translating their concern into consistent behavior (Benzehaf et al., 2025; Chen & Erfanian, 2026; Rahmania, 2024).

Empirical evidence indicates that the relationships among environmental knowledge, attitudes, awareness, and behavior are not always linear. Cahyani et al. (2025), for example, found that senior high school students demonstrated stronger environmental attitudes and behaviors than ecological knowledge, indicating an imbalance among the dimensions of environmental literacy. Similarly, Wibowo et al. (2023) reported that students displayed positive environmental attitudes and awareness, while their ecological behavior remained at a more moderate level. These findings suggest that positive perceptions of the environment do not automatically lead to scientifically informed or sustained environmental action. Environmental competence must therefore be deliberately developed through learning activities that connect ecological concepts with problems encountered in students' everyday environments (AlAli et al., 2025; Ardoin & Heimlich, 2021; Hsu, 2025). Contextual and locally grounded science learning can make abstract ecological principles more meaningful because students are encouraged to examine environmental phenomena that are geographically, culturally, and economically relevant to their communities (Al-Barakat et al., 2025; Delgado-Guerrero et al., 2025; Sotero et al., 2020).

The traditional aquaculture pond ecosystem of Sidoarjo provides a relevant context for developing and assessing students' eco-awareness. Traditional ponds represent dynamic socioecological systems in which water quality, aquatic organisms, human livelihoods, land use, waste management, and ecosystem sustainability are closely interconnected. Learning through this context allows students to analyze ecological processes while considering the consequences of human intervention in aquatic environments. Local environmental contexts can also promote deeper engagement because students are required to relate biological knowledge to actual environmental conditions rather than merely recalling decontextualized concepts (Ardoin & Heimlich, 2021; Jeronen & Jeronen, 2026; Xiong et al., 2025). Furthermore, locally situated environmental learning can encourage students to evaluate environmental decisions, formulate evidence-based arguments, and propose realistic solutions to ecological problems. Thus, the traditional pond ecosystem is not only a subject of biological study but also a meaningful setting for assessing how students understand, value, and respond to environmental sustainability.

However, the successful integration of eco-awareness into biology education depends on the availability of an appropriate assessment instrument. Assessment provides evidence regarding whether students have achieved the intended learning outcomes and whether educational interventions have influenced their knowledge, attitudes, and behavioral reasoning. An assessment instrument must be aligned with learning objectives, the characteristics of the construct being measured, the contextual content presented to students, and the cognitive demands of each item (Albu & Lindmeier, 2023; Levy-Feldman, 2025; Xin & Lo, 2025). In the assessment of eco-awareness, instruments should move beyond measuring factual recall and provide students with opportunities to interpret ecological evidence, evaluate environmental decisions, express attitudes, and justify possible courses of action. Essay-based items may be

particularly useful because they allow students to articulate their reasoning and reveal the connections among ecological knowledge, environmental judgment, and intended behavior. Nevertheless, such instruments can only produce meaningful evidence when their content relevance, empirical functioning, and scoring consistency have been systematically examined (Cruchinho et al., 2024; Mikkonen et al., 2022; Taber, 2018).

Instrument validity and reliability are therefore fundamental requirements in educational assessment. Content validation is needed to determine whether the items adequately represent the intended construct, learning content, item construction principles, and linguistic clarity. Empirical validation is required to examine whether individual items function consistently with the overall construct being measured, whereas reliability analysis provides evidence regarding the consistency of the scores generated by the instrument. These procedures should not be treated as merely statistical formalities because inadequately validated instruments may lead to inaccurate interpretations of students' abilities and inappropriate educational decisions (Boateng et al., 2018; Cook & Hatala, 2016). A systematic instrument-development framework, such as the MEASURE approach, offers an operational sequence that includes construct conceptualization, blueprint development, item construction, expert review, pilot testing, validity analysis, reliability estimation, and instrument refinement (Ullah et al., 2026; Valdivia-Salazar et al., 2025).

Although previous studies have examined environmental literacy, environmental attitudes, ecological behavior, and context-based environmental learning, several limitations remain. Existing studies have predominantly described students' general levels of environmental awareness or investigated relationships among knowledge, attitudes, and behavior, rather than developing assessment instruments specifically grounded in a local aquatic ecosystem (Benzehaf et al., 2025; Lye et al., 2024; Vieriu & Petrea, 2025). Other instrument-development studies have focused on critical thinking, general learning outcomes, or broad domains of knowledge, attitudes, and skills without specifically integrating ecological knowledge, ecological attitudes, and ecological behavior into an essay-based assessment situated in the traditional pond ecosystem of Sidoarjo (Musah & Wangila, 2024; A. M. Wibowo et al., 2024). Therefore, there remains a need for a context-specific instrument whose content relevance, empirical validity, and reliability are systematically evaluated. The novelty of the present study lies in the development and validation of an essay-based eco-awareness assessment instrument that integrates three dimensions ecological knowledge, ecological attitudes, and ecological behavior within the authentic context of Sidoarjo's traditional aquatic pond ecosystem. Accordingly, this study aims to determine the content validity, empirical validity, and reliability of an assessment instrument designed to measure the eco-awareness of Grade 10 senior high school students. Specifically, the study addresses whether the developed items adequately represent the intended eco-awareness construct and whether the resulting scores demonstrate acceptable empirical validity and internal consistency.

METHODS

This study employed a research and development design focused on the development and preliminary validation of an assessment instrument for measuring senior high school students' eco-awareness regarding the traditional aquatic pond ecosystem of Sidoarjo. The instrument-development process was adapted from the MEASURE approach proposed by Kalkbrenner (2021), which provides a systematic framework for defining a construct, developing an instrument blueprint, constructing items, obtaining expert judgments, conducting pilot testing, and evaluating the psychometric quality of the resulting scores. In the present study, the development procedure comprised seven sequential stages: development of the instrument blueprint, item construction, expert validation, instrument revision, pilot testing with students who were not members of the main research sample, validity and reliability analysis, and finalization of the valid and reliable instrument. This sequence was intended to ensure that the developed items were theoretically aligned with the eco-awareness construct, contextually

relevant to the traditional aquatic pond ecosystem of Sidoarjo, and empirically capable of producing sufficiently consistent scores.

The instrument blueprint was developed by aligning the intended biology learning objectives, the environmental context, the indicators of eco-awareness, the cognitive demands of the questions, and the expected responses. Eco-awareness was operationalized through three interrelated dimensions: ecological knowledge, ecological attitude, and ecological behavior. Ecological knowledge referred to students' understanding and interpretation of ecological concepts and environmental problems; ecological attitude represented their judgments, concerns, and positions toward environmental issues; and ecological behavior referred to their proposed or intended actions in response to ecological problems. Based on this blueprint, seven essay items were constructed to elicit students' reasoning rather than merely their ability to recall factual information. Items 1, 2, and 3 assessed ecological knowledge; Items 4 and 5 assessed ecological attitude; and Items 6 and 7 assessed ecological behavior. The items were designed as pretest and posttest questions using contextual stimuli, including environmental information, illustrations, graphs, and cases related to aquatic ecosystems and environmental sustainability.

The initial instrument was evaluated by four expert validators consisting of two university lecturers and two senior high school biology teachers. The involvement of both academic experts and practising teachers was intended to ensure that the instrument demonstrated theoretical relevance as well as suitability for classroom implementation. The validators examined each item in terms of content relevance, item construction, and linguistic clarity. Content relevance concerned the alignment of the items with the learning objectives, eco-awareness indicators, and aquatic ecosystem content. Item construction included the clarity of the stimulus, correspondence between the question and expected response, cognitive level, scoring guidance, and the absence of unintended clues. Linguistic evaluation addressed readability, grammatical accuracy, clarity of meaning, and appropriateness for Grade 10 students.

Each validation component was rated using a four-point Likert-type scale, in which a score of 1 indicated poor quality, 2 indicated fair quality, 3 indicated good quality, and 4 indicated very good quality. The validation score for each item was calculated by dividing the total score obtained from the validators by the maximum possible score and multiplying the result by 100%. The mean rating was also interpreted according to the criteria presented in Table 1. In addition to numerical ratings, the validators provided qualitative comments that were used to identify ambiguous wording, insufficient contextual information, inadequate visual stimuli, and inconsistencies between the items, answer keys, and scoring rubrics. The instrument was considered eligible for pilot testing after the required revisions had been completed.

Tabel 1. Interpretation Criteria for Content-Validation Scores

Mean score	Interpretation
1,00 – 1,75	Poorly valid
1,76 – 2,50	Moderately valid
2,51 – 3,25	Valid
3,26 – 4,00	Highly Valid

(Sumber : Joshi et al., 2021)

The instrument was subsequently revised on the basis of the validators' quantitative ratings and written recommendations. Revisions included improving the clarity and precision of the questions, refining the answer keys and scoring rubrics, removing potentially ambiguous expressions, and strengthening the stimuli by adding relevant species illustrations, environmental data, graphs, or diagrams where necessary. An item could be retained without modification, revised, replaced, or removed depending on the extent to which it represented the intended eco-awareness indicator. The revised version was then prepared for empirical pilot testing.

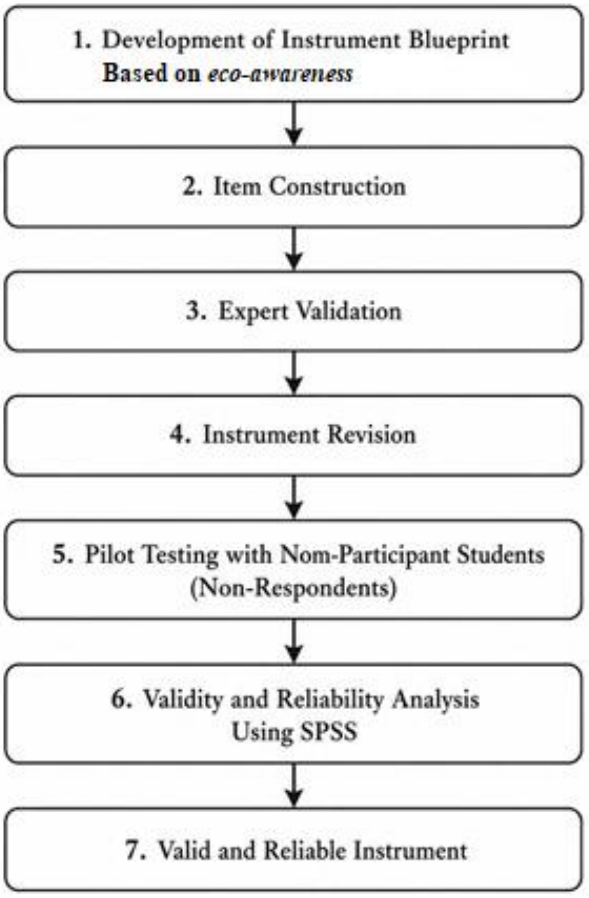

Figure 1. Flowchart of the Research Procedure

Pilot testing was conducted with 30 Grade 11 students from Class XI-10 at a senior high school in Sidoarjo. These students were selected because they had previously studied ecosystem concepts in Grade 10 and were therefore expected to possess sufficient background knowledge to respond meaningfully to the instrument. They were not included in the main research sample, thereby reducing the risk that prior exposure to the assessment items would influence the subsequent implementation of the instrument. The pilot sample was used exclusively to examine the empirical functioning of the seven items and the internal consistency of the resulting scores. To protect participant confidentiality, students' identities were replaced with numerical codes, such as S-1 to S-30, during data entry and analysis.

The students' responses were scored using the predetermined answer keys and analytic scoring rubrics. Each item had a specific maximum score determined by the complexity of the expected response and the eco-awareness indicator being assessed. Before statistical analysis, the scores were checked for completeness and accuracy. The empirical validity of each item was then evaluated using the Pearson product–moment item–total correlation in IBM SPSS Statistics version 27. The score for each item was correlated with the total instrument score to determine whether the item functioned consistently with the overall eco-awareness construct. With a pilot sample of 30 students and a significance level of .05, the critical correlation coefficient was set at

r=.361r = .361r=.361. An item was considered empirically valid when its calculated correlation coefficient exceeded the critical value and its two-tailed probability value was below .05. Items that did not satisfy these criteria were examined in relation to their content, wording, scoring rubric, response distribution, and conceptual alignment before a decision was made to revise or remove them.

The internal consistency of the instrument was estimated using Cronbach's alpha. Reliability analysis was conducted after excluding items that failed to demonstrate acceptable empirical validity because poorly functioning items could distort the estimation of score consistency. Cronbach's alpha was selected because the instrument consisted of multiple items intended to measure related dimensions of the broader eco-awareness construct. The obtained coefficient was interpreted using the categories presented in Table 2. Nevertheless, the reliability coefficient was considered together with the number of items, the multidimensional nature of eco-awareness, the item–total correlations, and the substantive coverage of the construct. Therefore, item-retention decisions were not based exclusively on the alpha coefficient but also on the theoretical and educational importance of each item.

Tabel 2. Interpretation of Instrument Reliability Levels

Cronbach's Alpha	Reliability level
0,00 ≤ r < 0,20	Low
0,20 ≤ r < 0,40	Fair
0,40 ≤ r < 0,60	Moderate
0,60 ≤ r < 0,80	Good
0,80 ≤ r < 1,00	Very Good

(Sumber: Sugiyo, 2017)

The final instrument was determined by integrating evidence from the instrument blueprint, expert validation, qualitative revision, empirical item–total correlations, and internal-consistency analysis. Items supported by both content-based and empirical evidence were retained, whereas items with inadequate empirical performance were revised or excluded. Through this procedure, the study generated an essay-based assessment instrument intended to produce valid and sufficiently reliable evidence of students' ecological knowledge, ecological attitudes, and ecological behavior within the context of Sidoarjo's traditional aquatic pond ecosystem.

RESULT AND DISCUSSION

Content Validity of the Eco-Awareness Assessment Instrument

Content validation was conducted to determine the extent to which the seven essay items represented the intended learning objectives, the three dimensions of eco-awareness, and the contextual characteristics of the traditional aquatic pond ecosystem of Sidoarjo. Four validators, consisting of two university lecturers and two senior high school biology teachers, evaluated each item in terms of content relevance, item construction, and linguistic clarity.

As presented in Table 4, the content-validation percentages ranged from 97.92% to 100%. Items P1 and P4 obtained validation percentages of 97.92%, whereas Items P2, P3, P5, P6, and P7 achieved the maximum percentage of 100%. All seven items were therefore classified as highly valid according to the predetermined interpretation criteria.

Table 4. Results of the Content Validation

No	Assessment aspect	Rerata score instrument penilaian Pretest dan posttes						
		P1	P2	P3	P4	P5	P6	P7
1	Content relevance	3,75	4	4	3,75	4	4	4

2	Item construction	4	4	4	4	4	4	4
3	Linguistic clarity	4	4	4	4	4	4	4
	Total Score	11,75	12	12	11,75	12	12	12
	Maximum score	12	12	12	12	12	12	12
	Percentage (%)	97,92	100	100	97,92	100	100	100
	Interpretation	Highly valid	Highly valid	Highly valid	Highly valid	Highly valid	Highly valid	Highly valid

Note: (P) = nomor soal

Although the quantitative results indicated a very high level of content validity, the validators recommended several minor revisions before the instrument was pilot-tested. These revisions included improving the precision of several questions, clarifying the wording of the answer keys, aligning the scoring rubrics more closely with the expected responses, and strengthening the contextual stimuli through the inclusion of relevant images, graphs, diagrams, and aquatic-species illustrations. The revisions were intended to reduce ambiguity and ensure that students' scores reflected their eco-awareness rather than their ability to interpret unclear questions.

Pilot-Testing Results

The revised instrument was pilot-tested with 30 Grade 11 students who had previously studied ecosystem concepts but were not included in the main research sample. The seven items had different maximum scores because they required different levels of reasoning and response complexity. The maximum total score was 100. The reported total scores ranged from 57 to 100, indicating substantial variation in students' performance. Several students obtained the maximum total score, whereas others demonstrated lower performance, particularly on Items P1, P4, P6, and P7. The variability in the total scores provided an empirical basis for examining whether each item could differentiate among students with different levels of eco-awareness.

Table 5. Results of the Instrument Pilot Testing

No	Participant	1	2	3	4	5	6	7	TOTAL
1	S-1	15	7	18	15	10	15	20	100
2	S-2	15	7	18	15	10	15	20	100
3	S-3	15	7	18	15	10	15	20	100
4	S-4	15	7	18	15	10	15	20	100
5	S-5	15	7	18	10	10	10	20	90
6	S-6	15	7	18	10	10	15	20	95
7	S-7	15	7	18	5	10	10	20	85
8	S-8	15	7	18	10	10	10	20	90
9	S-9	10	7	18	10	10	15	20	90
10	S-10	10	7	18	5	8	15	20	83
11	S-11	15	7	18	5	8	10	20	83
12	S-12	5	7	18	5	10	10	20	75
14	S-14	10	7	18	5	10	10	20	80
14	S-14	15	7	18	15	10	10	20	95
15	S-15	15	7	18	15	10	15	20	100
16	S-16	15	7	18	15	10	10	20	95
17	S-17	15	7	18	10	10	15	20	95
18	S-18	15	7	18	10	10	5	8	73
19	S-19	15	7	9	10	10	5	20	76
20	S-20	12	7	18	5	8	5	20	75
21	S-21	15	7	9	15	15	5	8	74
22	S-22	10	7	18	15	10	5	20	85
23	S-23	15	7	18	5	8	15	20	88
24	S-24	15	7	18	10	10	10	20	90
25	S-25	5	7	10	5	8	10	20	65

No	Participant	1	2	3	4	5	6	7	TOTAL
26	S-26	15	0	9	5	10	0	20	59
27	S-27	10	7	18	10	10	5	8	68
28	S-28	5	0	9	5	8	10	20	57
29	S-29	10	7	18	5	10	0	8	58
30	S-30	6	7	18	5	10	5	8	59

The response pattern also indicated that several items produced relatively homogeneous scores. For example, most students obtained the maximum or near-maximum score on P2, P3, P5, and P7. Such score concentration may reduce an item's capacity to distinguish students with high and low levels of the assessed construct. This issue was examined further through the empirical item-validity analysis.

Empirical Validity of the Assessment Items

The empirical validity of each item was evaluated using the Pearson product–moment correlation between the individual item score and the total instrument score. With 30 pilot participants and a significance level of .05, an item was considered valid when its calculated correlation coefficient exceeded the critical value of r=.361r = .361r=.361 and its probability value was lower than .05. As shown in Table 6, six of the seven items met the empirical-validity criteria. The item–total correlation coefficients for the valid items ranged from .484 to .774. Item P6 demonstrated the strongest relationship with the total score, r=.774r=.774r=.774, followed by P4, r=.675r=.675r=.675, and P1, r=.636r=.636r=.636. Item P5 did not meet the required criterion because it demonstrated a weak and nonsignificant relationship with the total score, r=.092r=.092r=.092, p=.629p=.629p=.629.

Table 6. Results of the Pearson Bivariate Item–Total Correlation Analysis

		Correlations							
		P1	P2	P3	P4	P5	P6	P7	TOTAL
P1	Pearson Correlation	1	.199	.187	.514**	.297	.205	.179	.636**
	Sig. (2-tailed)		.293	.323	.004	.111	.278	.344	.000
	N	30	30	30	30	30	30	30	30
P2	Pearson Correlation	.199	1	.613**	.290	.163	.283	-.120	.484**
	Sig. (2-tailed)	.293		.000	.121	.389	.129	.529	.007
	N	30	30	30	30	30	30	30	30
P3	Pearson Correlation	.187	.613**	1	.154	-.172	.384*	.045	.540**
	Sig. (2-tailed)	.323	.000		.417	.364	.036	.812	.002
	N	30	30	30	30	30	30	30	30
P4	Pearson Correlation	.514**	.290	.154	1	.521**	.304	.054	.675**
	Sig. (2-tailed)	.004	.121	.417		.003	.103	.778	.000
	N	30	30	30	30	30	30	30	30
P5	Pearson Correlation	.297	.163	-.172	.521**	1	-.210	-.439*	.092
	Sig. (2-tailed)	.111	.389	.364	.003		.265	.015	.629
	N	30	30	30	30	30	30	30	30
P6	Pearson Correlation	.205	.283	.384*	.304	-.210	1	.572**	.774**
	Sig. (2-tailed)	.278	.129	.036	.103	.265		.001	.000
	N	30	30	30	30	30	30	30	30
P7	Pearson Correlation	.179	-.120	.045	.054	-.439*	.572**	1	.535**
	Sig. (2-tailed)	.344	.529	.812	.778	.015	.001		.002
	N	30	30	30	30	30	30	30	30
TOTAL	Pearson Correlation	.636**	.484**	.540**	.675**	.092	.774**	.535**	1
	Sig. (2-tailed)	.000	.007	.002	.000	.629	.000	.002	
	N	30	30	30	30	30	30	30	30

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

The result for P5 indicated that students' scores on this item did not vary consistently with their overall eco-awareness scores. In addition, the original inter-item correlation matrix showed that P5 was negatively associated with several other items, particularly P7. This pattern suggests that P5 may have assessed a different response tendency, contained an unclear or inconsistent scoring direction, or provided insufficient score variability. Consequently, P5 should not be retained in its current form. The item requires substantive revision, rescoring, or removal before the instrument is administered in the main study.

Reliability of the Revised Instrument

Reliability analysis was conducted using the six items that fulfilled the empirical-validity criteria. The analysis yielded a Cronbach's alpha coefficient of .656 and a standardized alpha coefficient of .675, as presented in Table 7.

Table 7. Reliability of the Pretest and Posttest Instrument

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.656	.675	6

The obtained coefficient falls within the range of .60–.80, which was classified as good reliability according to the interpretation criteria adopted in the study. Nevertheless, from a broader psychometric perspective, an alpha coefficient of .656 should be interpreted as modest or provisionally acceptable rather than as evidence of high reliability. The result indicates that the six retained items demonstrated a reasonable degree of score consistency for an initial instrument-development study, but additional testing with a larger and more heterogeneous sample is required before the instrument can be used for high-stakes assessment.

Final Instrument and Example of a Retained Item

The integration of expert-validation results, pilot-testing data, empirical item validity, and reliability analysis resulted in a six-item eco-awareness assessment instrument. The retained items represented ecological knowledge, ecological attitude, and ecological behavior in the context of the traditional aquatic pond ecosystem of Sidoarjo. Item P5 was excluded from the preliminary final version because it did not satisfy the empirical-validity criteria.

Table 8 presents a revised example of one of the retained essay items. The example is based on Item P6, which demonstrated the highest item–total correlation. The stimulus and scoring rubric have been condensed to strengthen clarity, reduce unnecessary textual load, and align students' responses with the ecological-attitude and ecological-behavior dimensions.

Table 8. Example of a Retained Pretest and Posttest Item

Component	Description
Learning objective	Students are able to evaluate an environmental-management decision, identify its advantages and limitations, and propose an ecologically responsible alternative.
Eco-awareness indicator	Students evaluate a decision concerning the management of an invasive aquatic species and formulate an environmentally responsible response.
Contextual stimulus	A news-based case describes the mass removal of suckermouth catfish from polluted urban waterways. The species can accumulate pollutants and heavy metals, while the disposal of thousands of captured fish may generate ethical, ecological, and waste-management concerns.
Question	Mass capture and immediate disposal have been used to reduce the population of invasive suckermouth catfish.

Component	Description
	 Dianggap Mengancam, Gubernur DKI Jakarta Instruksikan Pembasmian Ikan Sapu- Sapu -...
	"KOMPAS.com — Suckermouth catfish are consumed in several Latin American countries, particularly in regions crossed by the Amazon River, their native habitat. In those areas, the fish are considered suitable for human consumption because the waters are relatively clean. Therefore, concerns regarding their consumption are more closely related to the environmental conditions in which they live than to the species itself. Suckermouth catfish harvested from Indonesian waters generally do not meet food-safety requirements. Although they are abundant and contain a relatively high level of protein, they should not be regarded as an alternative food source. 'Those who wish to consume suckermouth catfish should select fish obtained from relatively clean waters, such as rural rivers or aquaculture facilities. However, pregnant women, children, and older adults should avoid consuming them entirely to ensure maximum safety,' stated Surya Gentha Akmal, an expert on invasive aquatic species at IPB University, during a webinar on Saturday, 25 April 2026. As benthic feeders, suckermouth catfish absorb organic pollutants and heavy metals from river sediments, particularly in polluted urban waters. According to data from the Ministry of Environment, approximately 70.7% of monitored river sites in Indonesia are contaminated with heavy metals. Given the poor quality of many Indonesian waterways, suckermouth catfish may act as bioaccumulators because they ingest sediment from riverbeds, where heavy metals tend to accumulate. Source: KOMPAS.com, 'The Dilemma of Controlling Suckermouth Catfish in Indonesia.' At present, the mass capture of suckermouth catfish remains controversial because safe utilization methods have not yet been established and the effectiveness of large-scale removal is uncertain. 'Videos circulating on social media often show local communities capturing between 20,000 and 30,000 suckermouth catfish in a single day. The fish are then discarded in waste-disposal sites in quantities reaching several tonnes or buried alive. Is this approach truly effective? What position should you take in addressing this issue?'" Evaluate whether this measure is environmentally effective. State your position, justify it using ecological reasoning, and propose a more responsible management strategy.
Cognitive level	C4—Analyzing
Item number	P6
Item type	Essay
Expected response	Students should explain that population control may be necessary but that indiscriminate capture and disposal alone are insufficient. A comprehensive response may include safe utilization or processing of captured fish, pollution-source control, public education, prevention of the release of non-native species, monitoring of aquatic ecosystems, and consideration of animal welfare and

Component	Description
	environmental safety.
Scoring rubric	0 points: No answer or an irrelevant response. 5 points: States a position but provides limited justification. 10 points: States a position and provides scientifically relevant ecological reasoning. 15 points: Provides an evidence-based evaluation and proposes an integrated, feasible, and environmentally responsible solution.
Maximum score	15

The example demonstrates that the instrument was designed not only to assess students' recognition of environmental problems but also to elicit their ability to evaluate competing considerations, justify decisions, and propose responsible ecological actions.

Discussion

The findings provide preliminary evidence that the developed assessment instrument is capable of measuring students' eco-awareness within the context of Sidoarjo's traditional aquatic pond ecosystem. Expert evaluation produced content-validity percentages ranging from 97.92% to 100%, indicating strong agreement regarding the relevance of the content, item construction, and linguistic clarity. Nevertheless, empirical testing showed that only six of the seven items were statistically associated with the total score, while Item P5 produced a weak and nonsignificant item–total correlation. The contrast between the extremely high expert ratings and the failure of one item during pilot testing confirms that content validity and empirical validity represent complementary rather than interchangeable forms of evidence. Expert judgment can establish whether an item appears theoretically and pedagogically appropriate, but empirical analysis is still required to determine whether it functions consistently when administered to students. This interpretation is consistent with Kalkbrenner (2021) MEASURE approach, which positions theoretical blueprint development, expert review, participant recruitment, and validity and reliability evaluation as sequential components of instrument development.

The results are also comparable with the instrument-development study conducted by Astiwi et al. (2020), who developed an essay-based critical-thinking assessment instrument and reported strong content validity, valid item coefficients, and a reliability coefficient of .84. Both studies demonstrate the importance of integrating an assessment blueprint, essay items, expert judgment, empirical item testing, and reliability analysis. However, the reliability coefficient obtained in the present study, Cronbach's alpha = .656, was lower than that reported by Astiwi et al. This difference may be partly attributable to the substantially smaller number of retained items in the present instrument. Astiwi et al. used 20 items, whereas the current instrument retained only six items distributed across three dimensions of eco-awareness. Because internal-consistency estimates are affected by test length and the degree of similarity among items, a short instrument covering ecological knowledge, ecological attitudes, and ecological behavior may reasonably produce a more moderate coefficient. Accordingly, the present alpha value should be interpreted as preliminary or provisionally acceptable rather than as evidence of high reliability.

The multidimensional structure of the instrument is aligned with Levy-Feldman (2025) argument that educational assessment should adequately represent different domains of learning and meet the minimum requirements of validity, reliability, and objectivity. In the present study, the use of essay questions enabled students to explain ecological concepts, evaluate environmental decisions, and propose environmentally responsible actions. This design is potentially more informative than a conventional selected-response test because it captures the reasoning underlying students' answers. At the same time, essay-based assessment creates an additional requirement for transparent scoring criteria and consistent evaluator judgments. Although the developed instrument included answer keys and scoring rubrics, the study did not report inter-rater agreement or scorer reliability. Thus, the internal consistency of item scores does not by itself demonstrate that different scorers would assign equivalent scores to the same

student response. Future validation should therefore involve at least two independent raters and calculate an appropriate inter-rater reliability coefficient.

The empirical behavior of the items also supports the argument that eco-awareness should not be treated as a single, uniformly expressed attribute. Item P6, which required students to evaluate an environmental-management problem and propose an ecologically responsible solution, produced the strongest item–total correlation, r=.774r=.774r=.774. This result suggests that contextual problem-solving may capture the integration of knowledge, judgment, and intended action more effectively than items that assess only one component. In contrast, Item P5 obtained r=.092r=.092r=.092 and p=.629p=.629p=.629, suggesting that it did not operate consistently with the overall instrument. Its weak performance may have resulted from restricted score variation, ambiguous wording, socially desirable responses, an insufficiently discriminating rubric, or a mismatch between the item and the intended ecological-attitude indicator. Therefore, Item P5 should not merely be removed mechanically; its construct representation, response format, answer key, and scoring categories should be reviewed before a revised version is retested.

The need to measure ecological knowledge, attitudes, and behavior separately is reinforced by Cahyani et al. (2025). Their study of senior high school students found moderate environmental literacy, with relatively low ecological knowledge and cognitive-skill scores despite substantially higher environmental attitudes and behavior scores. They also found no significant relationship between environmental knowledge and students' attitudes or behaviors. These findings indicate that possessing positive environmental attitudes does not necessarily mean that students have adequate ecological understanding, while scientific knowledge does not automatically translate into consistent environmental behavior. The present instrument responds to this problem by incorporating different item groups for ecological knowledge, ecological attitudes, and ecological behavior rather than deriving eco-awareness from a single general score or attitude statement. However, because only six items were retained, the current instrument does not yet contain enough items to generate stable and separately reliable subscale scores for all three dimensions.

A similar comparison can be made with Saro et al. (2025), who reported that senior high school students demonstrated positive environmental attitudes and awareness but only neutral ecological behavior. Their results also showed meaningful relationships among attitudes, behavior, and awareness, particularly between ecological behavior and environmental awareness. The present findings extend this line of inquiry by developing an instrument that requires students not only to indicate environmental concern but also to justify decisions and formulate possible environmental actions through essay responses. This distinction is important because students can select socially desirable responses on an attitude questionnaire without demonstrating the reasoning needed to apply those attitudes to an authentic ecological dilemma. Consequently, the developed assessment may offer richer diagnostic information about the extent to which students can transform environmental concern into reasoned and feasible courses of action.

The contextual orientation of the instrument is consistent with Situmorang & Tarigan (2018) study, in which bottle-garden activities were used to cultivate students' environmental awareness. Their findings indicated that students developed a good understanding of several environmental-awareness dimensions, including ecological crisis, the balance of nature, limits to growth, and anti-anthropocentrism. While their study focused on environmental learning through direct school-based activities, the present study contributes an assessment mechanism for examining students' responses to environmental problems grounded in a locally relevant aquatic ecosystem. Together, these studies suggest that environmental education becomes more meaningful when ecological concepts are connected to concrete issues, organisms, resources, and environmental practices encountered in students' surroundings. The traditional pond ecosystem of Sidoarjo therefore functions not simply as background content but as an authentic socioecological setting through which students' ecological reasoning can be elicited and

evaluated.

The findings also correspond with Ardoin et al. (2020) systematic review of 105 environmental-education studies. Their synthesis showed that programs reporting direct environmental outcomes commonly focused on localized environmental issues, incorporated opportunities for action, involved collaboration with relevant stakeholders, and intentionally measured their outcomes. The present study addresses the measurement component by producing a context-specific instrument that connects biology learning with local aquatic environmental problems. However, whereas Ardoin et al. emphasized the importance of linking education to observable environmental action, the present instrument primarily measures written knowledge, evaluative judgment, and proposed behavior. It therefore assesses students' behavioral reasoning or intended ecological action rather than directly observed pro-environmental behavior. This distinction should be clearly acknowledged when interpreting the scores, because a well-reasoned written answer does not guarantee that the student will perform the corresponding behavior in real situations.

The principal novelty of this study lies in the development of an essay-based eco-awareness assessment that integrates ecological knowledge, ecological attitudes, and ecological behavior within the specific context of Sidoarjo's traditional aquatic pond ecosystem. Previous studies have generally measured environmental literacy through broad questionnaires, examined correlations among environmental constructs, or evaluated environmental-awareness interventions. In contrast, this study combines locally situated aquatic problems, higher-order essay questions, expert validation, empirical item analysis, and reliability testing within a single assessment-development process. The instrument also encourages students to interpret environmental information, evaluate competing ecological considerations, defend a position, and formulate contextually feasible responses. Thus, its contribution is not merely the production of another environmental-awareness questionnaire but the provision of a contextual diagnostic assessment that can reveal how students reason about local aquatic environmental problems.

The findings have several implications for biology education and educational assessment. For teachers, the six retained items can serve as a preliminary diagnostic instrument for identifying differences among students' ecological understanding, environmental judgment, and behavioral reasoning before and after instruction. The scoring results may help teachers determine whether students require additional support in interpreting ecological data, evaluating human impacts, or designing environmentally responsible solutions. For curriculum developers, the study demonstrates how local aquatic ecosystems can be incorporated into assessment tasks without separating scientific content from environmental responsibility. For schools and local education stakeholders, the instrument may support place-based environmental-learning programs related to water quality, pond sustainability, invasive species, pollution, aquaculture management, and biodiversity conservation. Nevertheless, the instrument should currently be used for formative or research purposes rather than high-stakes decision-making because the available psychometric evidence remains preliminary.

Several limitations must be considered when interpreting the findings. The pilot test involved only 30 students from a single class, which limits the stability of the correlation coefficients and the transferability of the findings to students from other schools, grade levels, and socioenvironmental settings. The involvement of four validators provided useful expert feedback, but a broader panel incorporating assessment specialists, aquatic-ecology experts, and environmental-education scholars could strengthen the content-validation process. The reliability coefficient of .656 indicates modest internal consistency and requires confirmation using a larger item pool and a more heterogeneous sample. The study did not examine inter-rater reliability, test–retest reliability, item difficulty, discrimination indices, response-process evidence, or the dimensional structure of the instrument. The sample size was also insufficient for exploratory or confirmatory factor analysis. In addition, the use of item–total correlations should be refined by reporting corrected item–total correlations that exclude the respective item from the total score. Finally, the locally specific content enhances contextual relevance but may

restrict direct application in regions with different aquatic ecosystems. Adaptation to other settings should therefore involve renewed contextual review, expert validation, pilot testing, and reliability analysis rather than assuming that the current evidence automatically applies across populations. The raw pilot dataset should also be rechecked before publication because the original table contains a duplicated participant code and one P5 score that exceeds the stated maximum score.

Conclusion

This study developed and provided preliminary validation evidence for an essay-based instrument designed to assess senior high school students' eco-awareness in the context of Sidoarjo's traditional aquatic pond ecosystem. Expert evaluation showed that all seven items demonstrated very high content validity, with validation percentages ranging from 97.92% to 100%, confirming their relevance in terms of ecological content, item construction, and linguistic clarity. Empirical testing with 30 students indicated that six items met the item-validity criteria, with item–total correlation coefficients ranging from .484 to .774, whereas one item required revision or removal because it did not function consistently with the overall instrument. The six retained items produced a Cronbach's alpha coefficient of .656, indicating acceptable internal consistency for preliminary and formative use. The main contribution of this study lies in integrating ecological knowledge, ecological attitudes, and ecological behavior into a locally contextualized essay assessment that requires students to interpret environmental information, evaluate ecological problems, justify decisions, and propose responsible actions. The instrument may therefore support biology teachers in diagnosing students' eco-awareness and evaluating context-based environmental learning; however, its use should remain limited to formative assessment and research until further validation is completed. Future studies should involve larger and more diverse samples, expand the number of items within each eco-awareness dimension, examine inter-rater and test–retest reliability, and evaluate the instrument's construct validity and transferability across different aquatic and socioenvironmental contexts.

References

Al-Barakat, A., AlAli, R., Alotaibi, S., Alrashood, J., Abdullatif, A., & Zaher, A. (2025). Science education as a pathway to sustainable awareness: Teachers' perceptions on fostering understanding of humans and the environment—A qualitative study. *Sustainability, 17*(15), Article 7136. https://doi.org/10.3390/su17157136

AlAli, R., Al-Barakat, A., Alrosaa, T., Alotaibi, S., Abdullatif, A., & Almughyirah, S. (2025). Science education and environmental identity: An integrative approach to fostering sustainability practices in primary school students. *Sustainability, 17*(19), Article 8883. https://doi.org/10.3390/su17198883

Albu, C., & Lindmeier, A. (2023). Performance assessment in teacher education research—A scoping review of characteristics of assessment instruments in the DACH region. *Zeitschrift für Erziehungswissenschaft, 26*(3), 751–778. https://doi.org/10.1007/s11618-023-01167-7

Ardoin, N. M., Bowers, A. W., & Gaillard, E. (2020). Environmental education outcomes for conservation: A systematic review. *Biological Conservation, 241*, Article 108224. https://doi.org/10.1016/j.biocon.2019.108224

Ardoin, N. M., & Heimlich, J. E. (2021). Environmental learning in everyday life: Foundations of meaning and a context for change. *Environmental Education Research, 27*(12), 1681–1699. https://doi.org/10.1080/13504622.2021.1992354

Begum, A., Wang, Q., Qayum, H., Khan, S., & Zaib, K. (2025). Fostering green behaviour through environmental education: A structural equation model of nature connectedness and social dominance orientation. *SAGE Open, 15*(3). https://doi.org/10.1177/21582440251378188

Benzehaf, B., Razkane, H., & Benzehaf, O. (2025). Exploring university students' perceptions and engagement in environmental awareness and sustainable practices. *Discover Environment, 3*(1), Article 247. https://doi.org/10.1007/s44274-025-00462-w

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers*

in Public Health, 6, Article 149. https://doi.org/10.3389/fpubh.2018.00149

Burgos-Espinoza, I. I., García-Alcaraz, J. L., Gil-López, A. J., & Díaz-Reza, J. R. (2025). Effect of environmental knowledge on pro-environmental attitudes and behaviors: A comparative analysis between engineering students and professionals in Ciudad Juárez, Mexico. *Journal of Environmental Studies and Sciences, 15*(4), 861–875. https://doi.org/10.1007/s13412-024-00991-5

Cahyani, R. P., Rahmat, A., Hamdiyati, Y., Amprasto, A., & Jamil, M. W. (2025). Does environmental knowledge predict attitudes and behaviors in environmental literacy? Evidence from senior high school students. *JPBI (Jurnal Pendidikan Biologi Indonesia), 11*(3), 828–840. https://doi.org/10.22219/jpbi.v11i3.41213

Chen, W., & Erfanian, S. (2026). Drivers of students' pro-environmental behavior in campus environments in Chinese universities. *Frontiers in Psychology, 17*, Article 1760915. https://doi.org/10.3389/fpsyg.2026.1760915

Cook, D. A., & Hatala, R. (2016). Validation of educational assessments: A primer for simulation and beyond. *Advances in Simulation, 1*(1), Article 31. https://doi.org/10.1186/s41077-016-0033-y

Cruchinho, P., López-Franco, M. D., Capelas, M. L., Almeida, S., Bennett, P. M., Miranda da Silva, M., Teixeira, G., Nunes, E., Lucas, P., & Gaspar, F. (2024). Translation, cross-cultural adaptation, and validation of measurement instruments: A practical guideline for novice researchers. *Journal of Multidisciplinary Healthcare, 17*, 2701–2728. https://doi.org/10.2147/JMDH.S419714

Delgado-Guerrero, Y. M., López-Ramírez, P. A., Viviescas-Beltrán, L. S., & Ríos-Reyes, C. A. (2025). Geoeducation for sustainable rural development: Enhancing environmental literacy through geoscience learning in Guaca, Santander, Colombia. *Sustainable Futures, 10*, Article 101520. https://doi.org/10.1016/j.sftr.2025.101520

Hsu, C.-H. (2025). Outdoor environmental education as a nature-based solution for "education" and "environment": A new conceptual framework and its pilot application in a coastal community case study in Taiwan. *Journal of Coastal Conservation, 29*(1), Article 13. https://doi.org/10.1007/s11852-025-01099-w

Jeronen, E., & Jeronen, J. (2026). Bibliometric and content analysis of sustainable education in biology for promoting sustainability at primary and secondary schools and in teacher education. *Education Sciences, 16*(2), Article 201. https://doi.org/10.3390/educsci16020201

Kalkbrenner, M. T. (2021). A practical guide to instrument development and score validation in the social sciences: The MEASURE approach. *Practical Assessment, Research, and Evaluation, 26*(1), 1–18.

Kang, J., & Teig, N. (2025). Which science matters? The role of biology, chemistry, physics, and earth science in shaping students' environmental attitudes and behaviours. *International Journal of Science Education*, 1–21. https://doi.org/10.1080/09500693.2025.2579584

Levy-Feldman, I. (2025). The role of assessment in improving education and promoting educational equity. *Education Sciences, 15*(2), Article 224. https://doi.org/10.3390/educsci15020224

Lin, S.-H., Sah, A. K., & Hong, Y.-M. (2025). Environmental literacy among the general public in Chiayi County, Taiwan. *Sustainability, 17*(7), Article 3108. https://doi.org/10.3390/su17073108

Lye, C.-T., Ng, T.-H., & Law, J.-W. (2024). The role of general and specific pro-environmental education in household waste management in Malaysia: Evidence from quantile regression. *Environmental Challenges, 15*, Article 100933. https://doi.org/10.1016/j.envc.2024.100933

Mikkonen, K., Tomietto, M., & Watson, R. (2022). Instrument development and psychometric testing in nursing education research. *Nurse Education Today, 119*, Article 105603. https://doi.org/10.1016/j.nedt.2022.105603

Musah, E., & Wangila, C. (2024). Inquiry-based ethnoecological conservation approach to improve students' critical thinking skills. *International Journal of Ethnoscience and Technology in Education, 1*(2), 97. https://doi.org/10.33394/ijete.v1i2.11240

Rahmania, T. (2024). Exploring school environmental psychology in children and adolescents: The influence of environmental and psychosocial factors on sustainable behavior in Indonesia. *Heliyon, 10*(18), Article e37881. https://doi.org/10.1016/j.heliyon.2024.e37881

Saro, J. M., Taray, J. D., Cinso, J. S., Ingaling, B. S., Enguio, M. J. A., & Hayon, C. P. B. (2025). Level of attitude, behavior, and awareness of senior high school students toward environmental conservation. *JPBI (Jurnal Pendidikan Biologi Indonesia), 11*(1), 111–124. https://doi.org/10.22219/jpbi.v11i1.37861

Sarwi, Fathonah, S., Winarto, Fatimah, S., Sari, E. F., & Fadilah, A. M. (2025). Ethnoscience and local wisdom in science education for conservation: A systematic review with bibliometric mapping. *Multidisciplinary Reviews, 9*(7), Article 2026298. https://doi.org/10.31893/multirev.2026298

Situmorang, R. P., & Tarigan, S. D. (2018). Cultivating students' environmental awareness by creating a bottle garden in school: A qualitative study. *JPBI (Jurnal Pendidikan Biologi Indonesia), 4*(3), 263–270. https://doi.org/10.22219/jpbi.v4i3.6785

Sotero, M. C., Alves, Â. G. C., Arandas, J. K. G., & Medeiros, M. F. T. (2020). Local and scientific knowledge in the school context: Characterization and content of published works. *Journal of Ethnobiology and Ethnomedicine, 16*(1), Article 23. https://doi.org/10.1186/s13002-020-00373-5

Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education, 48*(6), 1273–1296. https://doi.org/10.1007/s11165-016-9602-2

Ullah, R., Harrington, J., Farea, A., Otreba, M., Carroll, S., & McKenna, T. (2026). Digital maturity assessment tools for the construction industry: A PRISMA-ScR scoping review. *Buildings, 16*(1), Article 239. https://doi.org/10.3390/buildings16010239

Valdivia-Salazar, C., Serquén, O., Guevara, L., Bravo-Jaico, J., Alarcón, R., Germán, N., Aquino, J., Maquen-Niño, G. L. E., & Moreno Heredia, A. (2025). Development and validation of an instrument to assess the maturity of digital transformation in higher education institutions. *Frontiers in Education, 10*, Article 1691178. https://doi.org/10.3389/feduc.2025.1691178

Vieriu, A. M., & Petrea, G. (2025). The impact of artificial intelligence on students' academic development. *Education Sciences, 15*(3), Article 343. https://doi.org/10.3390/educsci15030343

Wibowo, A. M., Utaya, S., Wahjoedi, W., Zubaidah, S., Amin, S., & Prasad, R. R. (2024). Critical thinking and collaboration skills on environmental awareness in project-based science learning. *Jurnal Pendidikan IPA Indonesia, 13*(1), 103–115. https://doi.org/10.15294/jpii.v13i1.48561

Wibowo, N. A., Sumarmi, S., Utaya, S., Bachri, S., & Kodama, Y. (2023). Students' environmental care attitude: A study at Adiwiyata public high school based on the New Ecological Paradigm. *Sustainability, 15*(11), Article 8651. https://doi.org/10.3390/su15118651

Xin, J. J., & Lo, Y. Y. (2025). Developing the CLIL teacher assessment literacy inventory. *System, 133*, Article 103776. https://doi.org/10.1016/j.system.2025.103776

Xiong, Z., Song, Y., & Zhu, R. (2025). Pedagogical strategies for teaching environmental literacy in secondary school education: A systematic review. *Sustainability, 17*(20), Article 9104. https://doi.org/10.3390/su17209104

Zebua, N., Ziraluo, Y. P. B., & Zebua, E. N. K. (2025). Environmental conservation in biology learning: An effort to increase students' awareness of environmental issues. *International Journal of Advances in Educational Research, 1*(3), 145–156. https://doi.org/10.62941/ijaer.v1i3.114