Prisma Sains: Jurnal Pengkajian Ilmu dan Pembelajaran Matematika dan IPA IKIP Mataram
https://e-journal.undikma.ac.id/index.php/prismasains/index
e-mail: prismasains.pkpsm@gmail.com

October 2026. Vol. 14, No. 4
p-ISSN: 2338-4530
e-ISSN: 2540-7899
pp. 2847-2862

Development and Preliminary Evaluation of an Ethnopedagogy Based Science Module Integrating Bima Local Wisdom for Grade VIII Scientific Literacy

[1]*Rabiatul Adawiyah, [2]Nurzakiyyah, [3]Istin Fitriana Aziza, [4]Norma Juniati

[1] Chemistry Education, Universitas Nggusuwaru, Bima, Indonesia
[2] Enviromental Engineering, Universitas Nggusuwaru, Bima, Indonesia
[3] Computer Science, Universitas Nggusuwaru, Bima, Indonesia
[4] Biology Science Education, Universitas Islam Negeri Mataram, Mataram, Indonesia

*Corresponding author e-mail: rabiatuladwyh123@gmail.com

Received: August 2026; Revised: October 2026; Published: October 2026

Abstract

Local-wisdom science materials often introduce cultural examples without showing how community knowledge, classroom activities, and assessment are connected. This study develops and evaluates a Grade VIII science module that links bingka dolu, a Bima cake, and oi tua, a palm-sap context, to food additives and addictive substances. Its contribution is an explicit mapping of cultural claims to scientific questions, five ethnopedagogical stages, and three scientific-literacy competencies. The study uses an adapted 4D development framework and a one-group pretest–posttest design in two schools. Participants comprise 54 students, two teachers, and three academic validators; a separate pilot comprises 60 students. The assessment contains 18 constructed-response items. Mean module relevance is 0.92 on Aiken's V, and student and teacher questionnaire means are 3.56 and 3.84 on a four-point scale. Mean literacy scores increase from 50.10 to 74.07, with a paired change of 23.97 points, a 95% confidence interval of 20.12–27.82, and normalized gain of 0.48. Inquiry shows the smallest increase among the three competencies. These findings support a preliminary description of the module and within-group learning change. They do not establish a causal effect. Further cultural review, documented implementation, and evaluation using comparison groups are needed to determine the educational value and wider applicability of the design.

Keywords: Ethnopedagogy; Bima local wisdom; Scientific literacy; Food additives; Junior high school.

How to Cite: Adawiyah, R., Nurzakiyyah, N., Aziza, I. F., & Juniati, N. (2026). Development and Preliminary Evaluation of an Ethnopedagogy Based Science Module Integrating Bima Local Wisdom for Grade VIII Scientific Literacy. *Prisma Sains: Jurnal Pengkajian Ilmu Dan Pembelajaran Matematika Dan IPA IKIP Mataram*, *14*(4), 2847–2862. https://doi.org/10.33394/j-ps.v14i4.22447

https://doi.org/10.33394/j-ps.v14i4.22447

Copyright© 2026, Adawiyah et al.
This is an open-access article under the CC-BY License

INTRODUCTION

Scientific literacy requires learners to evaluate evidence, judge the credibility of scientific claims, and use knowledge in decisions that matter beyond the classroom (Allchin, 2022; Tasquier et al., 2022). Recent scholarship also examines whose knowledge and concerns enter those decisions (Guerrero & Sjöström, 2025). This module operationalizes a narrower, assessable scope through the three PISA 2018 science competencies: explaining phenomena scientifically, evaluating and designing scientific enquiry, and interpreting data and evidence scientifically (Organisation for Economic Co-operation and Development [OECD], 2019). Indonesia's PISA 2022 science mean of 383, against an OECD mean of 485, provides national context (OECD, 2023). Indonesian reviews discuss weaknesses in conceptual understanding, learning resources, and opportunities to use science in context (Fuadi et al., 2020; Yusmar & Fadilah, 2023). These findings motivate local investigation; they do not establish the needs of a particular Bima classroom.

The Grade VIII module provides a local basis for examining this problem. Its 56 pages

contain two 80-minute lessons, readings on bingka dolu and oi tua, collaborative questions, dialogue, reflection, and self-evaluation. Bingka dolu introduces ingredient functions and natural-colouring claims, whereas oi tua raises questions about changes in palm sap and fermentation. A review of the module identifies tasks that emphasize naming substances, with less explicit attention to comparing evidence, controlling variables, and qualifying conclusions. Some passages also equate natural origin with safety. These features establish a need to strengthen the product's treatment of evidence. Difficulties with hypotheses, experimental design, and conclusions have also been documented among secondary-school learners in Spain (Ayuso Fernández et al., 2022). That research offers a comparison, while the local rationale for the present study rests on the module analysis.

Ethnoscience concerns knowledge associated with cultural practices and its relationship to scientific explanation; ethnopedagogy concerns how cultural experience, identity, and participation inform teaching (Alwasilah et al., 2009; Sari et al., 2023). Culturally relevant pedagogy provides a foundational account of the relationship between academic learning and cultural competence (Ladson-Billings, 1995), while a recent science-education treatment emphasizes learners' cultural resources and community participation (Smith et al., 2022). This distinction has practical consequences. Chemistry examples from Zambia show that local practices can supply curriculum content (Chibuye & Singh, 2024), whereas a Lombok study identifies teachers' selection, interpretation, negotiation, and representation of community knowledge as substantive instructional processes (Bahtiar et al., 2026). Cultural identity also requires its own evidence, as illustrated by Andayani et al. (2023). Naming a traditional food therefore cannot, by itself, establish culturally responsive teaching or scientific literacy.

Recent reviews document substantial work on ethnoscience and contextual science learning (Hikmah et al., 2025; Putra & Wahyuni, 2025; Yanto et al., 2025). Table 1 compares 10 studies by cultural context, topic, pedagogy, and outcome. Grade VIII cultural materials, traditional-food projects, and local-wisdom modules are established approaches. The question addressed here is whether the link from a cultural claim to a scientific task and a scored literacy competency can be made explicit for both bingka dolu and oi tua. The study documents this alignment, reports academic review and perceived usability, and examines pretest–posttest change in explaining phenomena, evaluating and designing inquiry, and interpreting evidence. This scope locates the contribution in the documented design and the pattern of competency outcomes, without claiming that culture-based module development is itself new.

Table 1. Selected prior studies and implications for the present design

Study	Context and topic	Design or pedagogy	Outcome addressed	Boundary for the present work
Rahmawati et al. (2020)	Grade VII, Bolo, Bima; solar system	Five-stage ethnopedagogy; qualitative study	Literacy and cultural identity	Keep the local pedagogical precedent; specify a different topic and assessment chain.
Adawiyah et al. (2022)	Grade XI chemistry, Pujut	ADDIE; ethnoscience with CRTT	Expert judgments and student responses	Module feasibility precedes this work; specify the evidence demanded by each task.
Amalia et al. (2024)	Riau Malay contexts in science assessment	Instrument development	Contextual scientific-literacy assessment	Report competency alignment and item quality for the intended population.
Jufrida et al. (2024)	Grade VIII, Jambi; simple machines	Discovery learning; mixed methods	Literacy tests and implementation accounts	Use task mapping alongside observations; a test total alone cannot explain learning.

Study	Context and topic	Design or pedagogy	Outcome addressed	Boundary for the present work
Uslan et al. (2024)	Elementary schools, Kupang	Local-knowledge module; comparison groups	Scientific literacy and thinking skills	Comparative findings motivate stronger trials; they do not validate this Grade VIII design.
Haen et al. (2025)	Grade VIII; Bakumpai local wisdom	4D teaching-material development	Validity and practicality	A familiar culture and target grade are insufficient novelty; distinguish usability from learning.
Mardiana et al. (2025)	Junior-high science; local wisdom	4D e-module; expert validation	Content, presentation, language, graphics	Expert approval supports product review, not an independent causal-effect claim.
Syahdani and Tyas (2025)	Grade VIII, Sleman; jadah tempe	PjBL; non-equivalent comparison groups	Literacy and communication	Traditional-food literacy projects already exist; define the additional alignment logic.
Wahyuni et al. (2025)	Grade VII, Puger; Grenden Village	ADDIE science textbook	Pre–post literacy and user responses	Make competency-level reasoning explicit; gain scores are not directly comparable across studies.
Bahtiar et al. (2026)	15 science teachers across Lombok	Qualitative multiple-case study	Translation of Indigenous knowledge into teaching	Record whose knowledge is selected and how teachers interpret and represent it.

This study aims to develop and preliminarily evaluate an ethnopedagogy-based science module integrating Bima local wisdom, specifically *bingka dolu* and *oi tua*, into the Grade VIII topic of food additives and addictive substances. Specifically, the study seeks to document the alignment between cultural content, scientific concepts, the five ethnopedagogical stages, and scientific-literacy assessment; assess the module's relevance through academic expert review; describe students' and teachers' perceptions of its usability; and examine within-group changes in students' scientific literacy before and after module implementation. These changes are examined across three competencies: explaining phenomena scientifically, evaluating and designing scientific enquiry, and interpreting data and evidence scientifically.

METHOD

Research design and setting

The study follows the 4D development framework of Thiagarajan et al. (1974), with the present evaluation focused on Define, Design, and Develop. The definition and design stages are described through curriculum and module analysis; the development stage includes academic review, instrument analysis, and the school trial. Dissemination is outside the reported evaluation. Recent 4D applications provide relevant precedents (Haen et al., 2025; Mardiana et al., 2025; Martin et al., 2025). The field design compares the same students before and after teaching in each of two schools, conventionally O_1–X–O_2. O_1 and O_2 denote pretest and posttest assessment, and X denotes instruction using the module. There is no untreated comparison group or random allocation. As in one-group ethnoscience research reported by Dewi et al. (2021), the design describes within-group change but cannot isolate an instructional effect.

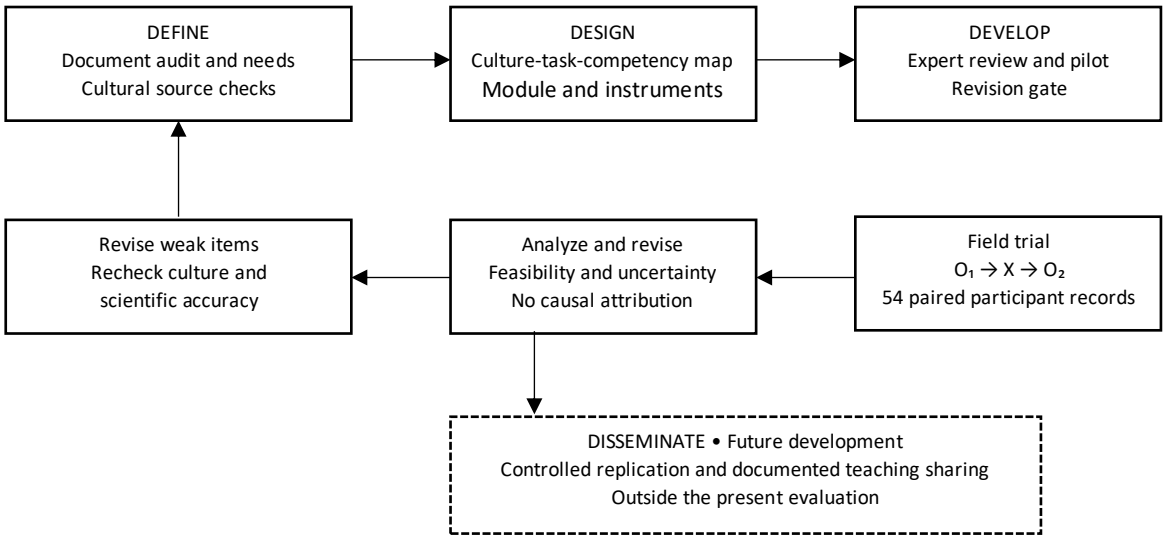

Figure 1. Adapted 4D workflow and the one-group trial. Solid boxes describe the design and trial; the dashed dissemination box is a future activity. The loop indicates revision, not evidence that repeated field trials occurred.

Participants and sampling

The school sample comprises 54 Grade VIII students: 36 at MTsN 1 Kota Bima, an Islamic junior high school, and 18 at SMP IT Insan Kamil, an integrated Islamic junior high school. These institutions are labelled School A and School B, respectively, in the results. The sample comprises one intact class per school, ages 13–14 years, purposive school selection based on curriculum fit and access, and one science teacher per class. The analysis uses complete paired records for 54 students. A separate pilot sample of 60 students is used for instrument analysis and is not included in the school-trial total.

Three academic validators assess the module. The selection criteria cover science education, chemistry content, and junior-high science teaching with instructional-design experience. Academic content review and cultural confirmation address different questions: the former evaluates instructional relevance, whereas the latter establishes the accuracy and attribution of community knowledge. Cultural claims in the module therefore require confirmation by practitioners with relevant food-preparation or palm-sap expertise.

4D development procedures

Define addresses curriculum requirements and the content of the module. The module analysis identifies two priorities: extending recall-oriented questions to require evidence interpretation, and separating cultural descriptions from scientific and health claims. Design links the cultural contexts to literacy competencies, learning activities, and scoring criteria. The five teaching stages are self-identification, content integration, collaboration, dialogue, and reflection, following Rahmawati et al. (2020) and their application in chemistry learning by Andayani et al. (2023). The analysis does not treat national PISA scores as evidence of the needs of these particular classrooms.

Develop is described through academic review, instrument diagnostics, the restricted school trial, and the product issues summarized in Table 3. Academic ratings inform revision but cannot establish cultural authenticity or classroom effectiveness. A content statement requires correction when it is inaccurate, poorly sourced, or misaligned with the intended competency, even if average ratings are favourable. Dissemination and evaluation of subsequent adoption are reserved for further development. The present findings concern the two participating schools.

Module content and cultural mapping

Bingka dolu and oi tua provide the module's cultural contexts; Mbojo is the local cultural designation used in the learning material. Table 2 links source claims to scientific

reconstruction, teaching activities, and assessment competencies. The descriptions are attributed to the module rather than presented as direct testimony from community members. Recipe variants, terminology, historical accounts, and consumption practices require practitioner confirmation. Learners can discuss another familiar household food practice and are not required to disclose family alcohol consumption. Community participation is central to culturally responsive teaching (Smith et al., 2022). Bahtiar et al. (2026) further distinguish teachers' selection, interpretation, and representation of community knowledge. These considerations define the limits of the present cultural interpretation.

Table 2. Alignment of cultural content, science, teaching stages, and assessment

Practice and source claim	Scientific reconstruction and concept	Stages and learning activity	Literacy competency and assessment specification
– Bingka dolu; pandan/suji and sugar are described as contributing colour, aroma, and sweetness. Source: prototype, printed pp. 26–29.	Distinguish ingredient function from proof of safety; compare a recipe with supplied ingredient information.	Self-identification and content integration: describe an observed recipe, then classify functions and justify the evidence.	Explain phenomena; items E1–E3. Identification alone earns less credit than an evidence-based explanation.
– Bingka dolu; the prototype links sugar to preservation and describes natural colouring as safer.	Investigate concentration and controlled conditions; natural origin is insufficient to establish safety or shelf life.	Collaboration and dialogue: critique a comparison that changes several variables, then redesign it.	Evaluate and design inquiry; I1–I3. Interpret evidence; D1–D3.
– Oi tua; the prototype describes palm sap changing during storage. Source: printed pp. 33–36.	Distinguish fermentation as a process from a measured substance concentration; neither taste nor pH alone establishes ethanol content.	Content integration and collaboration: explain possible change and identify measurements needed to evaluate a claim.	Explain phenomena; E4–E6. Evaluate and design inquiry; I4–I6.
– Oi tua; fresh or lightly fermented sap is presented as beneficial in the prototype. This is an unverified claim.	Evaluate provided pH and ethanol data separately; a table of these variables cannot establish overall food safety.	Dialogue and reflection: compare a prior claim with data, identify uncertainty, and state what further evidence is required.	Interpret data and evidence; D4–D6. Give a conclusion bounded by the supplied observations.

Instruments

The scientific-literacy assessment comprises 18 short constructed-response items, with six items for each competency and scores from 0 to 2. A score of 2 requires the correct conclusion and its justification; 1 denotes a correct conclusion with incomplete reasoning; 0 denotes an incorrect or unsupported response. The total score is the item sum divided by 36 and multiplied by 100; each competency score is the corresponding sum divided by 12 and multiplied by 100. The framework follows OECD (2019). Table 2 uses E, I, and D identifiers for explaining phenomena, evaluating and designing inquiry, and interpreting data and evidence. The need to align evaluation with literacy goals is discussed by Pratiwi et al. (2019), while contextual instrument development is illustrated by Maulida and Sunarti (2022) and

Amalia et al. (2024). The module's self-evaluation exercises and the scored literacy assessment are treated as distinct instruments.

Instrument quality is examined through academic relevance ratings, item facility, corrected item–total correlations, and internal consistency. Item facility is the mean item score divided by two; corrected item–total correlations describe discrimination. Instrument-development guidance emphasizes construct definition, expert review, cognitive interviews, and pilot evaluation (Boateng et al., 2018). Contextual relevance also requires empirical item checks (Maulida & Sunarti, 2022). Table 4 reports the available diagnostics for the pilot and school samples. Inter-rater agreement and response-process evidence are not established by these coefficients. Cronbach's alpha describes internal consistency and does not, by itself, demonstrate unidimensionality or construct validity (Taber, 2018).

The module review form contains 30 items: six each for content, presentation, language, appearance, and ethnopedagogical alignment. Its organization retains the concern with clear, usable, self-directed materials in Daryanto (2013), with explicit alignment criteria added for this product. Three academic raters score item relevance from 1 (not relevant) to 5 (highly relevant). The panel separately rates the 18 test items and 16 questionnaire items. Student and teacher usability questionnaires each contain 16 positively worded items, four per domain: content, appearance, language, and perceived usefulness. Responses range from 1 (strongly disagree) to 4 (strongly agree). Domain and overall means are first calculated within each respondent; variability is then calculated across respondents. These items assess perceived usability, not observed engagement or learning. Two teacher respondents support descriptive reporting only.

Implementation procedures

The module allocates two 80-minute teaching meetings to food additives and addictive substances. The first meeting centres on bingka dolu and the second on oi tua. Activities follow self-identification, content integration, collaboration, dialogue, and reflection. The one-group design places the pretest before instruction and the posttest after instruction. The lesson structure and assessment sequence provide the instructional framework. Differences in classroom implementation are considered in interpreting the school results.

Dialogue and reflection ask students to state a claim, identify supporting evidence, and explain a limitation. Teachers can model a comparison of explanations before students reason independently. Instructional-design research supports selecting and combining guidance and inquiry according to the learning context (de Jong et al., 2024). The use of explicit reasoning prompts also draws on higher-order science questions studied by Muhibbuddin et al. (2023). These sources support the instructional rationale. No observation-based fidelity or engagement score is reported, so differences in implementation cannot be evaluated from the questionnaire means.

Data analysis

Aiken's V is calculated for each item as the sum of $(r - l_0)$, divided by $n(c - 1)$, where r is a rating, $l_0 = 1$, $n = 3$, and $c = 5$ (Aiken, 1985). Domain means average six item coefficients. The development rules are $V < .80$: revise; $.80 \leq V < .90$: examine comments and alignment; $V \geq .90$: comparatively strong agreement. These are operational revision rules, not universal significance thresholds. With three raters, item estimates are coarse and uncertain; confidence intervals and qualitative concerns should inform judgment (Penfield & Giacobbi, 2004). A high domain average cannot override an inaccurate statement or establish cultural authenticity.

Questionnaire means are described using prespecified equal-width intervals: 1.00–1.75 very poor; >1.75–2.50 poor; >2.50–3.25 good; and >3.25–4.00 very good. These labels describe response locations on the scale, not independently validated benchmarks. Means, sample standard deviations, ranges of respondent means, and item profiles are reported. The primary learning summary is the mean paired change with a 95% t-based confidence interval.

Paired t tests operate on change scores, with Shapiro–Wilk screening supplemented by inspection of the distribution. Wilcoxon signed-rank calculations with a normal approximation provide a sensitivity analysis; ties and zero changes are handled explicitly. Because only two school clusters are represented, neither analysis establishes population-wide precision or corrects school-level dependence.

The standardized paired change d_z is mean change divided by the sample standard deviation of change; it is not a between-group treatment effect (Lakens, 2013). A percentile bootstrap with 10,000 paired-record resamples gives an interval for the overall d_z. Competency tests are exploratory and their three p values are adjusted by Holm's procedure. Normalized gain is computed from group means as (posttest mean − pretest mean)/(100 − pretest mean). Conventional categories are low below .30, moderate from .30 to below .70, and high at least .70 (Hake, 1998). Normalized gain is retained for comparability but is not used as a causal or universal effectiveness threshold (Nissen et al., 2018).

Pilot and field-trial item scores are analysed separately using IBM SPSS Statistics version 25. School-specific summaries retain the 36/18 distribution, while overall and competency analyses use the 54 paired student records. Results are presented as aggregate scores, instrument diagnostics, and questionnaire summaries.

Ethical considerations

The participants are junior-high school students. Student outcomes and questionnaire responses are presented as group summaries. Student names and individual identifiers are not included in the article.

RESULTS AND DISCUSSION

Characteristics and revision of the developed module

The module contains two cultural readings and a five-stage teaching sequence. Figures 2 and 3 show its cover, contents, cultural reading, collaborative activity, and self-evaluation. The module contains 56 physical PDF pages, and printed page numbers differ from PDF page positions. Its contents page repeats the bingka dolu title for the second activity, although that activity concerns oi tua. Some questions emphasize familiar names or preferred health conclusions. These features identify a need to distinguish recall, cultural interpretation, and reasoning from evidence. Table 3 relates the product issues to the revisions presented in the article and priorities for further module development.



Figure 2. Actual source prototype cover and content organization

A Cultural reading
PDF page 31



B Collaborative activity
PDF page 41



C Self-evaluation
PDF page 48



Figure 3. Actual cultural reading, collaborative task, and self-evaluation page

Table 3. Product review and implications for revision

Evidence or review issue	Revision presented in this article	Priority for further module development
– Contents page repeats Activity 1 title for Activity 2.	The description identifies oi tua as Activity 2 and distinguishes PDF from printed page positions.	Unify the activity titles and page numbering throughout the module.
– Printed p. 27 treats natural colouring as inherently safer.	Table 2 and the cultural-mapping discussion replace origin-based safety claims with evaluation of substance, exposure, and evidence.	Use consistent evidence-based safety explanations in the readings and answer keys.
– Printed pp. 34–36 imply fresh/lightly fermented oi tua is necessarily healthy.	The scientific reconstruction distinguishes cultural description from evidence of composition or safety.	Strengthen cultural attribution and distinguish customary beliefs from tested health claims.
– Several prototype questions ask mainly for names of additives.	Table 2 specifies explanation, controlled-comparison, data-interpretation, and bounded-conclusion tasks.	Align learning tasks and scoring rubrics with the reasoning required by each competency.
– Item M25 has the lowest relevance rating, V = .75.	The discussion identifies traceable cultural sources as a specific revision priority.	Involve relevant practitioners in reviewing cultural sources and explanations.
– Five stages are present but do not by themselves demonstrate reasoning.	The instructional description specifies claims, evidence, and limitations in dialogue and reflection.	Examine student explanations and classroom activity in subsequent evaluations.

Interpretation. Contextual science teaching and local-wisdom modules have earlier precedents (Fatmala et al., 2017; Setiawan et al., 2017). The recent literature makes the design task more demanding. Adawiyah et al. (2022) connect ethnoscience materials with CRTT, while Mardiana et al. (2025) report expert validation of a local-wisdom e-module. Haen et al. (2025) report validity and practicality for Grade VIII Bakumpai materials; Martin et al. (2025) examine a pottery-related chemistry e-module and additional learning outcomes. These studies support the feasibility of culturally contextualized materials, but their different outcomes

should not be treated as interchangeable evidence. The present design makes the cultural source, scientific question, expected reasoning, and scoring criterion inspectable together. It also turns an unsupported safety claim into an opportunity to evaluate evidence and source credibility, consistent with Allchin (2022).

Academic review and instrument diagnostics

Finding. The 30-item module form produces an overall mean V of 0.917. Domain coefficients are 0.944 for content, 0.931 for presentation, 0.903 for language, 0.944 for appearance, and 0.861 for ethnopedagogical alignment. The lowest individual coefficient is 0.750 for M25, the item concerning traceable cultural sources. Thus, the overall mean rounded to 0.92 conceals a substantive revision need. The corresponding means are 0.940 for the literacy test and 0.938 for the questionnaire. These values summarize academic relevance judgments, not independent evidence that the instruments or cultural accounts are valid.

Table 4. Instrument diagnostics

Instrument and sample	Statistic	Result	Interpretive limit
Literacy test pilot, n = 60	Alpha; item facility range	0.917; 0.308–0.717	Internal consistency estimates alone do not establish construct or response-process validity.
Literacy test pilot, n = 60	Corrected item–total correlation range	0.453–0.751	Consider item diagnostics before retaining or revising items.
Literacy test pilot, n = 60	Competency alpha: E / I / D	0.805 / 0.761 / 0.775	Six-item subscales require cautious interpretation.
Literacy test field trial, n = 54	Pretest / posttest alpha	0.847 / 0.802	Internal consistency changes with the score distribution.
Student questionnaire pilot, n = 60	Overall / four-domain alpha	0.890 / 0.780, 0.669, 0.652, 0.609	Several domain coefficients are modest despite a higher overall coefficient.
Student questionnaire field trial, n = 54	Overall alpha	0.825	Perceived usability is distinct from engagement and achievement.
Teacher questionnaire, n = 2	Reliability coefficient	Not estimated	Two responses are insufficient for a stable reliability analysis.

The weakest alignment domain requires attention because it concerns the central contribution. Documenting sources, correcting inaccuracies, and linking tasks to observable reasoning are substantive revision priorities. Amalia et al. (2024) emphasize contextual assessment, while Maulida and Sunarti (2022) report that only 15 of 22 initial items met their empirical criteria despite favourable expert review. Expert ratings and item performance therefore provide different evidence. Here, three four-item questionnaire-domain alpha coefficients are below .70, despite a higher overall coefficient. The domains are interpreted descriptively until further evidence supports their structure.

Perceived usability

Table 5. Teacher and student questionnaire results

Domain	Teachers n = 2 Mean (SD)	Teacher range	Students n = 54 Mean (SD)	Student range
Content	3.75 (0.00)	3.75–3.75	3.48 (0.39)	2.50–4.00
Appearance	4.00 (0.00)	4.00–4.00	3.63 (0.34)	2.75–4.00
Language	3.75 (0.00)	3.75–3.75	3.50 (0.38)	2.50–4.00

Domain	Teachers n = 2 Mean (SD)	Teacher range	Students n = 54 Mean (SD)	Student range
Perceived usefulness	3.88 (0.18)	3.75–4.00	3.62 (0.36)	2.50–4.00
Overall	3.84 (0.04)	3.81–3.88	3.56 (0.30)	2.88–4.00

Students average 3.56 (SD = 0.30) overall, with content rated lower than appearance. The teacher mean is 3.84 (SD = 0.04), but two teacher responses cannot characterize the wider teaching population. Reported student item means range from 3.44 to 3.69. Variation remains visible even though all domain means receive the same broad category label. The domain scores, dispersion, and respondent ranges in Table 5 therefore provide more information than the category alone.

Interpretation. Favorable usefulness ratings describe perceptions, not observed engagement, efficient classroom use, or stronger cultural identity. Ethnochemistry-media research reports user responses alongside literacy tests (Heliawati et al., 2022), demonstrating the need to distinguish those measures. Cultural identity in Andayani et al. (2023) and cultural-preservation attitudes in Jihannita et al. (2024) likewise involve separate outcomes. Evidence from ethnoscience-based PjBL also includes both comparative tests (Rusmansyah et al., 2023) and qualitative implementation accounts (Hidayah et al., 2024). Their findings cannot be reduced to a single claim that positive reactions cause learning. For this product, classroom observations and student explanations would clarify how the activities are used and where support is needed.

Preliminary learning change

Table 6. Paired literacy scores by school and overall

School	n	Pretest Mean (SD)	Posttest Mean (SD)	Paired change 95% CI	N-gain
School A	36	50.08 (19.95)	76.23 (14.82)	26.16 [21.33, 30.99]	0.52
School B	18	50.15 (17.32)	69.75 (15.48)	19.60 [13.13, 26.07]	0.39
Overall	54	50.10 (18.95)	74.07 (15.21)	23.97 [20.12, 27.82]	0.48

Finding. Overall scores increase from 50.10 (SD = 18.95) to 74.07 (SD = 15.21). Pretest scores range from 11.11 to 91.67 and posttest scores from 36.11 to 100.00. The mean paired change is 23.97 points (SD of change = 14.09), 95% CI [20.12, 27.82], $t(53) = 12.50$, $p < .001$. Shapiro–Wilk screening of total change yields $W = 0.985$, $p = .754$; Wilcoxon sensitivity analysis gives $W = 13.00$, $p < .001$. Fifty-two students improve and two decline. The paired standardized change is $d_z = 1.70$, with a participant-level bootstrap 95% interval [1.35, 2.21]. These inferential quantities describe the participant records under working independence assumptions and do not establish a causal effect of the module.

Table 7. Change by scientific-literacy competency

Competency	Pretest Mean (SD)	Posttest Mean (SD)	Change 95% CI	N-gain	d_z
Explain phenomena	54.01 (22.59)	82.41 (15.67)	28.40 [23.71, 33.08]	0.62	1.65
Evaluate and design inquiry	45.37 (21.46)	64.04 (20.91)	18.67 [12.54, 24.81]	0.34	0.83
Interpret data and evidence	50.93 (21.21)	75.77 (17.46)	24.85 [19.21, 30.48]	0.51	1.20

Interpretation. Inquiry has the smallest gain in the dataset, highlighting a competency-level pattern that is not fully represented by the total score. Ayuso Fernández et al. (2022) document difficulties with hypotheses, experimental design, and conclusions in secondary education, making these reasonable targets for diagnostic assessment. The short lessons and the prototype's emphasis on reading and discussion offer design-based hypotheses for a competency difference, not demonstrated mechanisms. Published contextual studies report learning changes in different settings: Dewi et al. (2021) examine university chemistry, and Wahyuni et al. (2025) examine a Grade VII textbook based on Grenden Village. Their results are design benchmarks, not confirmation of the present results. Coding of student responses

should examine whether students identify variables, request measurements, and recognize conclusions that exceed the evidence.

School A's normalized gain is .52 and School B's is .39, as reported in Table 6. Their pretest means are similar (50.08 and 50.15), while their posttest means are 76.23 and 69.75. The difference is descriptive; unequal sample sizes, exposure, teacher practices, and instrument functioning remain possible contextual explanations for future investigation. The present analysis does not identify a causal explanation. Normalized gain is sensitive to its denominator (Nissen et al., 2018), and its interpretation differs from that of standardized paired change. A moderate gain and a large paired dz can therefore coexist without establishing an instructional treatment effect.

Contribution and limitations

The study's contribution is a documented connection between cultural content, scientific reasoning tasks, and assessment competencies. The results show that improvement is uneven: explaining phenomena has the largest normalized gain, while evaluating and designing inquiry has the smallest. This pattern directs attention to the demands of inquiry tasks, but it does not establish that a particular cultural activity produced the difference. Jufrida et al. (2024) document Grade VIII ethnoscience implementation, while Syahdani and Tyas (2025) evaluate traditional-food project learning using a comparison condition. Uslan et al. (2024) and Verawati and Wahyudi (2024) offer further comparative contexts. These studies support a next step that tests the present mapping with classroom evidence, comparison groups, and delayed assessment. Longer-term evaluation is also relevant to the literature discussed by Yanto et al. (2025). Broader critical-literacy goals (Guerrero & Sjöström, 2025) would require outcomes beyond the three competencies assessed here.

The study uses small, unequal school samples in a restricted geographic setting and has no control group. Testing effects, maturation, concurrent instruction, teacher differences, regression to the mean, and repeated exposure remain alternative explanations for pretest–posttest change. Two school clusters do not support stable school-level causal or multilevel inference. Self-reported usability may reflect acquiescence or favourable responses to teachers. The reported internal-consistency coefficients do not replace scoring-agreement or response-process evidence. Cultural confirmation and observation-based implementation measures are also needed to evaluate how the module operates in practice. The findings are therefore limited to the reported sample and the outcomes measured.

CONCLUSION

The module connects bingka dolu and oi tua to scientific-literacy tasks through an explicit mapping of cultural content, science concepts, teaching activities, and assessment. The academic review and usability results support further development, while the pretest–posttest findings describe preliminary learning change. The smaller increase in inquiry identifies an area for refinement and further investigation. These results do not establish comparative effectiveness or justify generalization beyond the two schools. Cultural confirmation, instrument quality, and classroom implementation remain distinct requirements for evaluating the module's educational contribution.

RECOMMENDATION

Further development should strengthen inquiry tasks, document the sources of cultural explanations, and record revisions to the learner module and assessment. Subsequent evaluation should include independent scoring checks, classroom implementation records, multiple classes in each condition, and a suitable comparison design. Delayed and unfamiliar-context tasks would help assess retention and transfer. School differences should be investigated using documented instructional and participant characteristics rather than inferred from gain scores alone. Dissemination should follow product verification and include evaluation of adoption and classroom use.

ACKNOWLEDGMENTS
The research was supported by a research grant from the Ministry of Higher Education, Science and Technology through the 2026 Basic Research program. We are very grateful for the financial support provided. and all those who have supported the running of this research. Appreciation is also extended to the participating schools, teachers, students, and expert validators for their involvement in the research, the evaluation of learning products, and field implementation. The author also acknowledges the institutional support provided by Nggusuwaru University throughout the research process.

FUNDING INFORMATION
This research received funding from the Ministry of Higher Education, Science, and Technology through the 2026 Basic Research program.

AUTHOR CONTRIBUTIONS STATEMENT

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Rabiatul Adawiyah	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	
Nurzakiyyah						✓					✓	✓		
Istin Fitriana Aziza			✓		✓			✓			✓		✓	
Norma Juniati	✓	✓		✓					✓	✓		✓		

CONFLICT OF INTEREST STATEMENT
The authors declare that they have no known competing financial or personal interests that could have influenced the work reported in this article.

DATA AVAILABILITY
The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

Adawiyah, R., Andayani, Y., Rudyat, L., & Savalas, T. (2022). Pengembangan Modul Kimia Etnosains Terintegrasi Model Culturally Responsive Transformative Teaching (CRTT). *Journal of Classroom Action Research*, *4*(4). https://doi.org/10.29303/jcar.v4i4.2395

Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement, 45*(1), 131–142. https://doi.org/10.1177/0013164485451012

Allchin, D. (2022). Who speaks for science? *Science & Education*, *31*(6), 1475–1492. https://doi.org/10.1007/s11191-021-00257-4

Alwasilah, A. C., Suryadi, K., & Karyono, T. (2009). *Etnopedagogi: Landasan Praktek Pendidikan dan Pendidikan Guru*. PT Kiblat Buku Utama.

Amalia, D. V., Ilhami, A., Fuadiyah, S., & Kusumanegara, A. (2024). Development of a scientific literacy instrument based on Riau Malay ethnoscience in science subjects. *PEDAGOGIK: Jurnal Pendidikan, 11*(1), 1–18. https://doi.org/10.33650/pjp.v11i1.6382

Andayani, Y., Anwar, Y. A. S., Hakim, A., & Hidayanti, E. (2023). Ethnopedagogy approach in chemistry learning as an effort to strengthen students' cultural identity. In *Proceedings of the 1st Nusa Tenggara International Conference on Chemistry (NiTRIC 2022)* (pp. 163–173). Atlantis Press. https://doi.org/10.2991/978-94-6463-130-2_17

Ayu Fatmala, S., Sujana, A., & Maulana. (2017). Pembelajaran Kontekstual Untuk Meningkatkan Literasi Sains Siswa Sd Kelas V Pada Materi Peristiwa Alam. *Jurnal Pena Ilmiah*, *2*(1).

Ayuso Fernández, G. E., López-Banet, L., & Ruiz-Vidal, A. (2022). Students' performance in the scientific skills during secondary education. *Eurasia Journal of Mathematics, Science and Technology Education*, *18*(10), Article em2165. https://doi.org/10.29333/ejmste/12444

Bahtiar, Zulkarnaen, Ibrahim, Fadli, W., & Mustofa, H. A. (2026). How science teachers translate Indigenous knowledge into classroom science: A qualitative study in Lombok, Indonesia. *International Journal of Ethnoscience and Technology in Education*, *3*(2), 177–215. https://doi.org/10.33394/ijete.v3i2.22482

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

Chibuye, B., & Singh, I. S. (2024). Integration of local knowledge in the secondary school chemistry curriculum – A few examples of ethno-chemistry from Zambia. *Heliyon*, *10*(7), Article e29174. https://doi.org/10.1016/j.heliyon.2024.e29174

Daryanto. (2013). *Menyusun Modul Bahan Ajar Untuk Persiapan Guru Dalam Mengajar*. Penerbit Gava Media.

de Jong, T., Lazonder, A. W., Chinn, C. A., Fischer, F., Gobert, J., Hmelo-Silver, C. E., Koedinger, K. R., Krajcik, J. S., Kyza, E. A., Linn, M. C., Pedaste, M., Scheiter, K., & Zacharia, Z. C. (2024). Beyond inquiry or direct instruction: Pressing issues for designing impactful science learning opportunities. *Educational Research Review*, *44*, Article 100623. https://doi.org/10.1016/j.edurev.2024.100623

Dewi, C. A., Erna, M., Martini, Haris, I., & Kundera, I. N. (2021). The effect of contextual collaborative learning based ethnoscience to increase student's scientific literacy ability. *Journal of Turkish Science Education*, *18*(3), 525–541. https://doi.org/10.36681/tused.2021.88

Fatmala, S. A., Sujana, A., & Maulana. (2017). Pembelajaran kontekstual untuk meningkatkan literasi sains siswa SD kelas V pada materi peristiwa alam. *Jurnal Pena Ilmiah*, *2*(1), 211–220. https://ejournal.upi.edu/index.php/penailmiah/article/view/10656

Fernández, G. E. A., López-Banet, L., & Ruiz-Vidal, A. (2022). Students' performance in the scientific skills during secondary education. *Eurasia Journal of Mathematics, Science and Technology Education*, *18*(10). https://doi.org/10.29333/ejmste/12444

Fuadi, H., Robbia, A. Z., Jamaluddin, J., & Jufri, A. W. (2020). Analisis faktor penyebab rendahnya kemampuan literasi sains peserta didik. *Jurnal Ilmiah Profesi Pendidikan, 5*(2), 108–116. https://doi.org/10.29303/jipp.v5i2.122

Guerrero, G., & Sjöström, J. (2025). Critical scientific and environmental literacies: A systematic and critical review. *Studies in Science Education*, *61*(1), 41–87. https://doi.org/10.1080/03057267.2024.2344988

Haen, H., Fathul Zannah, & Tazkiyatunnafs Elhawwa. (2025). Pengembangan Bahan Ajar Bakar (Berbasis Kearifan Lokal Bakumpai) pada Pelajaran IPA untuk Meningkatkan Kemampuan Literasi Sains Siswa SMP Kelas VIII. *JURNAL PENDIDIKAN MIPA*, *15*(4), 1829–1836. https://doi.org/10.37630/jpm.v15i4.3665

Haen, Zannah, F., & Elhawwa, T. (2025). Pengembangan bahan ajar Bakar (berbasis kearifan lokal Bakumpai) pada pelajaran IPA untuk meningkatkan kemampuan literasi sains siswa SMP kelas VIII. *Jurnal Pendidikan MIPA*, *15*(4), 1829–1836. https://doi.org/10.37630/jpm.v15i4.3665

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics, 66*(1), 64–74. https://doi.org/10.1119/1.18809

Heliawati, L., Lidiawati, L., Adriansyah, P. N. A., & Herlina, E. (2022). Ethnochemistry-based Adobe Flash learning media using indigenous knowledge to improve students' scientific literacy. *Jurnal Pendidikan IPA Indonesia*, *11*(2), 271–281. https://doi.org/10.15294/jpii.v11i2.34859

Hidayah, A., Rokhimawan, M. A., & Suherman, R. (2024). Implementation of ethnoscience-based PjBL on science literacy learning outcomes. *Journal of Innovation in Educational and Cultural Research, 5*(3), 398–407. https://doi.org/10.46843/jiecr.v5i3.1278

Hikmah, N., Yohandri, Arsih, F., Azhar, M., & Razak, A. (2025). Uncovering The Potential Of Ethnoscience In Science Learning To Improve Students' Literacy: A Systematic-Literature Review (2014–2024). *Jurnal Pendidikan IPA Indonesia*, *14*(3), 550–562. https://doi.org/10.15294/jpii.v14i3.19591

Jihannita, J., Fadly, W., Ekapti, R. F., Luthfiana, D., & Widowati, A. (2024). The development of science module integrated with ethnoscience of Singo Barong mask to improve scientific literacy and cultural preservation attitudes. *Journal of Innovation in Educational and Cultural Research*, *5*(2), 356–363. https://doi.org/10.46843/jiecr.v5i2.790

Jufrida, Kurniawan, W., & Basuki, F. R. (2024). Ethnoscience learning: how do teacher implementing to increase scientific literacy in junior high school. *International Journal of Evaluation and Research in Education* , *13*(3), 1719–1730. https://doi.org/10.11591/ijere.v13i3.26180

Kurniawati, E., Waluyo, H. J., & Slamet, Y. B. M. (2013). *Developing a Model of Thematic Speaking Learning Materials Using Savi Approach (Somatic, Auditory, Visual, Intellectual) In Senior High School in Sambas Regency, West Kalimantan Province, Indonesia*. https://api.semanticscholar.org/CorpusID:131577438

Ladson-Billings, G. (1995). Toward a theory of culturally relevant pedagogy. *American Educational Research Journal, 32*(3), 465–491. https://doi.org/10.3102/00028312032003465

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology, 4*, Article 863. https://doi.org/10.3389/fpsyg.2013.00863

Mardiana, A., Waluyo, J., Dwi, P., Putra, A., & Artikel, R. (2025). Science E-Module Based On Local Wisdom To Improve Scientific Literacy And Critical Thinking Skills Of Junior High School Students Info Artikel Abstrak. *Paedagoria : Jurnal Kajian, Penelitian Dan Pengembangan Kependidikan*, *16*(3), 325–335. https://doi.org/10.31764

Martin, N., Andayani, Y., & Hakim, A. (2025). Pengembangan e-modul ajar terintegrasi ekowisata gerabah dengan model pembelajaran etnopedagogi untuk melatih kemandirian belajar dan berpikir kritis peserta didik. *Journal of Classroom Action Research*, *7*(4), 1696–1701. https://doi.org/10.29303/jcar.v7i4.13610

Maulida, F., & Sunarti, T. (2022). Pengembangan instrumen tes literasi sains berbasis kearifan lokal di Kabupaten Lamongan. *ORBITA: Jurnal Hasil Kajian, Inovasi, dan Aplikasi Pendidikan Fisika*, *8*(1), 52–65. https://doi.org/10.31764/orbita.v8i1.8337

Muhibbuddin, Artika, W., & Nurmaliah, C. (2023). Improving critical thinking skills through higher order thinking skills (HOTS)-based science. *International Journal of Instruction*, *16*(4), 283–296. https://doi.org/10.29333/iji.2023.16417a

Murtiningsih, I., Untari, A. D., & Luthfi, Z. F. (2024). Membangun Karakter Bangsa: Peran Pendidikan Kewarganegaraan dalam Pembentukan Generasi Berkualitas. *JURNAL ILMIAH KAJIAN PENDIDIKAN KEWARGANEGARAA*.

Nasution, J. E. (2024). Perencanaan Pembelajaran Pai Berbasis Higher Order Thinking Skills (Hots) Di Era Society 5.0: Strategi Dan Implementasi. *Jurnal Pendidikan Dan Pemikiran*, *19*(2).

Nissen, J. M., Talbot, R. M., Thompson, A. N., & Van Dusen, B. (2018). Comparison of normalized gain and Cohen's d for analyzing gains on concept inventories. *Physical Review Physics Education Research, 14*(1), Article 010115. https://doi.org/10.1103/PhysRevPhysEducRes.14.010115

OECD. PISA 2018 results (volume I): what students know and can do. Paris: OECD Publishing; 2019

Organisation for Economic Co-operation and Development. (2019). *PISA 2018 assessment and analytical framework.* OECD Publishing. https://doi.org/10.1787/b25efab8-en

PISA 2022 results (Volume I): The state of learning and equity in education. OECD Publishing. https://doi.org/10.1787/53f23881-en

Penfield, R. D., & Giacobbi, P. R., Jr. (2004). Applying a score confidence interval to Aiken's item content-relevance index. *Measurement in Physical Education and Exercise Science,*

8(4), 213–225. https://doi.org/10.1207/s15327841mpee0804_3

Pratiwi, S. N., Cari, & Aminah, N. S. (2019). Pembelajaran IPA abad 21 dengan literasi sains siswa. *Jurnal Materi dan Pembelajaran Fisika*, *9*(1), 34–42. https://jurnal.uns.ac.id/jmpf/article/view/31612

Putra, B. P., & Wahyuni, S. (2025). Integrasi kearifan lokal dalam pembelajaran IPA untuk meningkatkan literasi sains siswa: Kajian literatur. *Kalam Cendekia: Jurnal Ilmiah Kependidikan*, *13*(2), 1431–1439. https://doi.org/10.20961/jkc.v13i2.103518

Rahmawati, Y., Ridwan, A., Faustine, S., Syarah, S., Ibrahim, I., & Mawarni, P. C. (2020). Pengembangan literasi sains dan identitas budaya siswa melalui pendekatan etno-pedagogi dalam pembelajaran sains. *Edusains*, *12*(1), 54–63. https://doi.org/10.15408/es.v12i1.12428

Rusmansyah, Leny, & Sofia, H. N. (2023). Improving students' scientific literacy and cognitive learning outcomes through ethnoscience-based PjBL model. *Journal of Innovation in Educational and Cultural Research, 4*(1), 1–9. https://doi.org/10.46843/jiecr.v4i1.382

Sari, F. P., Maryati, & Wilujeng, I. (2023). Ethnoscience studies analysis and their integration in science learning: Literature review. *Jurnal Penelitian Pendidikan IPA, 9*(3), 1135–1142. https://doi.org/10.29303/jppipa.v9i3.2044

Setiawan, B., Innatesari, D. K., Sabtiawan, W. B., & Sudarmin, S. (2017). The development of local wisdom-based natural science module to improve science literation of students. *Jurnal Pendidikan IPA Indonesia*, *6*(1), 49–54. https://doi.org/10.15294/jpii.v6i1.9595

Smith, T., Avraamidou, L., & Adams, J. D. (2022). Culturally relevant/responsive and sustaining pedagogies in science education: Theoretical perspectives and curriculum implications. *Cultural Studies of Science Education*, *17*(3), 637–660. https://doi.org/10.1007/s11422-021-10082-4

Syahdani, E. P., & Tyas, R. A. (2025). Project-based ethnoscience learning: Elevating scientific literacy and communication through nutrient analysis of 'Jadah Tempe'. *Journal of Science Learning, 8*(2), 222–233. https://doi.org/10.17509/jsl.v8i2.82148

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

Tasquier, G., Knain, E., & Jornet, A. (2022). Scientific literacies for change making: Equipping the young to tackle current societal challenges. *Frontiers in Education*, *7*, Article 689329. https://doi.org/10.3389/feduc.2022.689329

Thiagarajan, S., Semmel, D. S., & Semmel, M. I. (1974). *Instructional Development for Training Teachers of Exceptional Children: A Sourcebook*. Leadership Training Institute/Special Education, University of Minnesota. https://books.google.co.id/books?id=K-2YpwAACAAJ

Uslan, Abdullah, N., Imami, M. K. W., & Aiman, U. (2024). The Effectiveness Of The Local Knowledge-Based Module (Lkbm) To Improve Students' Scientific Literacy And Thinking Skills. *Jurnal Pendidikan IPA Indonesia*, *13*(1), 147–161. https://doi.org/10.15294/jpii.v13i1.47561

Verawati, N. N. S. P., & Wahyudi. (2024). Raising the issue of local wisdom in science learning and its impact on increasing students' scientific literacy. *International Journal of Ethnoscience and Technology in Education*, *1*(1), 42–54. https://doi.org/10.33394/ijete.v1i1.10881

Wahyuni, S., Sukma, F. D., Rusdianto, R., & Suhada, I. (2025). Effectiveness A Science Textbook Based on Local Wisdom of Grenden Village to Improve the Science Literacy of Junior High School Students. *Unnes Science Education Journal*, *14*(1), 79–87. https://doi.org/10.15294/usej.v13i1.20924

WEFUSA. Unlocking the potential of the smart grid. AIP Conference Proceedings. 2015;1702. doi:10.1063/1.4938795.

World Health Organization. (2023). *Food additives.* https://www.who.int/news-room/fact-sheets/detail/food-additives

Yanto, N., Sari, N. I., & Yahya. (2025). Exploring scientific literacy in science classrooms: A literature study. *Venn: Journal of Sustainable Innovation on Education, Mathematics and Natural Sciences*, *4*(3), 164–173. https://doi.org/10.53696/venn.v4i3.292

Yusmar, F., & Fadilah, R. E. (2023). Analisis rendahnya literasi sains peserta didik indonesia: hasil pisa dan faktor penyebab. *Lensa (Lentera Sains): Jurnal Pendidikan IPA*, *13*(1), 11–19. https://doi.org/10.24929/lensa.v13i1.283