

WOMEN, EDUCATION, AND SOCIAL WELFARE
VOL. 3 NO. 4 (2026)

ISSN: 3064-2469

Evaluating Local-Wisdom-Based Teacher Professional Learning in Sumedang: Literacy – Numeracy Knowledge Gains and Inclusion Challenges

Ika Mustika Sari*✉, Winny Liliawati✉, Ridwan Efendi✉, Nuryana✉, Sumarna Surapranata✉, Suhendiana Noor✉, Intan Setiawati✉, M Gina Nugraha✉, Selly Feranie✉, and Muhamad Nur Dandy✉

To cite this article: I. M. Sari *et al.*, "Evaluating Local-Wisdom-Based Teacher Professional Learning in Sumedang: Literacy – Numeracy Knowledge Gains and Inclusion Challenges," *Women, Educ. Soc. Welf.*, vol. 3, no. 4, pp. 1346–1374, 2026. https://doi.org/10.70211/wesw.v3i4.849

Published online: November 06, 2026

Submit your article to this journal

View crossmark data

Full Terms & Conditions of access and use can be found at
https://journal.wiseedu.co.id/index.php/jiswel/about

OPEN ACCESS

https://doi.org/10.70211/wesw.v3i4.849

Women, Education, and Social Welfare Vol 3 No 4 (2026)

Check for updates

Research Article

Evaluating Local-Wisdom-Based Teacher Professional Learning in Sumedang: Literacy–Numeracy Knowledge Gains and Inclusion Challenges

Ika Mustika Sari*, Winny Liliawati, Ridwan Efendi, Nuryana, Sumarna Surapranata, Suhendiana Noor, Intan Setiawati, M Gina Nugraha, Selly Feranie, and Muhamad Nur Dandy

Received: September 23, 2026	Revised: October 07, 2026	Accepted: November 02, 2026	Online: November 06, 2026

Abstract

This programme evaluation examined knowledge-test changes, instructional products, and classroom implementation during a local-wisdom-based professional learning initiative for primary-school teachers in Sumedang Regency, Indonesia. From 29 August to 19 September 2026, 30 teachers from 26 public schools joined a planned 50-lesson-hour training and mentoring programme; 28 contributed complete paired assessments. The matched teachers' mean scores increased from 68.0 to 88.4 (mean difference 20.4; $t(27) = 7.65$, $p < .001$; paired-sample $d_z = 1.45$); the aggregate gain calculated from the two means was 0.64. Six documented deep-learning lesson packages containing worksheets and Minimum Competency Assessment-style tasks were produced. One Grade VI open lesson and questionnaires from 25 pupils yielded a reported composite response score of 77.4%, reflecting perceptions rather than demonstrated learning gains. Reported classroom difficulties included decimal place-value errors, the translation of word problems into mathematical expressions, unequal group participation, and limited explicit treatment of cultural values. With no comparison group, unverified item-level records and limited observation coverage, findings cannot establish causal impact, equity effects, or sustained pupil outcomes. The study offers an evidence-bounded account of knowledge learning, material development, and implementation constraints that can inform more rigorous and inclusive teacher development.

Keywords: Deep Learning; Literacy; Local Wisdom; Numeracy; Teacher Competence.

Publisher's Note:
WISE Pendidikan Indonesia stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright:
©
2026 by the author(s).
License WISE Pendidikan Indonesia, Bandar Lampung, Indonesia. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY 4.0) license (https://creativecommons.org/licenses/by/4.0/).



INTRODUCTION

Literacy and numeracy enable pupils to interpret information, reason quantitatively and apply knowledge in everyday situations. Indonesia continues to face substantial challenges in these areas: in PISA 2022, approximately 18% of 15-year-olds reached the minimum proficiency level in mathematics and 25% in reading, compared with OECD averages of 69% and 74%, respectively [1], [2]. These results provide national context, rather than a direct measure of primary pupils' achievement. In Sumedang Regency, an earlier school-transformation initiative reported an increase in the literacy index from 65.03 to 70.75 and in the numeracy index from 42.31 to 53.61 [3]. Although based on a separate programme, those findings reinforce the importance of sustained, context-sensitive support for foundational learning.

Teachers' capacity to design and assess meaningful learning experiences is central to this effort. Research on professional development highlights coherent content, opportunities for practice, feedback and alignment with teaching conditions [4], [5], [6], [7]. Coaching and professional learning communities can also support instructional improvement, although evidence of increased teacher knowledge should be distinguished from changes in classroom practice and pupil achievement [8], [9]. Accordingly, professional learning needs to connect conceptual understanding with the design, enactment and reflective refinement of lessons.

The Sumedang programme brought together Indonesia's Deep Learning (Pembelajaran Mendalam) approach—understood here as mindful, meaningful and joyful pedagogy rather than artificial intelligence—with Atikan Sunda (a Sundanese educational approach) and Pancawaluya (a Sundanese character-education framework). Familiar local foods, including Ubi Cilembu, Tahu Sumedang, Opak and Rengginang, provided potential contexts for reading and numeracy tasks. Culturally responsive and sustaining perspectives suggest that local experiences can make learning relevant when their use respects pupils' different backgrounds [10], [11], [12]. Nevertheless, such contexts require explicit links to disciplinary concepts and well-designed learning materials [13], [14], supported by assessment and instructional scaffolding [15], [16].

Contextual relevance is only one condition of meaningful participation. Teachers must also ensure that pupils can understand tasks, contribute ideas and receive appropriate support, including those less likely to lead group activities [17], [18], [19]. Research on programme implementation and teacher learning emphasises the value of examining how an instructional design is enacted, not merely whether training was delivered [20], [21]. The Sumedang activities therefore combined diagnostic assessment, the MERDEKA learning sequence, mentored lesson preparation, peer feedback and a plan–do–see open lesson to connect professional learning with initial classroom application.

This approach is consistent with accounts of teacher learning as a process shaped by instructional activities and school contexts [22], [23], [24]. Evidence on effective professional development similarly points to purposeful design, curricular relevance and opportunities to work with teaching materials [25], [26], [27], while reviews distinguish teacher characteristics, instructional practices and pupil outcomes as related but separate domains [28], [29]. In mathematics education, the social relevance of a context should not obscure conceptual precision [30]; clear feedback and guidance remain important for demanding tasks [31], [32].

These insights support an evaluation that considers teachers' assessment results alongside the quality of their learning products and their initial classroom use.

Indonesian studies have documented possibilities and challenges in connecting local wisdom with literacy, numeracy, instructional media and Deep Learning [33], [34], [35], [36], [37], [38]. Yet the presence of cultural examples in lesson plans does not, by itself, show how teachers translate them into purposeful learning activities or how pupils experience those activities. The more immediate question is how a short, practice-oriented programme supports teachers' instructional decisions and what difficulties become visible when those decisions are tried in class.

Against this background, the present study evaluates a local-wisdom-based literacy and numeracy programme for elementary teachers in Sumedang Regency. It examines initial learning needs, changes in teacher knowledge-assessment scores, instructional packages developed through mentoring, classroom observations and pupils' immediate responses. Attention to participation is especially important because active involvement cannot be inferred from the overall attractiveness of a lesson [39]. By examining these complementary sources of evidence, the study identifies both short-term achievements and areas requiring further instructional support, without treating the programme as proof of sustained improvements in pupil achievement.

METHODS

Research Design and Evaluation Framework

This is a descriptive community-programme evaluation with a single-group, within-teacher pretest/posttest component, artifact review and one observed Grade VI open lesson. The programme is the intervention context; the 28 complete pairs are the unit for the paired teacher test, 30 attendees and 26 schools describe programme reach, the six documented Deep Learning lesson-plan packages (Rencana Pembelajaran Mendalam; RPM) are artefact units, and the 25 pupils are questionnaire respondents from one class. No untreated comparison group, school-level model, measured pupil test change or formal equity evaluation is documented [5], [40].

Programme needs informed the training content; teachers' conceptual learning was assessed through a 25-item test; completed lesson packages documented instructional outputs; the open lesson offered an opportunity to observe initial classroom enactment; and pupil questionnaires captured immediate responses. Together, these stages form a practical sequence for programme evaluation, without establishing that any particular activity caused the observed changes. Within a Kirkpatrick-informed interpretation, questionnaire ratings represent reaction, test results indicate learning, and the observed lesson provides limited evidence of behaviour in practice. Sustained transfer and longer-term pupil outcomes were not assessed [20], [24].

The community service programme was implemented using a participatory, hands-on, and continuous mentoring approach. Teachers were not positioned merely as recipients of instructional materials but as designers and implementers of learning who identified local wisdom in their respective environments, developed learning materials, implemented them in the classroom, and engaged in reflection with peers. This approach was designed to connect

the strengthening of literacy and numeracy competencies with the implementation of contextual learning based on local wisdom.

Participants, Setting, and Programme Schedule

The programme initially allocated 50 teacher places, with 30 teachers ultimately attending (60.0% of planned capacity). Twenty-eight respondents completed the diagnostic questionnaire on 28 August 2026 and had paired pretest–posttest records; all 30 reportedly took the pretest on 29 August. The diagnostic questionnaire was administered before the formal programme period of 29 August–19 September 2026, and the posttest was administered on 17 September. The available records do not clarify the nomination process, the reasons some places remained unfilled, the profiles of the two participants without complete pairs, school-level recruitment coverage or individual attendance hours. The reported 50 lesson hours (jam pelajaran; JP) refer to the planned programme, while the duration of each JP and individual training exposure could not be independently established.

The programme combined face-to-face sessions at Puri Khatulistiwa Hotel in Jatinangor and a public primary school in Sumedang Regency with online meetings and independent school-based assignments. Classroom trials and the open lesson took place at participating schools. The pretest was administered on 29 August 2026, and the posttest on 17 September 2026.

Training Implementation and Mentoring Stages

The programme consisted of six stages: preparation, diagnostic assessment, teacher training, mentoring, learning practice (open lesson), and reflection. The implementation flow of these six stages is presented in Figure 1.

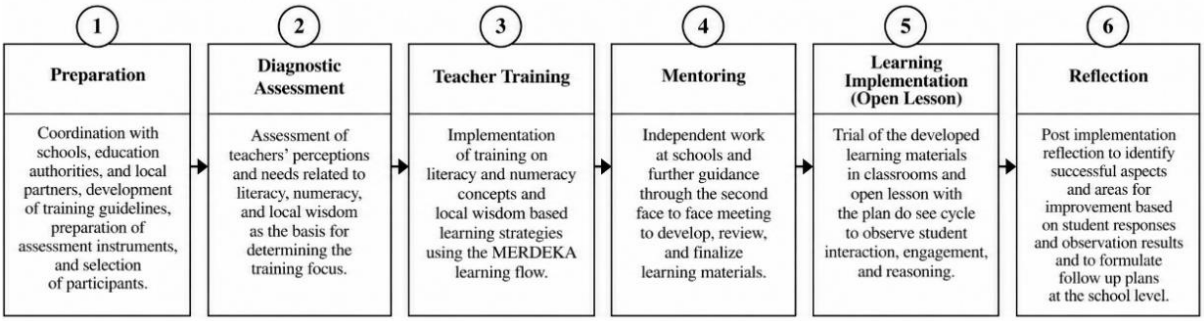


Figure 1. Literacy and Numeracy Program Implementation Flow

During the preparation stage, the team coordinated with schools, the Education Office, and local partners; developed training guidelines; prepared assessment instruments; and determined the participants. The diagnostic assessment stage was conducted through a teacher perception and needs questionnaire to map aspects of literacy, numeracy, and understanding of local wisdom. The assessment results were used as the basis for adjusting the emphasis of training materials and learning examples.

The teacher training stage used the Literacy and Numeracy Movement (Gerakan Literasi dan Numerasi; GLN) module following the MERDEKA sequence, namely Materi dan Misi (Materials and Missions), Eksplorasi dan Elaborasi (Exploration and Elaboration), Refleksi dan Rekoneksi (Reflection and Reconnection), Data dan Demonstrasi (Data and

Demonstration), Ekspresi dan Evaluasi (Expression and Evaluation), Komunikasi dan Kolaborasi (Communication and Collaboration), and Anjuran dan Aksi (Recommendations and Action). The four main topics covered were the Literacy and Numeracy Movement, Atikan Sunda Local Wisdom, the Minimum Competency Assessment (Asesmen Kompetensi Minimum; AKM), and Deep Learning, with the development of RPMs as the main product. The relationship among the training stages is visualized in Figure 2; the distribution of the 50 lesson hours (JP) is specified in the methodological tables below.

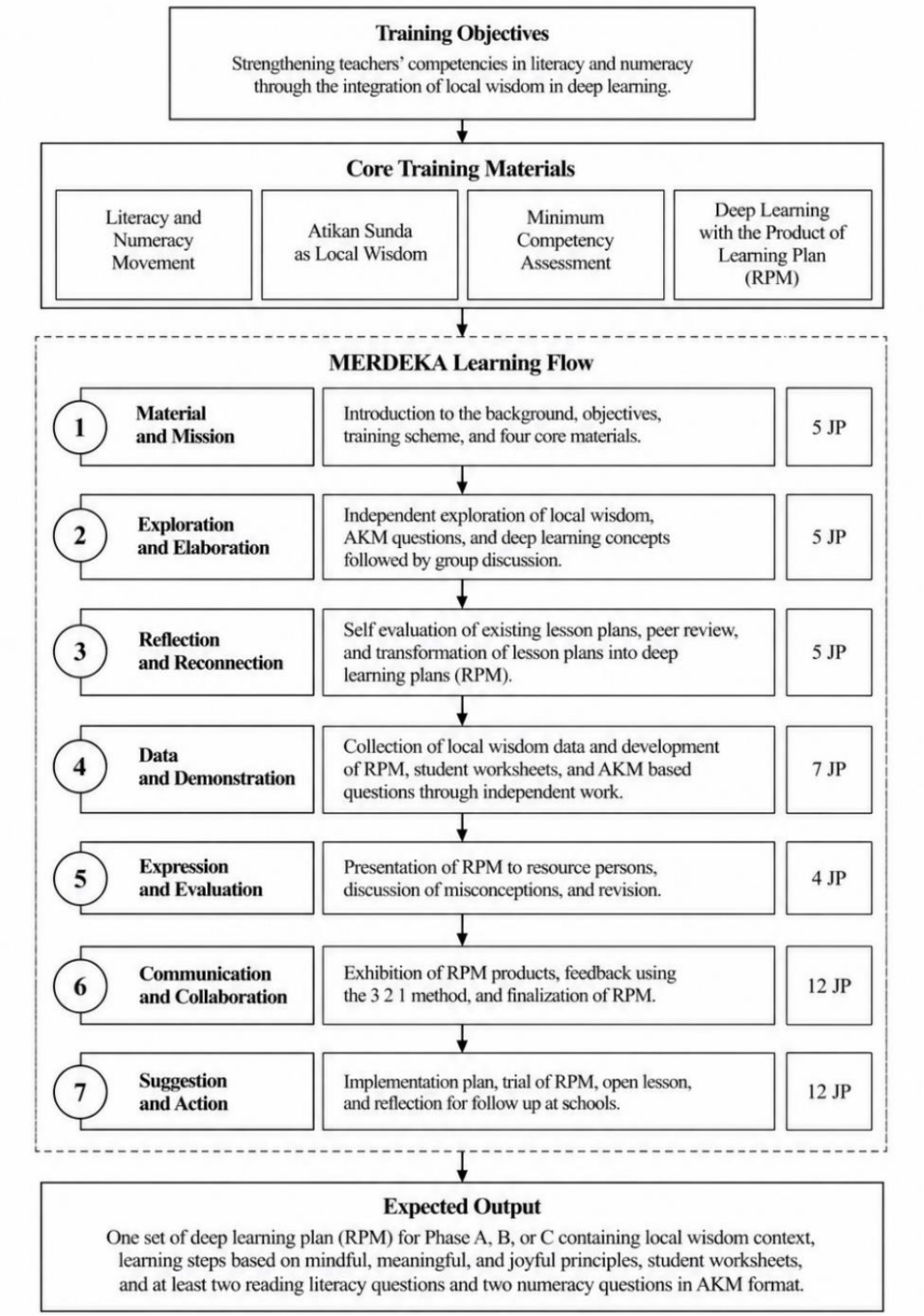


Figure 2. MERDEKA-Based Teacher Training Flow

During the mentoring stage, participants conducted independent work at their schools to identify local wisdom potential and develop RPMs, student worksheets (Lembar Kerja Peserta Didik; LKPD), and AKM-based questions. At the subsequent face-to-face meeting, the RPMs were presented to facilitators for feedback, revised accordingly, and then presented again through a showcase session using 3-2-1 peer feedback. Each participant was expected to produce one learning package containing Sumedang local-wisdom stimuli, the principles of mindful, meaningful, and joyful learning, a student worksheet (LKPD), and at least two reading literacy questions and two numeracy questions in the AKM style.

The learning practice stage was conducted through classroom trials of the RPM at the participants' respective schools and an open lesson following a plan-do-see cycle. Peer teachers and facilitators observed students' interactions, engagement, and reasoning while completing tasks that integrated AKM and local wisdom. Documentation of the programme implementation is presented in Figure 3.



(a) (b)

Figure 3. Documentation of the implementation of the local-wisdom-based Literacy and Numeracy Movement programme. (a) Opening of the programme by the Head of the Sumedang Regency Education Office; and (b) closing of the open lesson activity by the Coordinator of Education Supervisors of the Sumedang Regency Education Office.

The reflection stage was conducted after the learning activity by documenting aspects that had been successfully implemented and those that still required improvement. The reflection was complemented by data from the student response questionnaire and observations made during the *open lesson*. These findings were used to formulate recommendations and follow-up plans for teachers and schools.

Data Collection Instruments and Evidence Sources

Data were collected through a teacher diagnostic questionnaire, pretest-posttest assessments, an RPM assessment sheet, an open lesson observation sheet, a student response questionnaire, open-ended questions, and documentation. Programme evaluation was guided by the Kirkpatrick and Kirkpatrick training evaluation framework, which encompasses four aspects: reaction, learning, behavior, and results. Table 1 aligns the evidence sources with their units of analysis and interpretive limits. Table 2 reports the MERDEKA training stages and scheduled lesson hours, while Table 3 lists the evaluation instruments and analysis techniques.

Table 1. Alignment of evidence sources, units of analysis and interpretive limits

Evidence	Unit / time	Observed output	Interpretive boundary
Teacher diagnostic	28 teachers, 28 Aug 2026	Self-rated needs and participant profile	Self-selection/selection criteria and missingness unexplained
Initial pretest	30 teachers, 29 Aug 2026	Mean 68.3; SD 14.3	Unmatched baseline, not the paired analytical sample
Paired pre/post	28 complete teacher pairs; post 17 Sep	Means 68.0 and 88.4; t(27)=7.65	Within-group association; no control group
Mentoring artifacts	Six documented RPM packages	RPM/LKPD/AKM tasks	Package authorship and rubric audit unverified
Open lesson	One partner school; one Grade VI class	Observed decimal reasoning and participation	Short-term, single-class observation
Pupil experience	25 Grade VI respondents, post-lesson	Reported composite score 77.4%	Perception score, not a learning-gain estimate
Long-term equity/welfare	No evidence source	Not assessed	No gender effect, welfare effect or sustained transfer inference

Table 2. MERDEKA Based Training Activities

MERDEKA Stage	Lesson Hours (JP)	Meeting	Main Activities
Materials and Missions	5	Face-to-face 1	Background, objectives, training pattern, and four main topics
Exploration and Elaboration	5	Face-to-face 1	Exploration of local wisdom, AKM questions, Deep Learning, and discussion
Reflection and Reconnection	5	Face-to-face 1	Self-evaluation of lesson plans, peer review, and transformation of lesson plans into RPM
Data and Demonstration	7	Independent work	Data collection and development of RPM, LKPD, and AKM questions
Expression and Evaluation	4	Face-to-face 2	RPM presentation, misconception analysis, and revision

MERDEKA Stage	Lesson Hours (JP)	Meeting	Main Activities
Communication and Collaboration	12	Face-to-face 2	RPM showcase, 3-2-1 feedback, and RPM finalization
Recommendations and Action	12	Online, independent work, and face-to-face 3	RPM trial, open lesson, and follow-up reflection
Total	50		

Table 3. Program Evaluation Instruments and Data Analysis

Instrument	Respondents	Aspects Measured	Analysis
Diagnostic questionnaire	Teachers	Literasi, numerasi, dan kearifan lokal	Statistik deskriptif, persentase
25-item pretest-posttest	Teachers	Literacy and Numeracy Movement (GLN), Atikan Sunda, AKM, and Deep Learning	Rasch model, mean scores, N-Gain, and paired-samples t-test
RPM assessment sheet	Teachers	RPM, LKPD, and AKM questions	Rubric scores and percentages
Open lesson observation	Students	Interaction, engagement, and reasoning	Qualitative descriptive analysis
Student response questionnaire	Students	Literacy, numeracy, and local-wisdom-based learning	Percentages and categories

Data Analysis and Interpretation

Teacher knowledge was assessed using a reported 25-item instrument with five text stimuli and a 0–100 scoring scale. The aggregate normalized gain can be reconstructed from the reported group means as g_group = (mean_post − mean_pre) / (maximum − mean_pre) = (88.4 − 68.0) / (100 − 68.0) = 0.6375 ≈ 0.64. This value describes gain based on group means rather than the arithmetic mean of individual gains for the 28 paired respondents. Individual gain categories were reported separately; detailed item scores and the handling of near-ceiling pretest values were unavailable for further assessment [41].

$$< g >= \frac{S_{posttest} - S_{pretest}}{S_{maks} - S_{pretest}}$$

For respondents with valid paired scores, the individual normalized-gain formula is g_i = (post_i − pre_i) / (100 − pre_i), where pre_i < 100. The reported categories comprise high (g ≥ 0.70), moderate (0.30 ≤ g < 0.70), low nonnegative (0 ≤ g < 0.30), and negative (g < 0). Winsteps-based Rasch scores were also reported as 1.09 and 3.42 logits. Detailed information on the Rasch model, software version, common-item specification, model fit and processing

procedures was not available, limiting assessment of score comparability. The Rasch values are therefore presented as reported programme findings rather than as independently verified measurement results [42].

The teacher and pupil questionnaires used reported four-point ratings, with reverse scoring specified for negatively worded pupil items and percentage scores apparently calculated as obtained score divided by maximum possible score × 100. The item lists, scoring keys, domain weights, item-level valid denominators and category cutoffs were not available in the supplied source; the label "good" for 77.4% is retained as the programme's reported category, not asserted as externally validated. Open-ended responses and observation notes were summarised descriptively; without transcripts, observer identities, coding records or inter-observer agreement, an independently reproducible thematic analysis cannot be claimed [43]. Programme outcomes were described through attendance, teacher test results, six documented learning packages, open-lesson observations, pupil responses and proposed follow-up activities. These measures capture immediate learning and implementation experiences, while longer-term teacher practice and changes in pupil achievement require separate evaluation.

RESULTS AND DISCUSSION

Results

Results are organised by programme participation and teacher perceptions, initial learning needs, paired knowledge assessments, documented instructional materials, classroom enactment, pupil experiences and proposed follow-up. Each evidence source is reported using its own participant base and unit of analysis.

Teacher Participation and Programme Response

The programme enrolled 30 of the 50 teachers initially planned (60.0% of planned capacity) from 26 public elementary schools. The recorded programme-response and diagnostic data cover 28 teachers; the reason 20 planned places were unused and the reasons that two participants did not supply complete paired outcomes are undocumented. Consequently, the sample is not shown to represent teachers throughout Sumedang, and access barriers related to location, workload, caregiving, resources and connectivity cannot be estimated. The participant numbers, school coverage, and reported programme-response indicators are summarized in Table 4.

Table 4. Teacher Participation and Response to the Program

Indicator	Result
Planned number of teachers	50 teachers
Teachers participating in the program	30 teachers
Schools involved	26 schools
Teachers completing the post-program evaluation	28 teachers
Overall response to the program	93.6%
Quality and relevance of the materials	95.2%
Facilitator competence	96.2%

Indicator	Result
Impact on teacher competence	89.4%
Committee performance and facilities	93.6%

The overall teacher-response score was reported as 93.6%. Dimension scores were 96.2% for facilitator competence, 95.2% for material quality and relevance, 93.6% for committee performance and facilities, and 89.4% for perceived competence improvement. These values describe participants' assessments of the programme rather than independently observed teaching performance. The available summary does not include the item counts and denominators required to reproduce each percentage. The skew towards positive ratings is addressed as a possible response-bias concern in the discussion [44].

Figure 4 presents the four ratings alongside the overall score. The 6.8-point spread between the highest and lowest dimensions suggests generally favourable and relatively consistent responses, although ratings alone do not indicate the extent of subsequent instructional change. Perceived confidence and satisfaction are distinct from performance-based measures [5], [45].

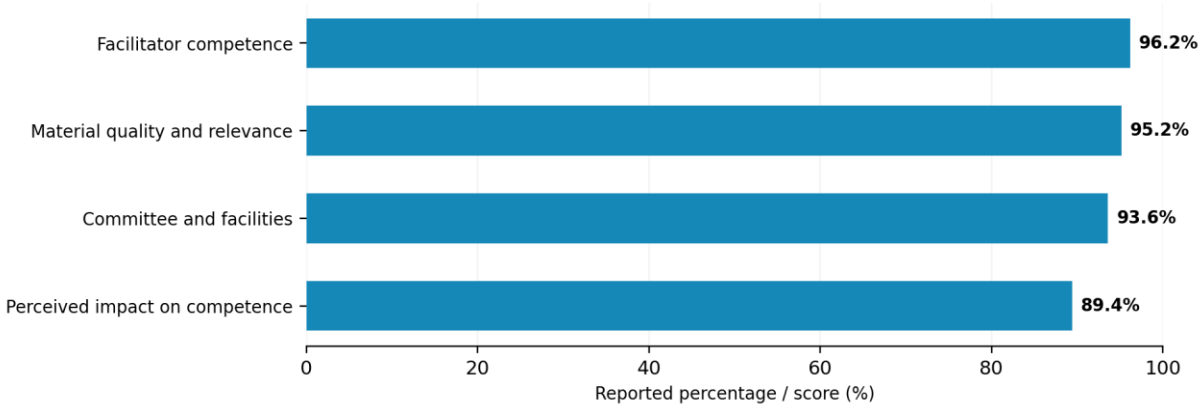


Figure 4. Teacher responses to the program implementation

Additional respondent-reported scores were 90.7% for understanding the Literacy and Numeracy Movement, 87.9% for confidence designing contextual lesson plans, 87.1% for perceived media-use capacity, and 91.9% for stated commitment to implementation. These figures describe self-reported perceptions and intentions, not independent observations of subsequent classroom practice.

The strongly positive response pattern merits some caution. According to the programme evaluation summary, 99.3% of selections were in the "agree" or "strongly agree" categories, with none in the two lowest categories. This distribution may reflect participants' positive experiences as well as response-style or social-desirability effects. Because confidentiality procedures and any separation between trainers and evaluators were not described, self-reported confidence is interpreted alongside, rather than as a substitute for, assessed learning [45], [46].

Open-ended responses provided more specific information regarding the aspects of the programme that participants considered most beneficial. Fifteen of the 28 teachers (53.6%)

identified Deep Learning and the development of RPMs integrated with local wisdom or Pancawaluya as the most beneficial topics. Six teachers (21.4%) highlighted local wisdom and Sundanese culture, while five teachers (17.9%) considered the integration of literacy and numeracy with local wisdom to be the most beneficial component. These findings indicate that teachers need practical learning experiences to translate literacy and numeracy concepts into learning materials that are appropriate to students' environmental and cultural contexts. This need for practical experience is consistent with the characteristics of effective professional development, which emphasize active learning, connections to practice, and opportunities to apply knowledge as important components of teacher competence development.

Eleven of 28 teachers (39.3%) identified the concrete integration of local wisdom or Pancawaluya into RPMs as their principal difficulty. Ten (35.7%) reported no significant difficulty, whereas seven (25.0%) cited time constraints. These responses identify different experiences within the participating group; they do not measure observed changes in instructional practice.

Initial Teacher Needs and Competencies

Data collection began with a diagnostic questionnaire completed by 28 teachers on 28 August 2026, one day before formal delivery began. The baseline knowledge test was completed by 30 teachers on 29 August. Because these instruments do not use the same respondent base or construct, their percentages must not be compared as interchangeable measures of competence. The diagnostic questionnaire provides 28-person characteristics; the all-attendee baseline score describes 30 participants. The paired analysis is limited to 28 persons with both assessments. The characteristics and baseline conditions of respondents are detailed in Table 5.

Table 5. Profile and Initial Conditions of Training Participants

Characteristic	Category	n	%
Gender	Female	19	67.90
	Male	9	32.10
Age	<30 years	2	7.10
	30–40 years	16	57.10
	41–50 years	9	32.10
	>50 years	1	3.60
Highest Educational Qualification	Bachelor's degree (S1)	26	92.90
	Master's degree (S2)	2	7.10
Teaching Experience	<5 years	2	7.10
	5–20 years	22	78.60
	>20 years	4	14.30
Teacher Certification	Certified	27	96.40
	Not certified	1	3.60
Literacy-Numeracy Training in the Previous 3 Years	None	21	75.00
	Once	5	17.90
	≥2 times	2	7.10

Characteristic	Category	n	%
Grade Levels Taught	Grades I–III	3	10.70
	Grades IV–VI	25	89.30

Among the 28 diagnostic respondents, 19 were women (67.9%) and nine were men (32.1%); 16 were aged 30–40 years (57.1%); 26 held bachelor's degrees (92.9%); 27 had teacher certification (96.4%); 22 reported 5–20 years of experience (78.6%); and 21 had received no literacy–numeracy training during the preceding three years (75.0%). These figures describe the respondent group. The evaluation did not include subgroup analyses of participation barriers or training effects, so potential differences by gender or other characteristics remain unexplored.

The questionnaire results showed that teachers' perceived competence was 71.3% for literacy, 73.2% for numeracy, and 74.6% for local wisdom. However, competence related to the production of learning materials was lower. Only 21.4% of teachers felt capable of developing AKM-equivalent literacy questions, 42.9% felt capable of developing contextual numeracy questions, and 35.7% felt capable of developing local-wisdom-based learning materials. In contrast, training needs reached 90.0% for literacy, 88.6% for numeracy, and 86.4% for local-wisdom-based learning materials. The comparison between perceived competence and training needs is presented in Figure 5.

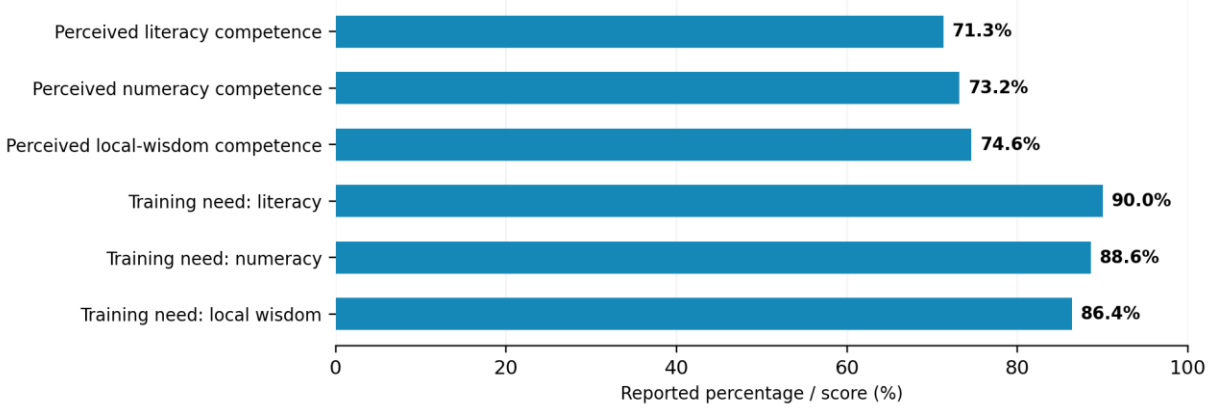


Figure 5. Teachers' Perceived Competence and Training Needs

Figure 5 compares self-rated domain competence with self-reported demand for additional training, two different constructs presented side-by-side for planning purposes only. A higher training demand does not represent a measured competence deficit of the same magnitude. Item wording, response scale and scores of the respondent-level comparison remain unavailable, limiting the interpretation of numeric gaps.

These needs were also reflected in participants' selection of priority training topics. Local-wisdom-based learning materials were the most frequently selected need (82.1%), followed by basic literacy and numeracy (78.6%) and active learning strategies (71.4%). The need for AKM question development and the use of instructional media or technology each reached 60.7%. This pattern indicates that teachers needed not only conceptual knowledge but also support in translating literacy and numeracy concepts into learning materials and instructional strategies that could be implemented in the classroom.

Teachers' actual competence was further reflected in the pretest results. The mean pretest score of the 30 teachers was 68.3 (SD = 14.3), with scores ranging from 44 to 92. Nine teachers (30.0%) scored below 60, 12 teachers (40.0%) scored between 60 and 79, and nine teachers (30.0%) scored 80 or higher. Across the stimuli, the highest percentage of correct responses was observed for Pancawaluya in science (86.0%), followed by AKM (83.3%) and story-based scientific inquiry (80.0%). Performance on the Deep Learning-based inquiry lesson plan reached 66.7%, whereas numeracy data analysis was substantially lower at 46.7%. A comparison of teachers' performance across the pretest stimuli is presented in Figure 6.

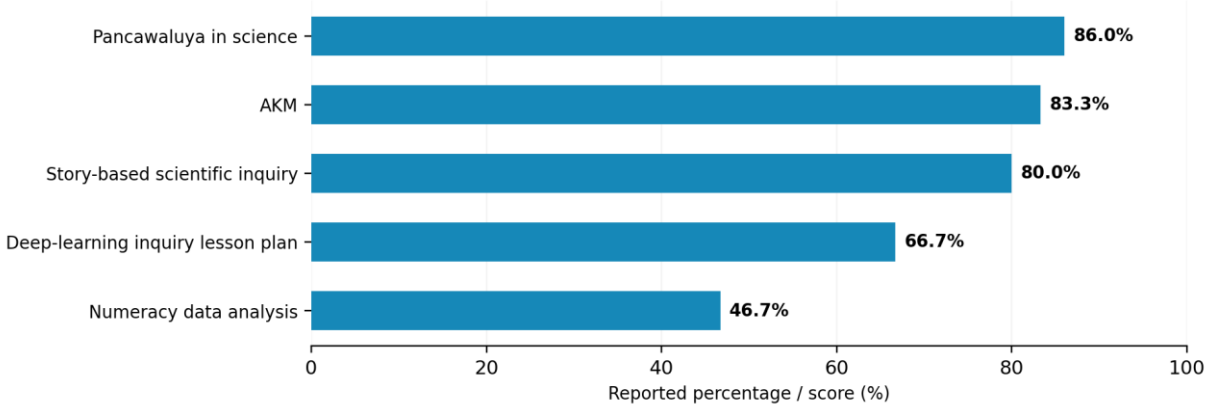


Figure 6. Teachers' Initial Competence Based on Pretest Stimuli

The all-attendee pretest showed heterogeneous performance and particularly weak numeracy data-analysis responses. The reported self-perceived numeracy score of 73.2% and numeracy-stimulus correct-response rate of 46.7% use different instruments and cannot be directly subtracted as if they were on one psychometric scale. The source reports that their correlation was weak and non-significant but does not provide the correlation estimate, p-value, confidence interval or paired data, so no reproducible association is asserted.

The baseline findings informed the programme's emphasis on numeracy reasoning, contextualised reading and problem solving, and the preparation of AKM-style tasks and locally grounded instructional materials. The perceived-competence ratings and the stimulus-specific test scores measure different constructs and were not pooled.

Improvement in Teacher Understanding

The paired analysis used 28 teachers with both measurements. The pretest mean was 68.0 (SD = 14.7) and the posttest mean 88.4 (SD = 11.2), a recorded increase of 20.4 points. These descriptive values differ appropriately from the full 30-participant pretest mean of 68.3 (SD = 14.3). The two cases without complete paired scores have not been characterised, and a missing-data sensitivity analysis cannot be performed without their records. Table 6 distinguishes the full-attendee pretest from the matched sample, and Table 7 summarizes the recorded teacher-understanding outcomes.

Table 6. Baseline and matched-analysis denominator reconciliation

Assessment / metric	N	Recorded summary	Use in inference
Pretest, all attendees	30	Mean 68.3; SD 14.3; 9/30 at score ≥80	Describes all programme attendees at baseline
Pretest, matched subset	28	Mean 68.0; SD 14.7	Paired change baseline
Posttest, matched subset	28	Mean 88.4; SD 11.2; 25/28 at score ≥80	Paired change posttest
Reported paired difference	28	Mean +20.4; t(27)=7.65; p<.001	95% CI ≈ [14.93, 25.87] from rounded summaries
Reported individual categories	28	High 17; moderate 5; low nonnegative 4; negative 2	Verbatim category counts; raw gains unavailable
Aggregate normalized gain	28	(88.4−68.0)/(100−68.0)=0.6375	Gain of means, not mean person-level gain
Rasch summary (unverified equating)	28	1.09→3.42 logits; difference 2.33	Do not treat as independent replication

Table 7. Evaluation Results of the Improvement in Teacher Understanding

Evaluation Indicator	Result
Number of participants	28 teachers
Mean score	Pretest = 68.0 Posttest = 88.4
Score increase	+20.4 points
N-Gain	0.64 (moderate)
Paired-samples t-test	t(27) = 7.65 p < 0.001
Effect size	d = 1.45 (large)
Rasch score	Pretest = 1.09 logit Posttest = 3.42 logit
Rasch score increase	+2.33 logit t(27) = 7.69 p < 0.001

The recorded proportion scoring at least 80 was 9/30 (30.0%) among all pretest participants and 25/28 (89.3%) in the posttest dataset. Since those percentages use different denominators, they are retained as separate cross-sectional descriptions and must not be presented as a within-person change in mastery without the matched baseline mastery count.

The reported paired t-test was $t(27) = 7.65$, $p < .001$, with a Shapiro–Wilk difference-score test $p = .66$. Using the published rounded mean difference and t value gives an approximate standard error of 2.67 and an approximate 95% confidence interval for the mean paired difference of 14.93–25.87 points. The paired standardized effect is $d_z = t/\sqrt{28} \approx 1.45$; inversion of the noncentral-t distribution gives an approximate 95% confidence interval of 0.91–1.97 for d_z. All intervals are algebraically reconstructed from rounded source summaries, not from the raw scores. A non-significant normality test does not prove normality, and the unavailable difference-score distribution prevents outlier checks [40], [44], [47].

The reported individual gain categories comprised 17 high (60.7%), five moderate (17.9%), four nonnegative low (14.3%), and two negative (7.1%) gains, summing to the 28 complete paired cases. Thus, 22/28 (78.6%) were in the high/moderate bins. The 0.64 value is the aggregate gain from group means; it is not necessarily the average of the 28 individual gains. Source-reported numeracy stimulus performance rose from 46.4% to 87.5% and the Deep Learning-based inquiry lesson-plan stimulus from 65.7% to 89.3%. Without the 28 row-level records, individual classifications and uncertainties cannot be independently recomputed.

A separate Rasch output in the original programme report stated mean person estimates of 1.09 logits pretest and 3.42 logits posttest (reported difference +2.33 logits; $t(27) = 7.69$, $p < .001$). These values are preserved. However, both Rasch summaries and raw percentages derive from the same item-response administration, and the report provides no calibration/equating, extreme-score or model-fit audit; therefore the logits do not establish an independent replication of learning gains. No claims of validated invariance, reliability or unidimensionality are made without the missing Rasch outputs [42].

The matched assessments document a within-teacher increase in recorded test scores. The available summaries do not permit independent verification of individual gain scores, differential item functioning or the Rasch model specification.

Mentoring and Learning Products

Mentoring sessions focused on drafting, reviewing and refining the instructional materials prepared by participating teachers before classroom application.

The mentoring records identify six documented RPM packages, accompanied by LKPD and AKM-style literacy and numeracy questions. Although the programme initially envisaged one package for each participating teacher, the available documentation does not clarify whether the six retained packages were jointly produced, selected for presentation or evaluated as final submissions. Accordingly, six documented outputs are described without assuming that every participant produced an individually assessed package. Their intended design included local contexts, reflective learning activities, and at least two reading and two numeracy questions.



Figure 7. Mentoring process and learning products integrating literacy, numeracy, and local wisdom. (a) RPM discussions in groups and plenary sessions; (b) presentation of the model teacher's lesson plan; (c) observation of learning practice; and (d) facilitator-led learning reflection.

The photographic evidence in Figure 7 documents the mentoring and presentation context, not validated artifact quality. Original and revised RPM/LKPD versions, dated assessor scores, rubric descriptors, the number of assessors and any inter-rater agreement are absent; consequently, the direction and magnitude of instructional-design improvement cannot be independently audited. The available evidence supports the narrower statement that six packages were documented after mentoring [14].

Implementation of Local Wisdom-Based Learning

One documented open lesson at a public elementary school in Sumedang brought together a model teacher, observers/facilitators and Grade VI pupils. Its decimal-learning tasks referenced Ubi Cilembu, Tahu Sumedang, Opak and Rengginang as culturally familiar local food examples, using the Pos Jawara 1–5 stations, an answer passport, challenge envelopes, illustrated LKPD and a digital resource named Jawara Desimal. These elements are described in the programme record. The original file does not identify the digital platform, device availability, connectivity, accessibility accommodations or whether all 26 schools were systematically observed; no digital usability or school-wide fidelity claim is made.

Observers described pupils' engagement with local foods and collaborative tasks, alongside difficulties translating narrative problems into equations. Sundanese vocabulary use was observed informally; no separately validated Sundanese reading or language assessment was described. Pupils who did not share the presumed cultural background or language might have faced additional barriers, but such barriers were not systematically assessed [12], [17].

In terms of numeracy, local food contexts provided opportunities for students to apply decimal operations in more concrete situations. Nevertheless, some students experienced difficulties when solving word problems that required more than one step. Difficulties were also observed in aligning decimal place values, understanding the equivalence between 0.8 and 0.80, and transforming verbal information into mathematical sentences. These findings indicate that local wisdom contexts need to be accompanied by adequate scaffolding so that students do not merely recognize the context but are also able to use it to construct conceptual understanding.



(a) (b)

(c)

Figure 8. Implementation and reflection of local wisdom-based learning through an open lesson. (a) Interviews with students after the learning session; (b) facilitator-led learning reflection; and (c) participants presenting the results of learning observations.

Reflection on the open lesson highlighted several directions for improvement: shorter and more clearly sequenced instructions in the module and LKPD, stronger use of concrete objects and visual representations to explain decimals, and a more balanced integration of reading and numeracy. The discussion also emphasised connecting local examples with specific learning goals rather than relying on cultural names or objects alone. Clearer group roles and opportunities for individual reflection were recommended to support broader pupil participation.

Table 8. Instructional design observations and recommendations from the open lesson

Component	Documented during the lesson	Proposed improvement (not yet verified)
Mathematics	Difficulty with 0.8 and 0.80, place value and multi-step word problems	Use place-value tables, visuals and gradual representations

Component	Documented during the lesson	Proposed improvement (not yet verified)
Literacy/AKM tasks	Contextual word problems and LKPD present; some text-to-equation difficulty	Shorten instructions and stage information extraction
Local wisdom	Foods used as task contexts; Pancawaluya not fully explored	Add explicit cultural-value prompts aligned to objectives
Participation	Some pupils followed more dominant peers	Rotate reader, calculator, explainer and reporter roles
Accessibility	Digital Jawara Desimal mentioned; access conditions unreported	Document device/offline, language and accessibility options
Reflection	Teacher feedback after one open lesson	Repeat lesson and track all pupils' task participation

Student Engagement and Learning Experiences

Following the Grade VI open lesson, 25 pupils completed a response questionnaire. The programme summary reported an overall questionnaire score of 77.4%, classified as "good". This is an aggregate response measure, not the proportion of pupils whose performance improved. The scoring rubric and respondent-level data needed to verify its weighting and cut-off were not available, and no pre-lesson or post-lesson literacy or numeracy performance test was administered. Table 9 presents the response indicators separately from pupil counts and highlights the interpretation appropriate to each measure.

Table 9. Student-response metric and arithmetic audit (n = 25 questionnaires)

Reported indicator	Source metric	Audit interpretation
Overall response	77.4%	Composite reported score; item weights/cutoff unavailable
Learning process	81.7%	Subscale percent score, not fraction of pupils
Numeracy attitudes	79.5%	Subscale percent score, not fraction of pupils
Local wisdom	78.8%	Subscale percent score, not fraction of pupils
Reading habits	69.1%	Subscale percent score, not fraction of pupils
Confused by pace	25%	Not an integer share of 25; verify item scoring or denominator
Afraid of arithmetic	22%	Not an integer share of 25; verify item scoring or denominator
Bored by lengthy text	54%	Not an integer share of 25; verify item scoring or denominator

Difficulty retelling	39%	Not an integer share of 25; verify item scoring or denominator

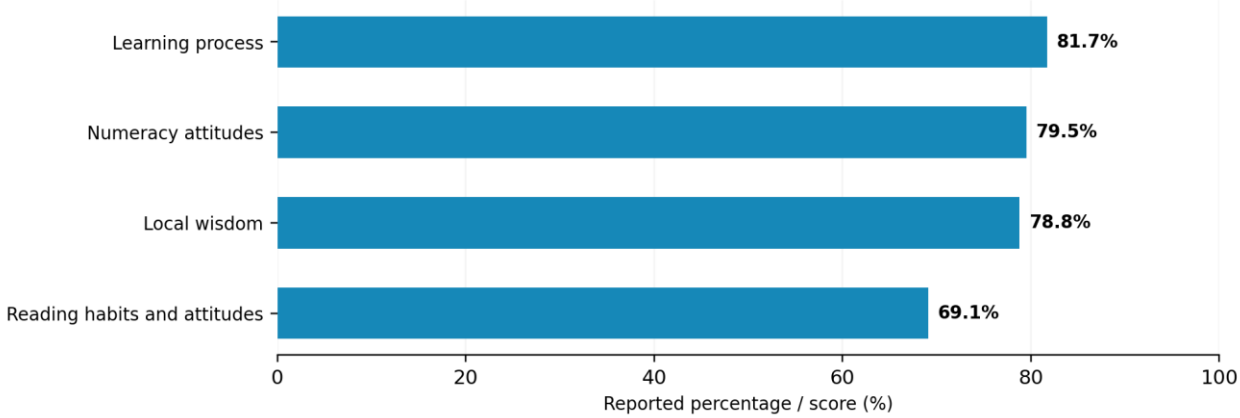


Figure 9. Student responses to local wisdom-based literacy and numeracy learning.

Reported domain scores were: learning process 81.7%, confidence and attitudes toward numeracy 79.5%, local wisdom-based learning 78.8%, and reading habits/attitudes 69.1%. These are questionnaire-domain score percentages, not fractions of pupils; the document does not provide the item maximums and weights needed for independent recomputation. The four-domain unweighted mean would not exactly reconstruct 77.4%, further underscoring that the reported composite needs its original scoring sheet.

The source states 96% enjoyment, 100% clarity of explanations, 96% confidence asking questions and 96% desire to repeat the lesson, all arithmetically possible with n = 25. It also reports "25% confused by fast pacing," which cannot be a simple number-of-respondents proportion with 25 complete responses; that value is retained as a source-reported indicator requiring item-level denominator or scoring clarification, not a verified count of pupils. Approval-seeking response bias cannot be excluded [46].

The source reports 100% perceived cultural familiarity, 96% positive attitudes to learning about the local area, 92% perceived support from cultural contexts and 52% difficulties making explicit conceptual connections. These are pupils' self-reports after one lesson and must not be interpreted as validated cultural knowledge growth. Explicit links between food stimuli, decimal properties and learning goals would need to be observed systematically before asserting culturally responsive practice [11], [12].

For numeracy, the source reports 100% stated confidence and 92% perseverance, alongside a 22% indicator of arithmetic fear. For literacy, the corresponding source figures are 88% enjoyment of books, 48% near-daily reading, 54% boredom with lengthy text and 39% difficulty retelling. Because 22%, 54% and 39% are not integer multiples of 4%, they cannot represent simple unweighted counts from exactly 25 complete pupils; the percentages are preserved verbatim as programme-reported item metrics requiring clarification of denominators, weighting or score-based origin. Importantly, reported self-confidence coexisted with observed difficulties in decimal equivalence and multi-step word problems, demonstrating why perceptions cannot substitute for knowledge assessment.

The post-lesson questionnaire describes pupils' immediate perceptions of one lesson. Because there was no comparable assessment of learning before or after the session, these reports cannot establish gains in literacy or numeracy achievement.

Challenges and Responses in the Open Lesson

The implementation of the open lesson revealed several challenges in applying local wisdom-based learning, particularly in conceptual understanding, translating contextual information into problem-solving processes, and ensuring equitable student engagement. The observation results showed that some students experienced difficulties in aligning decimal place values, understanding the equivalence between 0.8 and 0.80, pronouncing decimal numbers, and translating information presented in word problems into mathematical sentences. These conditions indicate that contexts closely connected to students' everyday lives do not automatically result in mathematical understanding; students still require support to connect contextual experiences with mathematical representations and procedures. This finding is consistent with Pratiwi and Ariawan [49] , who reported an association between mathematical understanding and reading comprehension when mathematical problems are presented through everyday-life and local wisdom contexts.

Open-lesson observers described uneven participation in some groups, where one or two pupils frequently proposed answers that others followed. This observation points to a classroom-organisation issue, although the frequency of individual contributions was not systematically counted. Assigning clearer roles for reading, calculation, explanation and reporting was recommended as a way to encourage wider participation; its effect would need to be examined in subsequent lessons [17].

Local foods including Ubi Cilembu, Tahu Sumedang, Opak and Rengginang were used as mathematical contexts. The observational record itself notes that Pancawaluya values were not explored in depth. Thus, culturally familiar examples were present, but instruction in the named local ethical values cannot be claimed as achieved. The findings show a distinction between contextual stimulus and sustained cultural-value pedagogy [10], [12].

Responses to these challenges included gradual assistance provided by the teacher, the use of visual representations of decimal numbers, and challenge-based activities that encouraged students to develop solution strategies collaboratively. Several aha moments emerged when students received appropriate assistance and were able to connect visual representations with the problems they were solving. These strategies demonstrate the importance of providing scaffolding when students move from everyday-life contexts toward more abstract mathematical representations. This approach is also consistent with Munafiroh et al. [50], who demonstrated that integrating local wisdom into mathematics learning can support elementary students' problem-solving abilities.

The subsequent reflection generated several recommendations for improvement, including simplifying the module and LKPD by providing more gradual instructions, strengthening decimal concepts through concrete objects, visual representations, and place-value tables, balancing literacy and numeracy activities, and using one or two local objects that are most familiar to students. In addition, group roles should be defined more clearly, and individual reflection activities should be strengthened so that all students have opportunities to articulate their understanding. These recommendations indicate that the *open lesson* functioned

not only as a learning demonstration but also as a reflective mechanism for identifying learning barriers and continuously improving the learning design.

Programme Reflection and Follow-Up

Proposed follow-up activities included teacher working groups (Kelompok Kerja Guru; KKG), additional open lessons, reading corners, item banks and dissemination of lesson materials. Their implementation would depend on collaboration among schools and teacher networks, sufficient time, appropriate resources and accessible participation arrangements. The available evaluation covers planning for these activities rather than their subsequent delivery or costs [9], [20].

Among the 28 teachers whose recommendations were summarised, 12 (42.9%) prioritised programme continuity; five (17.9%) recommended schedule adjustments; four (14.3%) requested additional teaching practice or lesson study; and three (10.7%) sought further learning materials. General appreciation and improvements in facilities or logistical information were each mentioned by two teachers (7.1%). Figure 10 displays these reported categories.

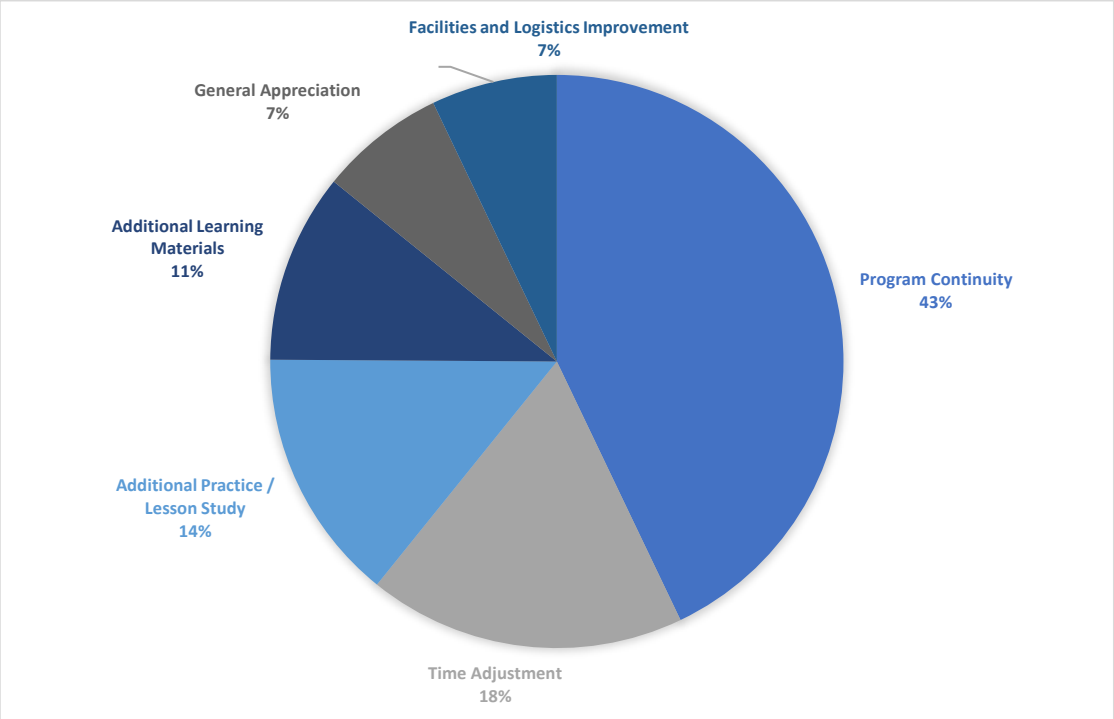


Figure 10. Distribution of teacher recommendations for program sustainability.

Discussion

The evaluation reveals a distinction between an exceptionally positive training experience and independently assessed teacher learning. Participants rated the facilitators and materials highly, but these reaction measures cannot establish instructional competence. The paired knowledge test provides a more direct indicator of short-term learning, while the six retained RPM packages and one open lesson describe subsequent steps in applying that learning. Treating the

measures separately is important: effective professional learning can involve changes in knowledge, resources and teaching, but progress in one domain does not guarantee progress in the others [5], [21], [29].

The 20.4-point increase in the matched teachers' mean test score is substantial within the observed group, and the reported paired test is consistent with a detectable difference. However, the single-group design does not distinguish programme-related learning from retesting, participant motivation or time-related influences. In addition, the absence of item-level records, model-fit information and a verifiable Rasch calibration prevents claims that the two reported score metrics independently validate one another. Future evaluations should retain anonymised paired responses and instrument specifications and, where feasible, include a comparison condition and delayed follow-up [8], [40], [42], [47].

The use of Sumedang food contexts gave teachers a concrete starting point for locally grounded literacy and numeracy tasks. The observed difficulties with decimal equivalence, place value and multi-step word problems nevertheless suggest that contextual familiarity alone was insufficient to secure mathematical understanding. Materials would benefit from carefully sequenced representations, shorter reading prompts and explicit links between cultural references and the target concept. This interpretation is consistent with work on culturally sustaining pedagogy, educative curriculum materials and instructional scaffolding [12], [14], [16], [30].

Inclusion emerged as a practical issue rather than a measured programme outcome. Observers noted that some pupils followed more vocal group members, and the study did not systematically record individual participation, language needs, accessibility arrangements or differential experiences. Although women represented 19 of the 28 diagnostic respondents, the data do not support conclusions about gender-specific training benefits or barriers. Rotating group roles, recording individual contributions and monitoring access to digital and non-digital materials would make future evaluations better able to address participation and equity [17], [18], [19], [39].

The teachers' requests for continued mentoring, revised scheduling and additional materials point to possible conditions for sustained implementation. A practical next phase would track the revision and classroom use of RPMs across participating schools, document the reach of KKG activities, and assess pupil reasoning before and after lessons rather than relying solely on satisfaction reports. Such monitoring would test whether the promising short-term findings translate into more dependable instructional practice, while keeping claims about pupil achievement and social welfare proportionate to the evidence [9], [20], [28], [48].

CONCLUSION

The Sumedang programme was associated with a substantial improvement in teacher knowledge-assessment scores among 28 participants with complete pretest and posttest records (68.0 to 88.4; mean difference 20.4 points; $t(27) = 7.65$, $p < .001$; $d_z \approx 1.45$; aggregate $g = 0.64$). Thirty teachers from 26 schools participated, against an initial plan of 50 places. The programme also produced six documented RPM learning packages, included an observed Grade VI open lesson and received generally positive responses from 25 pupils (reported overall score 77.4%). These findings capture related but distinct aspects of professional

learning: assessed teacher knowledge, instructional outputs and pupils' immediate perceptions. Changes in pupil literacy and numeracy achievement were not directly measured. Classroom observations nevertheless identified areas requiring further attention, including decimal concepts, multi-step word problems, uneven group participation and the translation of Pancawaluya values into explicit learning activities. Continued support through teacher working groups, lesson study and mentoring could help address these needs, although the proposed follow-up activities have yet to be evaluated. Future studies would benefit from complete assessment records, repeated classroom observations and measures of pupil learning and participation over time, enabling a fuller understanding of how locally grounded professional development contributes to classroom practice.

LIMITATIONS

The evaluation used a single-group design without random assignment, a comparison group, a delayed teacher test or a class-level pretest, which limits causal interpretation. Two of the 30 teacher attendees did not have complete paired assessments, and no comparison of completers and non-completers was available. Although 26 schools were represented, direct classroom evidence came from one open lesson and questionnaire responses from 25 pupils in a single Grade VI class. Detailed teacher test items, pupil questionnaire forms, response coding, individual scores, Rasch model outputs, lesson-hour definitions, implementation-fidelity records and artifact assessment sheets were not available for independent examination. As a result, the precision of some reported statistics and the consistency of scoring procedures could not be fully assessed. Teacher and pupil ratings may also reflect expectancy, satisfaction or ceiling effects. The programme did not measure subgroup differences, barriers to participation, language or support needs, longer-term instructional practice or wider pupil outcomes. Documentation of permissions for identifiable images likewise requires confirmation. These limitations confine the findings to short-term changes in assessed teacher knowledge, available instructional products and classroom experiences observed during the programme.

AUTHOR INFORMATION

Corresponding Author

Ika Mustika Sari – Department of Physics Education, Universitas Pendidikan Indonesia (Indonesia);
orcid.org/0000-0003-3272-1841
Email: ikams@upi.edu

Authors

Ika Mustika Sari – Department of Physics Education, Universitas Pendidikan Indonesia (Indonesia);
orcid.org/0000-0003-3272-1841

Winny Liliawati – Department of Physics Education, Universitas Pendidikan Indonesia (Indonesia);
orcid.org/0000-0002-9938-618X

Ridwan Efendi – Department of Physics Education, Universitas Pendidikan Indonesia (Indonesia);
orcid.org/0000-0002-3251-0908
Nuryana – Department of Physics Education, Universitas Pendidikan Indonesia (Indonesia);
orcid.org/0009-0002-6395-4373
Sumarna Surapranata – Department of Physics Education, Universitas Pendidikan Indonesia (Indonesia);
orcid.org/0009-0008-1209-1292
Suhendiana Noor – Department of Physics Education, Universitas Pendidikan Indonesia (Indonesia);
orcid.org/0009-0007-9991-3537
Intan Setiawati – Balai Besar Guru dan Tenaga Kependidikan Provinsi Jawa Barat (Indonesia);
orcid.org/0009-0004-1895-3990
M Gina Nugraha – Department of Physics Education, Universitas Pendidikan Indonesia (Indonesia);
orcid.org/0000-0001-6660-0306
Selly Feranie – Department of Physics, Universitas Pendidikan Indonesia (Indonesia);
orcid.org/0000-0002-1425-2685
Muhamad Nur Dandy – Department of Physics, Universitas Pendidikan Indonesia (Indonesia);
orcid.org/0009-0001-8300-6103

AUTHOR CONTRIBUTION

I.M.S.: Conceptualization, Methodology, Supervision, Project administration, Writing – review & editing; W.L. Project administration, Investigation, Data curation, Formal analysis, Writing – original draft, Writing – review & editing; R.E.: Methodology, Formal analysis, Investigation, Writing – original draft; N.: Conceptualization, Methodology, Supervision, Writing – review & editing; Sumarna Surapranata: Conceptualization, Methodology, Supervision, Writing – review & editing; S.N.: Methodology, Formal analysis, Investigation, Writing – original draft; Intan Setiawati: Methodology, Validation, Supervision, Writing – review & editing; M.G.N.: Methodology, Validation, Supervision, Writing – review & editing; S.F. and M.N.D.: Methodology, Formal analysis, Investigation, Writing – original draft. All nine authors reviewed and approved the final manuscript and accept responsibility for their respective contributions.

CONFLICT OF INTEREST

"The authors declare no conflict of interest."

ACKNOWLEDGMENTS

This study was supported by the Pusat Layanan Pembiayaan Pendidikan, Ministry of Primary and Secondary Education of the Republic of Indonesia (Grant No. 26146/1/A2/KU.07.00/2026). The authors also thank the participating Dinas Pendidikan

Kabupaten Sumedang for their valuable contributions to this programme. The authors acknowledge Siti Almaidah for manuscript-editing assistance and constructive editorial feedback.

DECLARATION OF USE OF AI IN SCIENTIFIC WRITING

The authors used ChatGPT to support language refinement and manuscript organisation. The authors reviewed the suggested text and retain full responsibility for the accuracy, integrity and final content of this article.

REFERENCES

[1] OECD, PISA 2022 Results (Volume I): The State of Learning and Equity in Education. 2023. https://doi.org/10.1787/53f23881-en

[2] I. Chang, "Early numeracy and literacy skills and their influences on fourth-grade mathematics achievement: A moderated mediation model", *Large-scale Assessments in Education,* vol. 11, Article 18, 2023. https://doi.org/10.1186/s40536-023-00168-6

[3] W. Wibawa, Y. M. Saputra, and I. Safari, "Evaluasi Programme Transformasi Sekolah Simpati (PTSS) Pada Jenjang Sekolah Dasar di Kabupten Sumedang", Journal of SPORT (Sport, Physical Education, Organization, Recreation, and Training), vol. 7, no. 3, 2023. https://doi.org/10.37058/sport.v7i3.8449

[4] H. Borko, "Professional development and teacher learning: Mapping the terrain", *Educational Researcher,* vol. 33, no. 8, pp. 3–15, 2004. https://doi.org/10.3102/0013189X033008003

[5] L. M. Desimone, "Improving impact studies of teachers' professional development: Toward better conceptualizations and measures", *Educational Researcher,* vol. 38, no. 3, pp. 181–199, 2009. https://doi.org/10.3102/0013189X08331140

[6] S. Sims et al., "Effective teacher professional development: New theory and a meta-analytic test", *Review of Educational Research,* vol. 95, no. 2, pp. 213–254, 2025. https://doi.org/10.3102/00346543231217480

[7] A. J. Visscher, N. Dmoshinskaia, M. Pellegrini, and A. Rey-Naizaque, "(When) do teacher professional development interventions improve student achievement? A meta-analysis of 128 high-quality studies", *Educational Research Review,* vol. 49, Article 100742, 2025. https://doi.org/10.1016/j.edurev.2025.100742

[8] M. A. Kraft, D. Blazar, and D. Hogan, "The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence", *Review of Educational Research,* vol. 88, no. 4, pp. 547–588, 2018. https://doi.org/10.3102/0034654318759268

[9] V. Vescio, D. Ross, and A. Adams, "A review of research on the impact of professional learning communities on teaching practice and student learning", *Teaching and Teacher Education,* vol. 24, no. 1, pp. 80–91, 2008. https://doi.org/10.1016/j.tate.2007.01.004

[10] G. Ladson-Billings, "Toward a theory of culturally relevant pedagogy", *American Educational Research Journal,* vol. 32, no. 3, pp. 465–491, 1995. https://doi.org/10.3102/00028312032003465
[11] G. Gay, "Preparing for culturally responsive teaching", *Journal of Teacher Education,* vol. 53, no. 2, pp. 106–116, 2002. https://doi.org/10.1177/0022487102053002003
[12] D. Paris, "Culturally sustaining pedagogy: A needed change in stance, terminology, and practice", *Educational Researcher,* vol. 41, no. 3, pp. 93–97, 2012. https://doi.org/10.3102/0013189X12441244
[13] D. L. Ball, M. H. Thames, and G. Phelps, "Content knowledge for teaching: What makes it special?", *Journal of Teacher Education,* vol. 59, no. 5, pp. 389–407, 2008. https://doi.org/10.1177/0022487108324554
[14] E. A. Davis and J. S. Krajcik, "Designing educative curriculum materials to promote teacher learning", *Educational Researcher,* vol. 34, no. 3, pp. 3–14, 2005. https://doi.org/10.3102/0013189X034003003
[15] P. Black and D. Wiliam, "Assessment and classroom learning", *Assessment in Education: Principles, Policy & Practice,* vol. 5, no. 1, pp. 7–74, 1998. https://doi.org/10.1080/0969595980050102
[16] J. van de Pol, M. Volman, and J. Beishuizen, "Scaffolding in teacher–student interaction: A decade of research", *Educational Psychology Review,* vol. 22, pp. 271–296, 2010. https://doi.org/10.1007/s10648-010-9127-6
[17] L. Florian and K. Black-Hawkins, "Exploring inclusive pedagogy", *British Educational Research Journal,* vol. 37, no. 5, pp. 813–828, 2011. https://doi.org/10.1080/01411926.2010.501096
[18] L. Florian and J. Spratt, "Enacting inclusion: A framework for interrogating inclusive practice", *European Journal of Special Needs Education,* vol. 28, no. 2, pp. 119–135, 2013. https://doi.org/10.1080/08856257.2013.778111
[19] M. Ainscow, "Inclusion and equity in education: Making sense of global challenges", *Prospects,* vol. 49, pp. 123–134, 2020. https://doi.org/10.1007/s11125-020-09506-w
[20] J. A. Durlak and E. P. DuPre, "Implementation matters: A review of research on the influence of implementation on programme outcomes and the factors affecting implementation", American Journal of Community Psychology, vol. 41, pp. 327–350, 2008. https://doi.org/10.1007/s10464-008-9165-0
[21] M. M. Kennedy, "How does professional development improve teaching?", *Review of Educational Research,* vol. 86, no. 4, pp. 945–980, 2016. https://doi.org/10.3102/0034654315626800
[22] V. D. Opfer and D. Pedder, "Conceptualizing teacher professional learning", *Review of Educational Research,* vol. 81, no. 3, pp. 376–407, 2011. https://doi.org/10.3102/0034654311413609
[23] B. Avalos, "Teacher professional development in Teaching and Teacher Education over ten years", *Teaching and Teacher Education,* vol. 27, no. 1, pp. 10–20, 2011. https://doi.org/10.1016/j.tate.2010.08.007
[24] T. R. Guskey, "Professional development and teacher change", *Teachers and Teaching,* vol. 8, no. 3, pp. 381–391, 2002. https://doi.org/10.1080/135406002100000512

[25] M. S. Garet, A. C. Porter, L. Desimone, B. F. Birman, and K. S. Yoon, "What makes professional development effective? Results from a national sample of teachers", *American Educational Research Journal,* vol. 38, no. 4, pp. 915–945, 2001. https://doi.org/10.3102/00028312038004915

[26] W. R. Penuel, B. J. Fishman, R. Yamaguchi, and L. P. Gallagher, "What makes professional development effective? Strategies that foster curriculum implementation", *American Educational Research Journal,* vol. 44, no. 4, pp. 921–958, 2007. https://doi.org/10.3102/0002831207308221

[27] L. Darling-Hammond, M. E. Hyler, and M. Gardner, Effective Teacher Professional Development (Number June). 2017. https://doi.org/10.54300/122.311

[28] E. López-Martín, B. Gutiérrez-de-Rozas, A. M. González-Benito, and E. Expósito-Casas, "Why Do Teachers Matter ? A Meta-Analytic Review of how Teacher Characteristics and Competencies Affect Students ' Academic Achievement", *International Journal of Educational Research,* vol. 120, no. June, 2023. https://doi.org/10.1016/j.ijer.2023.102199

[29] O. M. Ventista and C. Brown, "Teachers' professional learning and its impact on students ' learning outcomes : Findings from a systematic review", *Social Sciences & Humanities Open,* vol. 8, no. 1, 100565, 2023. https://doi.org/10.1016/j.ssaho.2023.100565

[30] R. Gutiérrez, "The sociopolitical turn in mathematics education", *Journal for Research in Mathematics Education,* vol. 44, no. 1, pp. 37–68, 2013. https://doi.org/10.5951/jresematheduc.44.1.0037

[31] J. Hattie and H. Timperley, "The power of feedback", *Review of Educational Research,* vol. 77, no. 1, pp. 81–112, 2007. https://doi.org/10.3102/003465430298487

[32] R. E. Mayer, "Should there be a three-strikes rule against pure discovery learning? The case for guided methods of instruction", *American Psychologist,* vol. 59, no. 1, pp. 14–19, 2004. https://doi.org/10.1037/0003-066X.59.1.14

[33] S. Almaidah, F. N. Asri, Y. Wicaksono, M. H. N. Arifin, A. S. Nurhalimah, and N. J. Fratiwi, "Inovasi Pembelajaran Inklusif Melalui Ethnophysics Flashcard Berbasis Augmented Reality di SLB Agro Industri Cisarua", *JPMB Jurnal Pengabdian Masyarakat Borneo,* vol. 9, no. 3, pp. 140–147, 2025. https://doi.org/10.35334/fnwmhy46

[34] P. D. Damayanti, A. M. B. Harsono, and A. Suriansyah, "Strategi Guru dalam Menerapkan Prinsip Deep Learning Melalui Kearifan Lokal Banjarmasin di Sekolah Dasar", *Pendas: Jurnal Ilmiah Pendidikan Dasar,* vol. 10, no. 4, 2025. https://doi.org/10.23969/jp.v10i04.38317

[35] A. Fauzan et al., "Realistic Mathematics Education (RME) to Improve Literacy and Numeracy Skills of Elementary School Students Based on Teachers' Experience", *Journal of Mathematics Education,* vol. 13, no. 2, pp. 301–316, 2024. https://doi.org/10.22460/infinity.v13i2.p301-316

[36] D. Irawan and Y. Cahyan, "Media Ajar Digital Berbasis Kearifan Lokal Kasumedangan: Analisis Kebutuhan Guru untuk Pembelajaran Menulis Berpendekatan Deep Learning", *Semantik,* vol. 15, no. 1, pp. 33–44, 2026. https://doi.org/10.22460/semantik.v15i1.p33-44

[37] A. R. Matondang, "Contextual Mathematics Learning Based on Local Wisdom to Improve Understanding of the Concept of Multiplication in Elementary School Students Pembelajaran Matematika Kontekstual Berbasis Kearifan Lokal untuk Meningkatkan Pemahaman Konsep Perkalian pada Si", *Jurnal Keilmuan Pendidikan Matematika,* vol. 4, no. 3, pp. 248–254, 2026. https://doi.org/10.47662/jkpm.v4i3.1110

[38] Y. Rahmawati, A. Mu'ti, S. Suyanto, N. Luthfi, and N. L. R. Herianingtyas, "Pembelajaran Mendalam: Transformasi Pembelajaran Menuju Pendidikan Bermutu", *Jurnal Penelitian Kebijakan Pendidikan,* vol. 18, no. 1, pp. 1–16, 2025. https://doi.org/10.24832/jpkp.v18i1.1281

[39] S. Freeman et al., "Active learning increases student performance in science, engineering, and mathematics", *Proceedings of the National Academy of Sciences,* vol. 111, no. 23, pp. 8410–8415, 2014. https://doi.org/10.1073/pnas.1319030111

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

[41] R. R. Hake, "Interactive-engagement versus traditional methods : A six-thousand-student survey of mechanics test data for introductory physics courses", *American Journal of Physics,* vol. 66, no. 1, pp. 64–74, 1998. https://doi.org/10.1119/1.18809

[42] W. J. Boone, J. R. Staver, and M. S. Yale, Rasch Analysis in the Human Sciences. Springer. 2014. https://doi.org/10.1007/978-94-007-6857-4

[43] V. Braun and V. Clarke, "Using thematic analysis in psychology", *Qualitative Research in Psychology,* vol. 3, no. 2, pp. 77–101, 2006. https://doi.org/10.1191/1478088706qp063oa

[44] G. Cumming, "The new statistics: Why and how", *Psychological Science,* vol. 25, no. 1, pp. 7–29, 2014. https://doi.org/10.1177/0956797613504966

[45] H. M. Ernst, A. Prinz-weiß, J. Wittwer, and H. M. Ernst, Discrepancy between performance and feedback affects mathematics student teachers ' self-efficacy but not their self-assessment accuracy. (January). 2025. https://doi.org/10.3389/fpsyg.2024.1391093

[46] R. L. Wasserstein and N. A. Lazar, "The ASA's statement on p-values: Context, process, and purpose", *The American Statistician,* vol. 70, no. 2, pp. 129–133, 2016. https://doi.org/10.1080/00031305.2016.1154108

[47] S. B. Morris and R. P. DeShon, "Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs", *Psychological Methods,* vol. 7, no. 1, pp. 105–125, 2002. https://doi.org/10.1037/1082-989X.7.1.105

[48] A. Humaidi, D. Huriaty, and R. L. Agustina, "Strengthening Literacy and Numeracy Competencies Based on Local Wisdom for Teachers to Design Contextual Learning", *Bubungan Tinggi: Jurnal Pengabdian Masyarakat,* vol. 6, no. April 2024, pp. 538–548, 2024. https://doi.org/10.20527/btjpm.v6i3.10428

[49] I. M. Pratiwi and V. A. N. Ariawan, "Introducing local wisdom by numbering and reading", *Journal of Physics: Conference Series,* vol. 1521, no. 3, 032013, 2020. https://doi.org/10.1088/1742-6596/1521/3/032013

[50] V. Munafiroh, S. P. Kawuryan, and S. Munawwarah, "The effect of discovery learning integrated with Lampung local wisdom on students' mathematical problem-solving

skills across gender", *Jurnal Penelitian Ilmu Pendidikan,* vol. 17, no. 2, pp. 99–106, 2024. https://doi.org/10.21831/jpip.v17i2.90981