EDUMATSAINS
JURNAL PENDIDIKAN MATEMATIKA DAN SAINS
E-ISSN:2527-7235
Volume 11 Issue. 2, October 2026, pp 241-257
http://ejournal.uki.ac.id/index.php/edumatsains

Teachers' Needs Analysis for Designing an Inclusive Mathematics Assessment Training Program

N. Sinta Mustika Aryanti[1*], Zainal Arifin[2], Budi Setiawan[3]

[1,2,3]Pengembangan Kurikulum FIP Universitas Pendidikan Indonesia

e-mail: sintamustikaaryanti@upi.edu

Article Info

Article history:
Received: August 30th, 2026
Revised: October 25th, 2026
Accepted: October 26th, 2026
Available online: October 31st, 2026

https://doi.org/10.33541/edumatsains.v11i2.8715

Abstract

Inclusive mathematics classrooms require teachers to collect, interpret, and act on assessment evidence from students with diverse needs, yet assessment training programs are seldom designed from systematic evidence of what teachers already do and still need. Using the analysis and exploration phase of Educational Design Research (EDR), this study identified teachers' needs as empirical requirements for designing an assessment training program in inclusive mathematics education. Data were collected from 42 elementary school teachers in Bandung, Indonesia, through a needs-analysis questionnaire and from three public inclusive elementary schools through document analysis and semi-structured interviews with principals. Questionnaire data were analyzed descriptively, qualitative data were analyzed thematically, and all sources were triangulated. Teachers reported frequent assessment practice across five domains, and most had never attended training specific to inclusive mathematics assessment. Documents and interviews, however, showed that assessment evidence was seldom recorded, interpreted, and used systematically for individualized instruction. Diagnostic assessment was both the strongest perceived competency need and the most frequently selected training content, followed by assessment adaptation and differentiation. Teachers preferred hands-on workshops and usable modules, and most requested post-training mentoring. These convergent findings were translated into requirements for a training program that is practical (teachers produce assessment artifacts), contextualized (authentic school cases), adaptive (adjustment to student characteristics), and sustainable (a one-day workshop followed by mentoring). The study contributes a triangulated, design-oriented needs-analysis model that links reported practice,

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

documented gaps, and teacher preferences to the subsequent EDR design phase.

Keywords: educational design research, needs analysis, assessment training program, inclusive mathematics education, elementary school teachers

1. Introduction

Inclusive education requires teachers to give every student an equitable opportunity to demonstrate learning. In mathematics, learner diversity involves differences in prior knowledge, conceptual understanding, communication, and ways of demonstrating competence. Indonesian teachers in inclusive elementary schools already use various accommodations but report difficulty in meeting students' needs because of limited pedagogical knowledge (Azizah et al., 2024). This difficulty is most consequential in assessment: without valid evidence of what each student knows and how each student can show it, instructional adjustment remains guesswork.

The assessment challenge is twofold. First, assessment must be flexible. Teachers can provide multiple ways for students to demonstrate mathematical understanding (Porta et al., 2026), and designs grounded in Universal Design for Learning can widen these options (Aryanti et al., 2026), although the consistency of teachers' adjustments varies (Dewodo & Agbetorwoka, 2025). Second, assessment evidence must be used: the effectiveness of formative assessment depends on connecting evidence with learning intentions and subsequent instruction (Kang & Lam, 2024; Maskos et al., 2025). In Indonesia, diagnostic assessment is expected to identify students' initial abilities, yet teachers' readiness to implement it varies (Jannah et al., 2025), and teachers find it difficult to map individual abilities from the results (Siburian & Mukhlis, 2024).

Both demands concern teacher assessment literacy, the capacity to design, implement, interpret, and use assessment in context (Pastore, 2023; Xu & Brown, 2016). Assessment literacy is situated and multidimensional (DeLuca et al., 2019; Estaji et al., 2024; Fu, 2024; Hull & Vígh, 2025) and, in diverse classrooms, requires sensitivity to learner diversity (Wiese & Nortvedt, 2023). Teachers in inclusive settings report ongoing professional learning and curriculum adaptation among their most frequent support needs (Chow et al., 2024), and prospective mathematics teachers feel least ready to adapt teaching methods (Chiani et al., 2025). The competency gap is therefore specific: adapting assessment and converting its evidence into individualized instructional decisions.

In this article, *professional development* refers to the broader, continuing process of teachers' professional learning, whereas *training program* refers to the structured intervention designed in this study: a core workshop with planned post-training mentoring. Effective professional development offers relevant content, active engagement, contextualization, and continued support (Geng & Mohd Yusoff, 2022; Kennedy, 2016; Pastore & Mincu, 2024; Sims & Fletcher-Wood, 2021) and can strengthen teachers' formative assessment literacy (Li et al., 2023), whereas short in-service courses on inclusion are difficult to sustain (Nseibo et al., 2022). Because teachers value training that responds to their perceived needs (Estaji, 2024; Giraldo et al., 2023; Umam & Indah,

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

2020), evidence of their current practices, gaps, and preferences is needed before a program is designed.

Previous studies fall into three strands. First, needs analyses of assessment literacy have been conducted outside inclusive mathematics, for example with English for Academic Purposes instructors (Estaji, 2024) and teachers of English to young learners (Umam & Indah, 2020). Second, studies of inclusive education have examined teachers' general support needs (Chow et al., 2024) and pedagogical competencies (Azizah et al., 2024) rather than their assessment needs in mathematics. Third, Indonesian studies have been intervention-oriented, developing diagnostic mathematics instruments (Deviana et al., 2024), assisting teachers in compiling them (Roisatul MN et al., 2024), or training teachers in a special school (Faradiba et al., 2024); Jannah et al. (2025) examined readiness for diagnostic assessment without focusing on inclusive mathematics, and Li et al. (2023) evaluated an existing program rather than analyzing needs before design. None of these studies combined reported practices, documentary evidence, perceived needs, training preferences, and school-level perspectives into requirements for designing an assessment training program in inclusive mathematics.

This study therefore conducted the analysis and exploration phase of an Educational Design Research (EDR) project to answer four questions: (1) What assessment practices do teachers report? (2) What competency gaps emerge when these reports are compared with documents and interviews? (3) What training content, format, resources, and support do teachers need? (4) What design requirements follow? The study contributes empirically, through a triangulated profile of assessment practice and needs; methodologically, through a needs-analysis model that distinguishes current practice from training need; and practically, through a design-requirements matrix for the subsequent design phase.

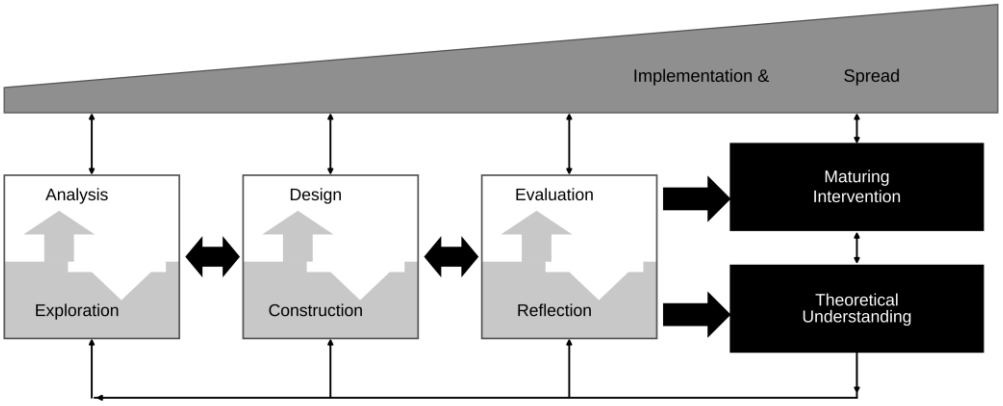
2. Methods

2.1 Research Design

EDR integrates the investigation of an authentic educational problem with the development of a practical intervention and the generation of design knowledge (McKenney & Reeves, 2012). In the generic model of McKenney and Reeves (2012) (Figure 1), three flexible and iterative phases (analysis and exploration, design and construction, and evaluation and reflection) interact with practice through implementation and spread and yield a maturing intervention and theoretical understanding. In this project, the first phase comprises problem analysis, needs analysis, and the formulation of design requirements; the second, prototype development; and the third, expert validation, a limited try-out, and refinement.

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

Figure 1. Generic model for conducting educational design research



Source: McKenney and Reeves (2012).

This article reports the analysis and exploration phase for three reasons. EDR is problem-driven, so an intervention must be grounded in a well-understood problem; no contextual evidence was available on how the participating teachers implement, document, and use assessment; and the resulting design requirements serve as criteria for prototype development and subsequent expert validation. The later phases will be reported separately.

2.2 Participants and Research Context

The population consisted of elementary school teachers in Bandung, Indonesia, where the City Education Office has stated that all schools are inclusive (Ramdani, 2021), although implementation quality varies (Ratri & Herawati, 2024). Teachers were selected using purposive sampling: the questionnaire was distributed openly to elementary teachers, and only responses from teachers who met the inclusion criterion were analyzed, namely classroom teachers teaching mathematics in Grades 1–6 in inclusive schools whose classes include learners with diverse characteristics and needs. The sample comprised 42 teachers from 23 schools (22 public elementary schools and one *madrasah ibtidaiyah*). Most were female (78.6%) and had more than ten years of teaching experience (64.3%), and all grade levels were represented (Table 1).

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

Table 1. Characteristics of participating teachers (N = 42)

Characteristic	Category	n	%
Gender	Female	33	78.6
	Male	9	21.4
Grade level taught	Grade 1	7	16.7
	Grade 2	6	14.3
	Grade 3	7	16.7
	Grade 4	5	11.9
	Grade 5	7	16.7
	Grade 6	10	23.8
Teaching experience	< 5 years	5	11.9
	5–10 years	10	23.8
	> 10 years	27	64.3
Specific training on assessment for inclusive mathematics	Never attended	33	78.6
	Attended	9	21.4

Three public elementary schools were selected purposively for document analysis and principal interviews because they enrolled diverse learners, including students with special educational needs, in regular classes and agreed to provide access to their assessment documents. Participation was voluntary and based on informed consent, with the permission of the participating schools. The schools are referred to as Schools A, B, and C.

2.3 Instruments

Questionnaire. Indicators were derived from the literature on assessment literacy and inclusive assessment and from the curriculum's assessment stages and organized into a blueprint (Table 2). Content validity was established by two experts, a curriculum development expert and an education supervisor (*penilik*) from the Bandung City Education Office, who rated all research instruments (questionnaire, document-analysis checklist, and interview protocol) on ten indicators covering clarity, content accuracy, measurability, language, and usability using a four-point scale (1 = not appropriate to 4 = highly appropriate). Aiken's V (Aiken, 1985) ranged from 0.67 to 1.00 across the indicators (mean V = 0.80), and both experts judged the instruments suitable for use with minor revisions. The lowest values concerned the clarity of the blueprint indicators, measurability, and usability (V = 0.67); accordingly, the blueprint and items were aligned, indicator statements were reworded, question wording was simplified and matched with the answer options, and terminology was made consistent. The questionnaire contained 30 items in three parts: (A) 15 items on factual practice on a four-point Likert scale (1 = never, 2 = rarely, 3 = frequently, 4 = consistently); (B) 10 items on perceived competency needs on a four-point Likert scale (1 = not needed at all, 2 = not needed, 3 = needed, 4 = very needed), with responses reported in three categories (not needed, needed, very needed) because no teacher selected "not needed at all"; and (C) 5 items on training

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

preferences (content, methods, and resources, with multiple responses allowed; duration; and post-training assistance). Teacher characteristics were recorded in an identity section. Item validity and reliability were examined using the responses of the 42 teachers. In Part A, corrected item–total correlations ranged from 0.46 to 0.81, all above the critical value ($r = 0.304$; $df = 40$; $\alpha = .05$), and Cronbach's α was 0.92. In Part B, nine items had corrected item–total correlations of 0.31–0.80 and Cronbach's α was 0.81. The item on using results for instructional follow-up did not correlate with the other need items ($r = -0.04$); it was retained for descriptive reporting because it captures a distinct aspect of need (Section 3.5), and removing it would raise α to 0.86. Part C, which consists of nominal choices, was not subjected to these analyses.

Table 2. Blueprint of the needs-analysis questionnaire

Part	Domain	Indicators	Items
A. Factual practice	Diagnostic assessment	Identifying initial abilities, difficulties, and learning needs	A1–A3
	Formative assessment	Monitoring progress; providing feedback	A4–A5
	Summative assessment	Assessing and evaluating attainment of learning objectives	A6–A7
	Adaptation, differentiation, and accessibility	Adjusting techniques, forms, and difficulty; differentiating tasks; providing accessible assessment	A8–A12
	Use of assessment results	Using results for remediation, enrichment, and improving instruction	A13–A15
B. Perceived needs	Ten competencies	Listed in Table 4	B1–B10
C. Training preferences	Program characteristics	Content, methods, resources, duration, assistance	C1–C5

Document analysis. Following Bowen (2009), documents from each school for the 2026/2027 academic year (learning modules, assessment instruments, diagnostic materials, formative and summative records, rubrics and scoring guidelines, remedial and enrichment plans, follow-up records, and individual learning programs) were appraised with a 16-component checklist based on the assessment cycle and coded as *complete*, *partial*, or *not found*.

Interviews. Semi-structured interviews with the three principals followed a protocol of 20 main questions on the implementation of diagnostic, formative, and summative assessment; alignment of objectives, indicators, and instruments; item forms, rubrics, and scoring guidelines; adaptation, differentiation, and accessibility; use of results for remediation, enrichment, and follow-up; and teachers' competency needs and expected training design. The protocol was validated by the same two experts as part of the instrument validation. The interviews were conducted face-to-face by the first author between July and August 2026 and were audio-recorded with the principals' consent; responses to each question were documented in interview records together with the researcher's initial interpretation.

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

2.4 Data Analysis

Questionnaire data were analyzed using frequencies, percentages, and means. Domain scores were the means of their items and were interpreted with equal intervals (width = (4 − 1)/4 = 0.75): 1.00–1.75 never, 1.76–2.50 rarely, 2.51–3.25 frequent, and 3.26–4.00 very frequent. Competency needs were ranked with a weighted mean score, WMS = (4 × n very needed + 3 × n needed + 2 × n not needed + 1 × n not needed at all)/N (range 1.00–4.00), which reflects the full response distribution. Interview transcripts and document notes were analyzed using reflexive thematic analysis (Braun & Clarke, 2021) through familiarization, coding, generating, reviewing, and defining themes, and reporting, combining deductive codes from the assessment cycle with inductive codes. All coding was conducted by the first author. Consistent with reflexive thematic analysis, which treats coding as an interpretive process rather than one requiring inter-coder agreement (Braun & Clarke, 2021), no agreement coefficient was calculated. Trustworthiness (Nowell et al., 2017) was supported by triangulation across the questionnaire, documents, and interviews and by an audit trail consisting of the codebook, analytic memos, and the document checklist.
Finally, a triangulation matrix compared reported practice, perceived need, documentary evidence, and principals' views for each competency. A competency was classified as a *convergent need* when several sources indicated a gap and as a *hidden need* when teachers rated it as less needed but documents or interviews revealed a gap. Thus, a low practice score was not automatically treated as the highest priority.

3. Result and Discussion

3.1 Context for Training Program Design

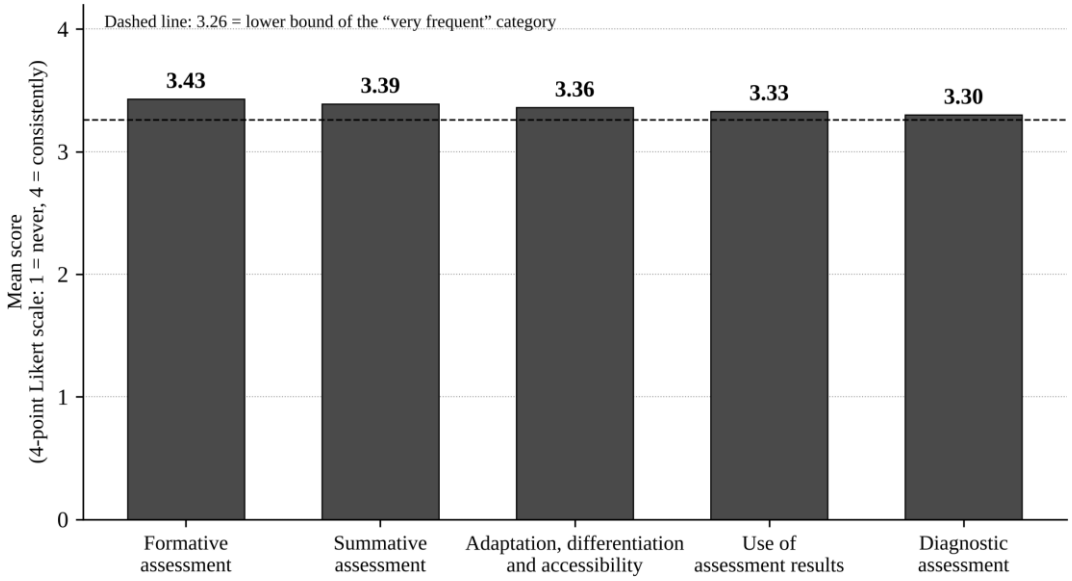
Of the 42 teachers, 33 (78.6%) had never attended training specific to assessment in inclusive mathematics (Table 1), although they reported frequent assessment practice (Section 3.2). This limited access may reflect the short and inconsistent nature of in-service training for inclusion (Nseibo et al., 2022), teachers' unmet need for ongoing professional learning (Chow et al., 2024), limited preparation in initial teacher education (Chiani et al., 2025), and the focus of existing diagnostic-assessment training on differentiated learning rather than inclusive mathematics (Deviana et al., 2024; Roisatul MN et al., 2024). The program should therefore build on existing practice rather than assume no prior experience, and it should be easy to attend and continue at school level.

3.2 Assessment Practice as the Starting Point

All five practice domains fell within the very frequent category (M = 3.30–3.43; Figure 2), with formative assessment highest (3.43) and diagnostic assessment lowest (3.30).

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

Figure 2. Teachers reported factual assessment practices by domain



Note. N = 42. Domain scores are item means (Table 2, Part A) on a four-point Likert scale (1 = never, 4 = consistently). The dashed line marks the lower bound of the "very frequent" category (3.26).

Because the range is narrow (0.13), the profile was treated as a baseline rather than a ranking of needs, and the lowest score alone was not interpreted as the principal need. Even for diagnostic assessment, 95.2% of teachers reported frequently or consistently identifying initial abilities, 85.7% identifying learning difficulties, and 92.9% using results to determine learning needs. The design issue is therefore whether teachers can document, interpret, and use diagnostic evidence systematically.

3.3 Evidence of Competency Gaps

Comparing teachers' reports with documentary evidence revealed the key design consideration. Table 3 presents the results of the 16-component document checklist for the three schools.

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

Table 3. Documentary evidence of assessment practice across Schools A, B, and C

Checklist component	School A	School B	School C
1. Diagnostic assessment of initial abilities, difficulties, and needs	Partial	Partial	Partial
2. Formative assessment records of learning progress	Partial	Complete	Complete
3. Summative assessment of attainment	Partial	Partial	Not found
4. Consideration of students' characteristics and needs	Partial	Partial	Complete
5. Alignment of learning objectives with assessment indicators	Complete	Complete	Complete
6. Blueprint mapping indicators to items	Complete	Not found	Not found
7. Assessment form suited to learning objectives	Complete	Complete	Complete
8. Instrument suited to the targeted competencies	Complete	Complete	Partial
9. Item form suited to the targeted competencies	Complete	Complete	Partial
10. Rubrics and scoring guidelines	Complete	Partial	Partial
11. Assessment form adapted to students' characteristics	Complete	Partial	Complete
12. Differentiated assessment	Complete	Not found	Complete
13. Accessibility adjustments	Complete	Partial	Complete
14. Results used for remediation	Complete	Partial	Partial
15. Results used for enrichment	Partial	Not found	Not found
16. Results used for instructional follow-up	Partial	Partial	Complete

Note. Complete = present and complete; Partial = present but incomplete; Not found = absent from the documents analyzed.

All three schools aligned assessment forms and indicators with learning objectives (components 5 and 7), but diagnostic documentation was only partial in every school (component 1), and the use of results for enrichment was partial or absent (component 15). The schools also differed. School A had a blueprint, adaptive and differentiated instruments, rubrics, and remedial and enrichment programs, yet its recapitulation sheets for formative and summative results and its lists of students receiving remediation or enrichment were left blank, so there was no evidence that actual results had informed decisions. School B lacked a blueprint, differentiated assessment, and enrichment records, and individual diagnostic results were not found. School C's individual learning program planned adaptation, differentiation, and accessibility and linked evaluation to follow-up, but it had no blueprint, no summative instrument, and an initial-assessment section that was not completed in detail. The principals likewise emphasized teachers' need to identify student needs, process assessment information, develop instruments and rubrics, and use results to adjust instruction, consistent with evidence that Indonesian teachers find it difficult to map individual abilities from diagnostic assessment (Siburian & Mukhlis, 2024). The central gap is thus not the absence of assessment but the limited recording and use of assessment evidence for individualized decisions, so the training should follow an assessment cycle that links planning, evidence collection, recording, interpretation, feedback, and follow-up.

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

3.4 Competency Priorities

Table 4. Teachers perceived competency needs (N = 42)

Competency	Very needed n (%)	Needed n (%)	Not needed n (%)	WMS	Rank
Diagnostic assessment	28 (66.7)	14 (33.3)	0 (0.0)	3.67	1
Differentiated assessment	22 (52.4)	19 (45.2)	1 (2.4)	3.50	2
Assessment instrument development	20 (47.6)	21 (50.0)	1 (2.4)	3.45	3
Analysis for remediation	19 (45.2)	23 (54.8)	0 (0.0)	3.45	3
Assessment adaptation	19 (45.2)	20 (47.6)	3 (7.1)	3.38	5
Analysis for enrichment	15 (35.7)	26 (61.9)	1 (2.4)	3.33	6
Summative assessment	13 (31.0)	28 (66.7)	1 (2.4)	3.29	7
Accessible assessment	13 (31.0)	26 (61.9)	3 (7.1)	3.24	8
Formative assessment and feedback	10 (23.8)	30 (71.4)	2 (4.8)	3.19	9
Use of results for instructional follow-up	13 (31.0)	22 (52.4)	7 (16.7)	3.14	10

Note. WMS = weighted mean score (range 1.00–4.00); tied scores share the same rank.

Diagnostic assessment was the strongest perceived need (WMS = 3.67), followed by differentiated assessment (3.50), instrument development, and analysis for remediation (both 3.45). Use of results for instructional follow-up ranked lowest (3.14) and was the only competency that a notable proportion of teachers (16.7%) considered not needed.

3.5 Priority Content for the Training Program

Table 5. Training content selected by teachers compared with perceived competency needs

Training content	n	%	Rank	Rank in Table 4
Diagnostic mathematics assessment	25	59.5	1	1
Assessment adaptation according to student characteristics	21	50.0	2	5
Mathematics assessment instrument development	12	28.6	3	3
Use of assessment results for instructional follow-up	9	21.4	4	10
Differentiated mathematics assessment	9	21.4	4	2
Assessment-result analysis	8	19.0	6	3 and 6
Formative assessment	6	14.3	7	9
Summative assessment	4	9.5	8	7

Note. Multiple responses were allowed; percentages do not total 100%. Assessment-result analysis corresponds to analysis for remediation and enrichment in Table 4.

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

Comparing Tables 4 and 5 shows three patterns. First, diagnostic assessment ranked first in both tables and instrument development third, providing convergent evidence for a core diagnostic module. Second, assessment adaptation was the second most selected content and differentiated assessment the second strongest need; as complementary responses to diversity, they should form one module. Third, use of results for follow-up is a hidden need: it ranked lowest in perceived need and, unlike the other need items, did not correlate with them (Section 2.3), yet 21.4% of teachers selected it as content and documents showed that the recorded use of results, particularly for enrichment, was partial or absent in all three schools (Table 3), so it should be integrated across modules. Formative and summative assessment ranked low in both tables but had the highest reported practice (Figure 2), suggesting familiarity; they should serve as a foundation for diagnostic and adaptive tasks rather than as separate modules. Such needs-based tailoring characterizes effective assessment-literacy programs (Estaji, 2024; Geng & Mohd Yusoff, 2022; Pastore & Mincu, 2024).

3.6 Training Format, Resources, and Support

Teachers preferred hands-on workshops (81.0%), practical instrument development (64.3%), and training modules (81.0%), followed by sample instruments (47.6%), case studies (19.0%), video tutorials (16.7%), and templates or blueprints (14.3%). For the face-to-face component, 66.7% preferred one day, 26.2% two days, and 7.1% three days, while 26 teachers (61.9%) stated that they needed post-training assistance and 16 (38.1%) did not. The preferred face-to-face duration must therefore be distinguished from the total duration of professional development: because most teachers wanted continued support after a short workshop, the results support a blended, sustained model in which a one-day workshop is followed by structured mentoring, rather than a one-day program as such. Because more than one-third of teachers did not request assistance, mentoring should be flexible, with core follow-up assignments for all participants and more intensive support for those who need it.

The assistance is specified with five components, whose intensity will be set in the design phase: (1) follow-up assignments in which teachers develop or adapt an instrument, use it in class, and record the results; (2) online mentoring with feedback on these products; (3) peer discussion of cases and products; (4) school-based coaching by the facilitator or principal; and (5) a professional learning community, ideally within existing teacher working groups. This model reflects evidence that effective professional development provides continued support for application (Kennedy, 2016; Li et al., 2023; Sims & Fletcher-Wood, 2021), that one-time courses are difficult to sustain (Nseibo et al., 2022), and that teachers' perceptions can inform continuing professional development (Giraldo et al., 2023; Umam & Indah, 2020).

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

3.7 Initial Design Requirements

Table 6 and Figure 3 summarize how the findings were translated into design requirements.

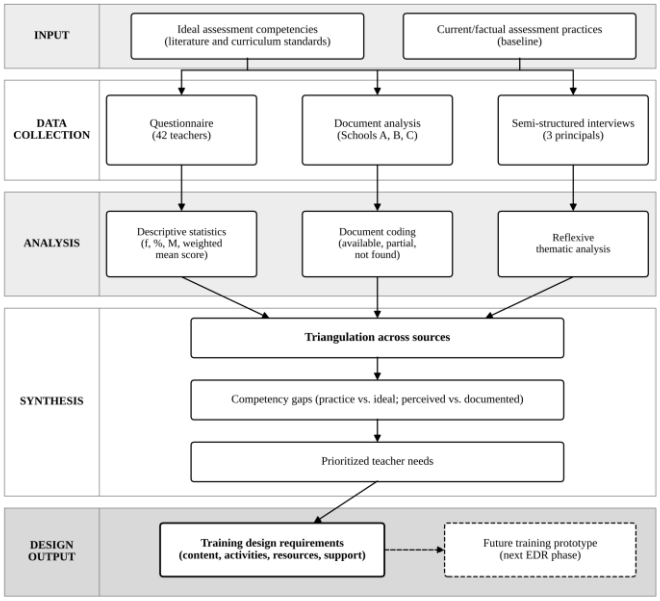
Table 6. Synthesis of teachers' needs and initial training design requirements

Identified condition/gap	Evidence	Design implication
Diagnostic results not systematically documented and analyzed	Q, D, I	Core module on diagnostic assessment and learner-needs mapping
Adaptation not consistently individualized; differentiation varies	Q, D, I	Combined module on adaptive and differentiated assessment
Instruments and rubrics vary in completeness	D, I	Practice in developing blueprints, items, rubrics, and scoring guidelines
Evidence not transformed into instructional information	D, I	Practice in interpreting assessment results
Use of results for enrichment and follow-up weakly documented (hidden need)	D, I; low need in Q	Feedback, remediation, enrichment, and follow-up across all modules
Preference for practical learning and usable resources	Q, I	Hands-on workshop with authentic cases; module, sample instruments, templates
Most teachers request support after training (61.9%)	Q, I	One-day workshop followed by flexible mentoring and a professional learning community

Note. Q = questionnaire; D = document analysis; I = interviews.

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

Figure 3. Needs-analysis process and derivation of training design requirements



The program should be organized around an integrated assessment cycle: identifying learner needs, planning, instrument development, adaptation and differentiation, evidence collection and recording, interpretation, feedback, remediation or enrichment, and follow-up. Its four characteristics follow from the findings: it is *practical* because teachers preferred hands-on learning, *contextualized* because documentary gaps differed across schools (Table 3), *adaptive* because adaptation and differentiation were priority needs (Tables 4 and 5), and *sustainable* because most teachers (61.9%) requested post-training assistance and two principals noted that a single training session would not be sufficient. Compared with earlier needs analyses based mainly on teachers' perceptions (Estaji, 2024), triangulation revealed a gap that self-reports alone would have missed; compared with Indonesian studies that developed instruments or assistance programs (Deviana et al., 2024; Roisatul MN et al., 2024), this study specifies what a program must contain before it is developed. These findings are initial design requirements rather than a completed program, and the next phase will use them to develop and iteratively refine a prototype (McKenney & Reeves, 2012).

4. Conclusion

4.1 Conclusion and Contribution

Teachers reported frequent assessment practice, but documents and interviews revealed gaps in documenting, interpreting, and using assessment evidence, in individualized adaptation, and in rubric development. Diagnostic assessment emerged as the core content, supported by adaptation

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

and differentiation, instrument development, result analysis, and instructional follow-up, to be delivered through hands-on workshops, usable modules, and mentoring. The main contribution is a triangulated, design-oriented needs-analysis model that distinguishes practice from need, reveals hidden needs, and yields a design-requirements matrix for the next EDR phase.

4.2 Practical Implications

Assessment training for inclusive mathematics should be organized around the full assessment cycle with diagnostic assessment at its core, require teachers to produce and use assessment artifacts based on authentic school cases, and follow a blended model of a one-day workshop, mentoring, and a professional learning community. Principals should support the systematic recording of individual assessment results and follow-up decisions.

4.3 Limitations and Future Research

The sample was small (42 teachers) and drawn from one city; qualitative data were coded by a single researcher; questionnaire data were self-reported, and item validity and reliability were estimated from the main-study responses rather than from a separate try-out; document analysis and interviews covered only three schools and their principals; teacher characteristics were limited to gender, grade level, and teaching experience; the analysis was descriptive; and the program's effectiveness was not evaluated. Future research should develop, validate, try out, and evaluate the program, using larger samples and classroom observations.

5. Acknowledgments

The authors thank the participating teachers and principals for their time and cooperation.

6. References

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

Aryanti, N. S. M., Munir, Famelia, & Yunita, S. (2026). Exploring the potential alignment between VR/AR and Universal Design for Learning for mathematics accessibility in elementary schools: A systematic literature review. *EduMatSains: Jurnal Pendidikan, Matematika dan Sains*, *10*(4), 54–68. https://doi.org/10.33541/edumatsains.v10i4.7632

Azizah, N., Mumpuniarti, Rudiyati, S., & Evans, D. (2024). Elementary teachers' pedagogical competencies in supporting students with learning difficulties. *International Journal of Evaluation and Research in Education*, *13*(2), 723–730. https://doi.org/10.11591/ijere.v13i2.26345

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal*, *9*(2), 27–40. https://doi.org/10.3316/QRJ0902027

Braun, V., & Clarke, V. (2021). One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative Research in Psychology*, *18*(3), 328–352. https://doi.org/10.1080/14780887.2020.1769238

Chiani, S. H., Putriawati, W., & Rahmawati. (2025). The influence of student development understanding on mathematics education students' readiness to manage diverse classrooms. *EduMatSains: Jurnal Pendidikan, Matematika dan Sains*, *10*(2), 52–60. https://doi.org/10.33541/edumatsains.v10i2.6955

Chow, W. S. E., de Bruin, K., & Sharma, U. (2024). A scoping review of perceived support needs of teachers for implementing inclusive education. *International Journal of Inclusive Education*, *28*(13), 3321–3340. https://doi.org/10.1080/13603116.2023.2244956

DeLuca, C., Coombs, A., MacGregor, S., & Rasooli, A. (2019). Toward a differential and situated view of assessment literacy: Studying teachers' responses to classroom assessment scenarios. *Frontiers in Education*, *4*, 94. https://doi.org/10.3389/feduc.2019.00094

Deviana, T., Sulistyani, N., Roisatul MN, F., & Maulyda, M. A. (2024). Pengembangan instrumen asesmen diagnostik matematika untuk mendesain pembelajaran berdiferensiasi di sekolah dasar. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, *13*(4), 1269–1280. https://doi.org/10.24127/ajpm.v13i4.9773

Dewodo, C. Y., & Agbetorwoka, A. (2025). Teachers' adaptation of assessment practices for learners with hearing impairments in inclusive schools, Ketu South, Ghana. *International Journal of Studies in Inclusive Education*, *2*(3), 23–29. https://doi.org/10.38140/ijsie.v2i3.2179

Estaji, M. (2024). Perceived need for a teacher education course on assessment literacy development: Insights from EAP instructors. *Asian-Pacific Journal of Second and Foreign Language Education*, *9*, 50. https://doi.org/10.1186/s40862-024-00272-2

Estaji, M., Banitalebi, Z., & Brown, G. T. L. (2024). The key competencies and components of teacher assessment literacy in digital environments: A scoping review. *Teaching and Teacher Education*, *141*, 104497. https://doi.org/10.1016/j.tate.2024.104497

Faradiba, S. S., Sunismi, Ilmi, Y. I. N., & Fuat. (2024). Empowering special needs children: Community outreach for inclusive mathematics education. *International Journal of Community Service Learning*, *8*(3), 226–235. https://doi.org/10.23887/ijcsl.v8i3.80816

Fu, R. (2024). Exploring K–12 teachers' assessment literacy and self-efficacy in China. *Canadian Journal for New Scholars in Education*, *15*(2). https://journalhosting.ucalgary.ca/index.php/cjnse/article/view/79917

Geng, J., & Mohd Yusoff, N. (2022). The shared features of effective improvement programmes for teachers' assessment literacy. *Pedagogies: An International Journal*, *18*(4), 710–727. https://doi.org/10.1080/1554480X.2022.2084398

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

Giraldo, F., Escalante-Villa, D., & Isaza-Palacio, D. (2023). English teachers' perceptions of a language assessment literacy course. *HOW*, *30*(2). https://doi.org/10.19183/how.30.2.750

Hull, P., & Vígh, T. (2025). Teachers' assessment literacy: A descriptive literature review. *Hungarian Educational Research Journal*, *15*(2), 210–225. https://doi.org/10.1556/063.2024.00317

Jannah, M., Maryani, I., & Santosa, A. B. (2025). Kesiapan guru sekolah dasar dalam implementasi asesmen diagnostik untuk mendukung pembelajaran berdiferensiasi. *Ideguru: Jurnal Karya Ilmiah Guru*, *10*(1). https://doi.org/10.51169/ideguru.v10i1.1467

Kang, M., & Lam, R. (2024). Understanding university English instructors' assessment literacy: A formative assessment perspective. *Language Testing in Asia*, *14*, 52. https://doi.org/10.1186/s40468-024-00323-y

Kennedy, M. M. (2016). How does professional development improve teaching? *Review of Educational Research*, *86*(4), 945–980. https://doi.org/10.3102/0034654315626800

Li, Z., Yan, Z., Chan, K. K. Y., Zhan, Y., & Guo, W. Y. (2023). The role of a professional development program in improving primary teachers' formative assessment literacy. *Teacher Development*, *27*(4), 447–467. https://doi.org/10.1080/13664530.2023.2223595

Maskos, K., Schulz, A., Oeksuez, S. S., & Rakoczy, K. (2025). Formative assessment in mathematics education: A systematic review. *ZDM–Mathematics Education*, *57*, 679–693. https://doi.org/10.1007/s11858-025-01696-x

McKenney, S., & Reeves, T. C. (2012). *Conducting educational design research*. Routledge. https://www.routledge.com/Conducting-Educational-Design-Research/McKenney-Reeves/p/book/9780415618038

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods*, *16*(1), 1609406917733847. https://doi.org/10.1177/1609406917733847

Nseibo, K., Vergunst, R., McKenzie, J., Kelly, J., Karisa, A., & Watermeyer, B. (2022). Learning together: The dynamics of training teachers for disability inclusion in South Africa. *Frontiers in Education*, *7*, 999814. https://doi.org/10.3389/feduc.2022.999814

Pastore, S. (2023). Teacher assessment literacy: A systematic review. *Frontiers in Education*, *8*, 1217167. https://doi.org/10.3389/feduc.2023.1217167

Pastore, S., & Mincu, M. (2024). A feasible balance? The Italian teachers' standpoint on assessment literacy, assessment practice, and teacher professional development. *European Journal of Teacher Education*, *47*(2), 226–245. https://doi.org/10.1080/02619768.2024.2338853

Porta, T., Quane, K., & Gilbert, D. (2026). 'Nobody is prioritised; they are all prioritised': Teachers' application and conceptualisation of Universal Design for Learning in inclusive mathematical learning environments. *Learning Environments Research*, *29*, 25. https://doi.org/10.1007/s10984-026-09581-4

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

Ramdani, A. R. (2021, June 3). *Anak berkebutuhan khusus bisa mendaftar di semua sekolah Kota Bandung*. Merahputih.com. https://www.merahputih.com/post/read/anak-berkebutuhan-khusus-bisa-mendaftar-di-semua-sekolah-kota-bandung

Ratri, T. M., & Herawati, N. I. (2024). Implementasi pelaksanaan pendidikan inklusi di sekolah reguler Kota Bandung. *Jurnal Lensa Pendas*, *9*(1), 96–109. https://doi.org/10.33222/jlp.v9i1.3524

Roisatul MN, F., Sulistyani, N., Maulyda, M. A., & Deviana, T. (2024). Pendampingan penyusunan instrumen asesmen diagnostik matematika untuk mendesain pembelajaran berdiferensiasi di sekolah dasar. *SELAPARANG: Jurnal Pengabdian Masyarakat Berkemajuan*, *8*(1), 771–779. https://doi.org/10.31764/jpmb.v8i1.22258

Siburian, D. A. N. B., & Mukhlis, M. (2024). Implementasi asesmen diagnostik dalam Kurikulum Merdeka di SMAN 1 Kabun. *Ideguru: Jurnal Karya Ilmiah Guru*, *9*(3), 1921–1929. https://doi.org/10.51169/ideguru.v9i3.1095

Sims, S., & Fletcher-Wood, H. (2021). Identifying the characteristics of effective teacher professional development: A critical review. *School Effectiveness and School Improvement*, *32*(1), 47–63. https://doi.org/10.1080/09243453.2020.1772841

Umam, A., & Indah, Y. A. (2020). Exploring in-service TEYL teachers' assessment literacy: Implication for continuing professional development. *JEES (Journal of English Educators Society)*, *5*(1). https://doi.org/10.21070/jees.v5i1.364

Wiese, E., & Nortvedt, G. A. (2023). Teacher assessment literacy in culturally and linguistically diverse classrooms: A Norwegian case study. *Teaching and Teacher Education*, *135*, 104357. https://doi.org/10.1016/j.tate.2023.104357

Xu, Y., & Brown, G. T. L. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education*, *58*, 149–162. https://doi.org/10.1016/j.tate.2016.05.010

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia