

Designing Formative Assessment Tools for Primary Moral Education: Teachers' Current Practices and a Practice-Oriented Framework

Hang Nguyen Thi Thu¹ and Le To Quyen²

¹Faculty of Primary Education, Thai Nguyen University of Education, Thai Nguyen University, Thai Nguyen, Vietnam.

²Dan Tien II Primary School, Dan Tien Commune, Thai Nguyen Province, Vietnam.

Abstract: This study investigated primary teachers' current practices, barriers, and professional-development needs in designing and using formative assessment tools for Grade 3 Moral Education in Vietnam and developed a practice-oriented framework for classroom use. A cross-sectional survey was conducted with 120 teachers from primary schools in Thai Nguyen, Tuyen Quang, Phu Tho, and Cao Bang provinces. Data from a five-point questionnaire were analyzed using frequencies, percentages, means, and standard deviations. Teachers strongly endorsed standards-aligned, multi-source formative assessment (overall $M = 4.03$), but their practice remained dominated by oral questioning, situation-based tasks, and observation without structured recording. Only 12.5% reported independently designing complete assessment tools. Analytic rubrics ($M = 1.72$), holistic rubrics ($M = 1.58$), portfolios ($M = 1.61$), peer-assessment forms ($M = 2.09$), and rating scales ($M = 2.05$) were seldom used. Technical design activities, including translating learning outcomes into observable indicators, differentiating performance levels, piloting tools, and specifying interpretation procedures, were also infrequent. Major barriers included limited time, difficulty assessing attitudes and behavior, insufficient exemplars, and linguistic diversity. Drawing on formative-assessment theory and the survey findings, the proposed seven-stage framework connects learning outcomes, evidence models, assessment tasks, fit-for-purpose tools, observable criteria, accessibility, feedback, and iterative revision. The findings indicate that sustainable assessment reform requires collaborative professional learning that develops teachers' assessment-design literacy rather than the distribution of templates alone.

Keywords: Formative assessment; assessment tools; moral education; primary education; rubrics; teacher assessment literacy; Vietnam.

INTRODUCTION

Classroom assessment is formative when evidence of student learning is interpreted and used to improve subsequent teaching and learning decisions rather than merely to certify achievement (Black & Wiliam, 1998, 2009). This function is especially important in primary education because clear criteria, timely feedback, and opportunities to revise work contribute to self-regulation and productive learning habits. Formative assessment therefore depends on the alignment among intended learning outcomes, evidence-generating tasks, interpretive criteria, feedback, and instructional action (National Research Council, 2001; Schildkamp *et al.*, 2020).

Moral education presents distinctive assessment demands. Learning outcomes extend beyond remembering rules or identifying right and wrong; they include recognizing moral issues, expressing attitudes, selecting and explaining responses, acting responsibly, and regulating behavior in meaningful situations. Character-education scholarship similarly conceptualizes moral development as an integration of moral knowing, moral feeling, and moral action (Lickona, 1991). A single written response may reveal knowledge or an intended choice but cannot support broad conclusions about stable attitudes or behavior. Assessment in this domain should therefore focus

on specific performances and combine evidence across tasks, contexts, and time.

Vietnam's 2018 General Education Curriculum positions Moral Education as a compulsory primary-school subject contributing to five core qualities—patriotism, compassion, diligence, honesty, and responsibility—and to competencies related to behavioral regulation, personal development, and participation in social and economic life (Ministry of Education and Training [MOET], 2018a, 2018b). Circular No. 27/2020/TT-BGDDT emphasizes continuous assessment through teacher feedback, student self-assessment, peer participation, and appropriate cooperation with families (MOET, 2020). These requirements create a need for usable tools such as checklists, rating scales, rubrics, observation forms, self-assessment sheets, peer-assessment forms, and portfolios.

An assessment task is not the same as an assessment tool. A scenario, question, role-play, or learning product creates an opportunity for students to demonstrate learning; a tool specifies how the resulting evidence will be recorded, classified, interpreted, and used. Without explicit criteria, observable indicators, differentiated performance descriptions, and interpretation

guidance, teachers' judgments may remain impressionistic. This risk is particularly important in moral education, where global labels such as good, active, or well-behaved can conflate a learner's identity with a specific action.

The challenge is heightened in linguistically and culturally diverse classrooms. Grade 3 students may understand a situation or identify an appropriate response yet have difficulty explaining their reasoning in Vietnamese. When assessment depends heavily on reading and writing, results may partly reflect language proficiency rather than the intended moral-learning outcome. Fair tools should preserve the target standard while allowing appropriate response modes such as oral explanation, pictures, choice cards, role-play, or observed action (CAST, 2024).

Although research has established principles for formative assessment, rubric design, feedback, and self-assessment (Andrade *et al.*, 2008; Brookhart, 2013, 2018; Hattie & Timperley, 2007), fewer studies have examined how primary teachers translate moral-education learning outcomes into structured classroom tools, particularly in schools serving ethnic-minority communities. This study therefore examined teachers' perceptions, current practices, barriers, and professional-development needs and used the findings to develop a practice-oriented framework. Four questions guided the study: (1) How do teachers perceive the purposes and requirements of formative assessment tools? (2) What methods, tools, and design practices do

they report? (3) What barriers and support needs do they identify? and (4) What practical framework can be derived from the theoretical principles and empirical findings?

MATERIALS AND METHODS

Research Design

The study used a cross-sectional descriptive survey design. The principal data source was a structured teacher questionnaire containing fixed-response and open-ended items. The study focused on teachers' perceptions, reported practices, barriers, and professional-development needs related to formative assessment tools in Grade 3 Moral Education.

Participants and Context

Participants were 120 teachers who were currently teaching Grade 3 or had recently taught Grade 3 Moral Education in primary schools in Thai Nguyen, Tuyen Quang, Phu Tho, and Cao Bang provinces. Purposive convenience sampling was used to include teachers from urban, rural, mountainous, socioeconomically disadvantaged, and ethnic-minority contexts. Seventy-eight participants (65.0%) were current Grade 3 teachers and 42 (35.0%) had recent Grade 3 teaching experience. Sixty-nine teachers (57.5%) worked in mountainous or disadvantaged schools, while 51 (42.5%) worked in relatively advantaged locations. Sixty-four teachers (53.3%) taught in a class or school in which ethnic-minority students represented at least half of enrolment.

Table 1: Participant Profile

Participant characteristic	n	%
Currently teaching Grade 3	78	65.0
Previously taught Grade 3	42	35.0
Mountainous/disadvantaged school	69	57.5
Relatively advantaged location	51	42.5
≥50% ethnic-minority students	64	53.3

Instrument and Data Analysis

The questionnaire covered six domains: (a) perceptions of the purposes and requirements of assessment tools; (b) frequency of assessment methods and formats; (c) use of checklists, rating scales, rubrics, observation forms, self-assessment, peer assessment, feedback forms, and portfolios; (d) design and adaptation practices; (e) use of assessment evidence for feedback, instructional adjustment, progress monitoring, and family communication; and (f) barriers and professional-development needs.

Frequency items ranged from 1 (never) to 5 (very frequently); agreement items ranged from 1 (strongly disagree) to 5 (strongly agree); and barrier/need items ranged from 1 (none/not needed) to 5 (very high/very necessary). Quantitative data were summarized using frequencies, percentages, means, and standard deviations. Mean scores were interpreted as very low (1.00–1.80), low (1.81–2.60), moderate (2.61–3.40), high (3.41–4.20), or very high (4.21–5.00). For multiple-response items, percentages were calculated against the full sample; therefore, totals could exceed 100%.

Ethical Considerations

Participation was voluntary. The introductory statement in the anonymous online questionnaire explained the study purpose, confidentiality, and the voluntary nature of participation before teachers decided whether to proceed. The questionnaire was distributed through a Google Forms link and did not request names, email addresses, telephone numbers, school names, or other directly identifying information. Respondents were not required to sign in. Submission of the completed questionnaire constituted informed consent. Data were used only for research purposes, stored securely, and reported in aggregate form.

RESULTS

Teachers' Perceptions of Assessment-Tool Requirements

Teachers expressed positive perceptions of formative assessment tools (overall $M = 4.03$). The strongest endorsements concerned alignment with learning outcomes ($M = 4.28$), focusing comments on behavior rather than labeling students ($M = 4.24$), combining evidence about cognition, attitudes, choices, and behavior ($M = 4.21$), and using results for feedback and instructional adjustment ($M = 4.19$). The lowest item concerned distinguishing checklists, rating scales, and rubrics ($M = 3.42$), indicating that broad support for assessment reform coexisted with uneven technical knowledge.

Table 2: Teachers' Perceptions of Assessment-Tool Requirements

Selected perception item	M	SD
Tools should align with learning outcomes	4.28	0.64
Comments should target behavior, not label students	4.24	0.65
Combine cognition, attitude, choice, and behavior evidence	4.21	0.69
Use results for feedback and instructional adjustment	4.19	0.67
One written test is insufficient	4.16	0.73
Adapt to linguistic and cultural characteristics	4.12	0.72
Distinguish checklists, rating scales, and rubrics	3.42	0.86

Sources of Assessment Tools and Common Assessment Methods

Teachers mainly obtained assessment materials from the Internet (83.3%), colleagues or professional groups (75.0%), and textbooks or teacher guides (70.8%). Only 12.5% reported

independently designing complete tools. This pattern indicates substantial dependence on available materials rather than systematic tool development.

Table 3: Sources of Assessment Tools Used by Teachers

Source of assessment tools	%
Internet	83.3
Colleagues/professional groups	75.0
Textbooks/teacher guides	70.8
Complete independent tool design	12.5

Observation ($M = 4.32$), oral questioning ($M = 4.18$), and written questions or tasks ($M = 4.05$) were frequently used. Practice or role-play

assessment ($M = 2.54$) and portfolios or behavioral journals ($M = 1.92$) were much less common.

Table 4: Frequency of Common Assessment Methods

Assessment method	M	SD
Observation	4.32	0.66
Oral questioning	4.18	0.70
Written questions/tasks	4.05	0.75
Practice/role-play assessment	2.54	0.88
Portfolio/behavior journal	1.92	0.79

Use of Structured Formative Assessment Tools

Situation-based questions or exercises were the most frequently used resource ($M = 4.24$), followed by product or learning-task forms ($M =$

3.56). In contrast, structured tools for recording and interpreting evidence were used infrequently. Self-assessment forms, checklists, observation forms, peer-assessment forms, and rating scales

were all in the low range. Analytic rubrics, portfolios, and holistic rubrics were in the very-low range. The contrast suggests that assessment practice was often task-centered: teachers used

activities that elicited responses but less often employed explicit criteria and procedures for interpreting the evidence.

Table 5: Frequency of Selected Assessment Tools

Assessment tool	M	SD	Level
Situation-based questions/tasks	4.24	0.68	Very high
Product/task assessment form	3.56	0.82	High
Self-assessment form	2.36	0.91	Low
Checklist	2.28	0.88	Low
Teacher observation form	2.13	0.86	Low
Peer-assessment form	2.09	0.84	Low
Rating scale	2.05	0.85	Low
Analytic rubric	1.72	0.76	Very low
Portfolio/behavior journal	1.61	0.72	Very low
Holistic rubric	1.58	0.70	Very low

Assessment-Tool Design and Use of Evidence

Teachers more frequently identified an assessment purpose ($M = 3.54$) and examined learning outcomes ($M = 3.41$) than completed technical design steps. Identifying evidence-generating tasks was moderate ($M = 2.76$). Differentiating evidence of cognition, attitudes, choices, and behavior ($M = 2.42$), translating outcomes into observable indicators ($M = 2.31$), adapting language for students with limited Vietnamese ($M = 2.26$), writing recording and interpretation guidance ($M = 2.14$), and describing performance levels behaviorally ($M = 2.08$) were infrequent. Piloting before wider use was least frequent ($M = 1.62$). Teachers rated their tools higher for alignment ($M = 3.42$) and feasibility ($M = 3.39$) than for limiting subjectivity ($M = 2.46$), inter-rater consistency ($M = 2.41$), and clear differentiation among performance levels ($M = 2.38$).

Immediate feedback was relatively frequent ($M = 4.03$), but sustained use of evidence was less developed. Identifying achieved and improvable elements was moderate ($M = 3.18$), as was adjusting questions, tasks, or teaching methods ($M = 3.05$). Opportunities to revise a task ($M = 2.87$), individualized support ($M = 2.74$), and combining multiple pieces of evidence before judging

behavior ($M = 2.68$) were less frequent. Student self-assessment ($M = 2.44$), progress tracking ($M = 2.31$), and guided peer assessment ($M = 2.18$) were low.

Barriers and Professional-Development Needs

The leading barriers were lack of time ($M = 4.31$), difficulty describing clear quality levels ($M = 4.18$), difficulty assessing attitudes and behavior ($M = 4.16$), insufficient exemplars and guidance ($M = 4.12$), and difficulty converting learning outcomes into observable indicators ($M = 4.08$). Other high-level difficulties included large classes, adapting tools for students with limited Vietnamese, collecting evidence across contexts and time, organizing self- and peer assessment, and storing and using assessment data.

Professional-development needs were uniformly very high. Teachers prioritized topic- or lesson-specific exemplars ($M = 4.72$), guidance on designing checklists, rating scales, and rubrics ($M = 4.64$), guidance for assessing cognition, attitudes, choices, and behavior ($M = 4.61$), collaborative design-pilot-revision activities ($M = 4.55$), and using evidence for feedback and instructional adjustment ($M = 4.48$).

Table 6: Leading Barriers and Professional-Development Needs

Highest-rated barriers	M	Highest-rated support needs	M
Lack of time	4.31	Lesson/topic tool exemplars	4.72
Describing quality levels	4.18	Designing checklists/scales/rubrics	4.64
Assessing attitudes and behavior	4.16	Assessing cognition-attitude-choice-behavior	4.61
Insufficient exemplars/guidance	4.12	Collaborative design, pilot, and revision	4.55
Outcomes to observable indicators	4.08	Feedback and instructional adjustment	4.48

Proposed Practice-Oriented Framework

The findings show a gap between endorsement and enactment. Teachers recognized the value of standards-aligned, multi-source assessment, yet

practice centered on readily available tasks and intuitive judgments. The proposed framework translates formative-assessment theory and the identified needs into seven connected stages.

Table 7: Practice-Oriented Framework for Designing and Using Formative Assessment Tools

Stage	Core teacher actions	Expected contribution
1. Clarify purpose and outcome	Identify the immediate purpose and specify the moral-learning claim.	Focused and aligned assessment.
2. Specify the evidence model	Distinguish evidence of knowledge, attitude, choice/reasoning, and enacted behavior.	Avoids overgeneralization from one response.
3. Design an evidence-generating task	Select a scenario, role-play, product, routine, or authentic activity that elicits the target evidence.	Connects the learning claim with observable performance.
4. Select a fit-for-purpose tool	Match checklists, rating scales, rubrics, observation forms, or self/peer forms to the intended interpretation.	Ensures functional tool selection.
5. Write observable criteria and levels	Describe specific speech, actions, choices, explanations, or products; distinguish levels by quality, independence, accuracy, or consistency.	Improves clarity and consistency.
6. Ensure accessibility and feasibility	Use concise wording, suitable response modes, and culturally familiar contexts without changing the target standard.	Reduces irrelevant language barriers and workload.
7. Use evidence, provide feedback, and revise	Combine evidence where appropriate, specify the next step, adjust instruction, and revise the tool after use.	Completes the formative cycle.

For example, when assessing how students handle disagreement with peers, the teacher first specifies whether the target is recognizing the disagreement, choosing a constructive response, explaining the choice, or demonstrating respectful communication. A scenario can elicit recognition and reasoning; a role-play can elicit listening and respectful language; and a brief self-assessment can support reflection. An analytic rubric may evaluate the scenario response, while a checklist records observable interaction behaviors. Students with limited written Vietnamese may respond orally, use illustrated options, or demonstrate the response through role-play, provided that the same criteria are retained.

Implementation should be supported through collaborative professional learning. Teacher groups can analyze one learning outcome, identify evidence, design a tool, pilot it, examine the resulting student work, and revise the criteria. Schools can maintain a shared tool bank that records the target outcome, evidence type, task, criteria, interpretation guidance, accessibility adaptations, and revision history. In large classes, teachers can rotate observation groups and prioritize a small number of indicators rather than attempting to score every student on every criterion in every lesson.

DISCUSSION

The results reveal a clear gap between teachers' endorsement of formative-assessment principles and their technical enactment in Grade 3 Moral Education. Teachers valued alignment with learning outcomes, multiple evidence sources, behavior-focused feedback, and instructional use, yet practice remained dominated by oral questioning, situation-based tasks, and observation. The low use of rubrics, rating scales, observation forms, self-assessment, and peer assessment indicates that assessment was often task-centered rather than evidence-centered. This finding is consistent with the view that formative assessment depends on teachers' capacity to elicit, interpret, and use evidence, not simply on the availability of activities or templates (National Research Council, 2001; Schildkamp *et al.*, 2020).

The findings also explain why providing ready-made forms alone is unlikely to produce sustained change. Although teachers strongly requested lesson-specific exemplars, they also needed guidance in translating learning outcomes into observable indicators, distinguishing performance levels, piloting tools, and using results for feedback and instructional adjustment. Professional learning should therefore combine

carefully selected examples with collaborative cycles of analysis, design, classroom trial, review of evidence, and revision. Such work develops transferable assessment-design literacy rather than dependence on decontextualized forms.

The limited use of self- and peer assessment deserves attention because formative assessment requires students to become users of criteria and evidence. Rubric-referenced self-assessment can improve primary students' performance when criteria are explicitly taught and applied to concrete work (Andrade *et al.*, 2008; Vasileiadou & Karadimitriou, 2021). For Grade 3 students, these practices should use a small number of behavioral indicators, modeled examples, and guided language for respectful reflection rather than simple self-scoring.

The results also highlight an important fairness issue. When moral understanding is assessed primarily through written Vietnamese, performance may be affected by language proficiency. Alternative modes such as oral responses, visual prompts, choice cards, and role-play can reduce construct-irrelevant barriers when they preserve the same learning outcome and criteria (CAST, 2024). This adjustment does not lower expectations; it improves the validity of the inference made from student performance.

The study has limitations. The purposive convenience sample was drawn from four northern Vietnamese provinces, and self-report data may be affected by social desirability and differences in teachers' understanding of assessment terminology. The descriptive design does not establish causal relationships, and the proposed framework has not yet been evaluated as a professional-development intervention. Future studies should test the framework in diverse school settings and examine changes in tool quality, consistency of teacher judgments, feedback practices, student understanding of criteria, and feasibility over time.

CONCLUSION

Primary teachers demonstrated strong support for formative assessment but limited experience with the structured design and systematic use of assessment tools. Their practice depended heavily on readily available tasks and common methods, while criteria development, performance-level description, piloting, progress tracking, self-assessment, and peer assessment remained uncommon. The proposed seven-stage framework

responds to this gap by connecting learning outcomes, evidence models, tasks, fit-for-purpose tools, observable criteria, accessibility, feedback, and iterative revision. Its central implication is that assessment reform should move beyond distributing templates toward sustained, collaborative development of teachers' assessment-design literacy and evidence-informed classroom practice.

Acknowledgments

The authors sincerely thank the primary school teachers who voluntarily participated in the survey.

Declarations

Ethics approval and consent to participate: Participation was voluntary. The introductory statement in the anonymous online questionnaire explained the purpose of the study, confidentiality, and the voluntary nature of participation. Submission of the completed questionnaire constituted informed consent. No directly identifying information was collected.

Consent for publication: Not applicable.

Competing interests: The authors declare that they have no competing interests.

Funding: This research received no external funding.

Author contributions: Hang Nguyen Thi Thu: conceptualization, survey-instrument design, translation, writing-original draft, and correspondence and manuscript submission. Le To Quyen: survey coordination, development of the theoretical framework, and contribution to manuscript writing. Both authors reviewed and approved the final manuscript.

REFERENCES

1. Andrade, H. L., Du, Y., & Wang, X. "Putting rubrics to the test: The effect of a model, criteria generation, and rubric-referenced self-assessment on elementary school students' writing." *Educational Measurement: Issues and Practice* 27.2 (2008): 3-13.
2. Black, P., & Wiliam, D. "Assessment and classroom learning." *Assessment in Education: principles, policy & practice* 5.1 (1998): 7-74.
3. Black, P., & Wiliam, D. "Developing the theory of formative assessment." *Educational Assessment, Evaluation and Accountability (formerly: Journal of personnel evaluation in education)* 21.1 (2009): 5-31.
4. Brookhart, S. M. "How to create and use rubrics for formative assessment and grading." *Ascd*, (2013).

5. Brookhart, S. M. "Appropriate criteria: Key to effective rubrics." *Frontiers in education*. Vol. 3. Frontiers Media SA, (2018).
6. CAST. "CAST Universal Design for Learning Guidelines version 3.0." (2024).
7. Hattie, J., & Timperley, H. "The power of feedback." *Review of educational research* 77.1 (2007): 81-112.
8. Lickona, T. "Educating for character: How our schools can teach respect and responsibility." *Bantam*, (1992).
9. Ministry of Education and Training. "General education curriculum: Overall curriculum [Issued under Circular No. 32/2018/TT-BGDDT]." *Hanoi, Vietnam*. (2018a).
10. Ministry of Education and Training. "General education curriculum for civic education [Issued under Circular No. 32/2018/TT-BGDDT]." *Hanoi, Vietnam*. (2018b).
11. Ministry of Education and Training. "Circular No. 27/2020/TT-BGDDT promulgating regulations on the assessment of primary school students." *Hanoi, Vietnam*. (2020).
12. National Research Council. "Knowing what students know: The science and design of educational assessment." *National Academies Press*. (2001).
13. Schildkamp, K., Van Der Kleij, F. M., Heitink, M. C., Kippers, W. B., & Veldkamp, B. P. "Formative assessment: A systematic review of critical teacher prerequisites for classroom practice." *International journal of educational research* 103 (2020): 101602.
14. Vasileiadou, D., & Karadimitriou, K. "Examining the impact of self-assessment with the use of rubrics on primary school students' performance." *International Journal of Educational Research Open* 2 (2021): 100031.

Source of support: Nil; **Conflict of interest:** Nil.

Cite this article as:

Thu, H. N. T. & Quyen, L. T. "Designing Formative Assessment Tools for Primary Moral Education: Teachers' Current Practices and a Practice-Oriented Framework." *Sarcouncil Journal of Education and Sociology* 5.8 (2026): pp 1-7.