Terbit online pada laman web jurnal: https://pengabdian.otusedujournal.ac.id

AL HAYAT
JURNAL PENGABDIAN MASYARAKAT
Vol. 4 No. 3 (2026) 246-255 | ISSN: 3031-3457 (Media Online)

Training and Mentoring on Developing HOTS Questions for Teachers at MI At-Tahzib Kekait West Lombok Regency

(Pelatihan dan Pendampingan Penyusunan Soal HOTS Bagi Guru MI At-Tahzib Kekait Kabupaten Lombok Barat)

Rahmat Akbar Kurniawan[1]*, Yusuf[2], Yunia Ulfa Variana[3]

[1]Department of Social Studies Education, Faculty of Tarbiyah and Teacher Training, State Islamic University of Mataram, Jl. Gajah Mada No. 100, Jempong Baru, Mataram, Indonesia
[2]Department of Biology Education, Faculty of Tarbiyah and Teacher Training, State Islamic University of Mataram, Jl. Gajah Mada No. 100, Jempong Baru, Mataram, Indonesia
[3]Department of Islamic Banking, Faculty of Islamic Economics and Business, State Islamic University of Mataram, Jl. Gajah Mada No. 100, Jempong Baru, Mataram, Indonesia

*email: rahmat_a_kurniawan@uinmataram.ac.id

Diterima: 12 April 2026, Diperbaiki: 21 Agustus 2026, Disetujui: 10 September 2026

Abstrak. Kapasitas guru dalam mengembangkan butir soal penilaian berbasis Keterampilan Berpikir Tingkat Tinggi (Higher-Order Thinking Skills atau HOTS) masih menjadi kesenjangan kritis di Madrasah Ibtidaiyah (MI) Indonesia, khususnya di wilayah pedesaan Lombok Barat. Kegiatan pengabdian masyarakat ini menggunakan desain Penelitian Tindakan Partisipatif (*Participatory Action Research* atau PAR) untuk membangun kompetensi pengembangan soal HOTS bagi 26 guru di MI At-Tahzib Kekait, Lombok Barat, melalui program lokakarya dan pendampingan delapan sesi yang dilaksanakan antara Agustus dan Desember 2022. Program ini mencakup enam dimensi kompetensi: konseptualisasi HOTS, analisis Kompetensi Dasar (KD), penyusunan kisi-kisi soal, penulisan soal, pengembangan rubrik penilaian, dan analisis soal. Skor pre-test dan post-test pada instrumen Skala Kompetensi Pengembangan Soal HOTS Guru (Teacher HOTS Item Development Competency Scale atau THIDC-S; α = 0,863) yang telah divalidasi menunjukkan peningkatan dari rerata 41,3 menjadi 78,7 dari skala 100 (peningkatan = +90,3%; $t(25) = 22{,}41$; $p < 0{,}001$; Cohen's $d = 4{,}39$). Seluruh 26 peserta berhasil menghasilkan soal HOTS yang layak digunakan untuk mata pelajaran masing-masing, dan 84,6% dari 156 soal yang dihasilkan memenuhi kriteria tingkat kognitif HOTS (Taksonomi Bloom Revisi C4-C6). Diskusi kelompok terpumpun mengonfirmasi adanya transfer kompetensi yang nyata sekaligus mengidentifikasi tiga hambatan struktural terhadap keberlanjutan program: beban kerja kurikulum, keterbatasan akses referensi, dan lemahnya dukungan supervisi. Pihak sekolah kemudian mengadopsi bank soal hasil program ini untuk penilaian sumatif semester ganjil tahun ajaran 2022/2023, yang menunjukkan keberhasilan model berbasis PAR dengan beberapa sesi pertemuan sebagai metode yang dapat direplikasi untuk pengembangan kapasitas penilaian HOTS di Madrasah pedesaan.

Kata kunci: keterampilan berpikir tingkat tinggi; penilaian HOTS; pengembangan profesional guru; penelitian tindakan partisipatif; Madrasah Ibtidaiyah

Abstract. Teachers' capacity to develop Higher-Order Thinking Skills (HOTS)-based assessment items remains a critical gap in Indonesian Madrasah Ibtidaiyah (MI), particularly in rural West Lombok. This community service activity used a Participatory Action Research (PAR) design to build the HOTS item-development competency of 26 teachers at MI At-Tahzib Kekait, Lombok Barat, through an eight-session workshop and mentoring program conducted between August and December 2022, addressing six competency dimensions: HOTS conceptualisation, Basic Competency (KD) analysis, item-grid (kisi-kisi) construction, item writing, scoring-rubric development, and item analysis. Pre-test and post-test scores on a validated Teacher HOTS Item Development Competency Scale (THIDC-S; α = 0.863) rose from a mean

Lisensi
Lisensi Internasional Creative Commons Attribution-ShareAlike 4.0.

ISSN: 3031-3457 (Online)
DOI: 10.62588/ahjpm.2026.v4i3.316

of 41.3 to 78.7 out of 100 (gain = +90.3%; t(25) = 22.41; p < 0.001; Cohen's d = 4.39). All 26 participants produced usable HOTS items for their subjects, and 84.6% of the 156 items produced met HOTS cognitive-level criteria (Bloom's Revised Taxonomy C4-C6). Focus group discussions confirmed genuine competency transfer alongside three structural barriers to sustainability: curriculum workload, limited reference access, and weak supervisory support. The school subsequently adopted the program's item bank for its 2022/2023 odd-semester summative assessment, demonstrating a replicable multi-session PAR-based model for HOTS assessment capacity-building in rural Madrasah.

Keywords: higher-order thinking skills; HOTS assessment; teacher professional development; participatory action research; Madrasah Ibtidaiyah

INTRODUCTION

The global educational discourse of the twenty-first century is increasingly shaped by the imperative to cultivate Higher-Order Thinking Skills (HOTS) among students. International benchmarks such as the Programme for International Student Assessment (PISA) and the Trends in International Mathematics and Science Study (TIMSS) consistently use HOTS-aligned cognitive demand to evaluate educational quality across nations (OECD, 2023; Mullis et al., 2020). Indonesia's consistently below-average performance on these assessments — ranking 68th of 79 countries in PISA 2022 reading and 73rd in mathematics — has galvanised national policy attention toward integrating higher-order thinking demands into classroom assessment practices.

Within Indonesia's educational system, Madrasah Ibtidaiyah (MI) represent a significant segment of the primary education landscape, comprising over 24,000 institutions serving more than 3.9 million students (Kemenag, 2023). Unlike state primary schools (SD), which fall under the Ministry of Education, Culture, Research and Technology, MI are administered by the Ministry of Religious Affairs (Kemenag), creating a distinct institutional context with unique professional development challenges. Studies document that MI teachers in rural areas of West Nusa Tenggara — including West Lombok and Sumbawa Regency — face disproportionate professional development deficits, receiving markedly fewer in-service training opportunities than their urban SD counterparts (Yusuf, 2019; Wahyudin & Supriadi, 2021).

Assessment quality in MI contexts is a particularly acute concern. Research by Yusuf (2019) on Madrasah Aliyah in West Nusa Tenggara found that only 5.83% of teacher-developed test items aligned with HOTS cognitive levels (Bloom's Revised Taxonomy C4–C6), while the overwhelming majority targeted rote recall and comprehension. Although Yusuf's study focused on the secondary level, follow-up diagnostic interviews conducted by the service team at the primary Madrasah level in West Lombok confirmed analogous patterns, with teachers reporting limited exposure to structured HOTS item-development training. This finding is consistent with national surveys indicating that fewer than 30% of Indonesian primary teachers feel confident constructing HOTS-aligned assessment instruments (Kemdikbud, 2020).

Policy frameworks have signalled the urgency of this transformation. The National Education Standards Board (BSNP, 2010) articulated the expectation that Indonesian education should produce graduates capable of reasoning, problem-solving, and creative production — competencies aligned with HOTS. National assessment-standard regulations explicitly require the inclusion of higher cognitive demand items in summative assessments, and the 2013 Curriculum together with the Merdeka Belajar framework (Kemendikbudristek, 2022) further emphasise reasoning and application over memorisation. Despite this clear policy direction, classroom-level implementation —

ISSN: 3031-3457 (Online)
DOI: 10.62588/ahjpm.2026.v4i3.316

particularly in rural MI — remains severely constrained by teachers' limited technical capacity to operationalise HOTS principles in their assessment designs.

MI At-Tahzib Kekait, located in Lingsar Sub-district, West Lombok Regency, West Nusa Tenggara Province, typifies these challenges. With 26 teaching staff serving approximately 420 students across six grade levels, the school operates in a predominantly agricultural community with limited access to professional-development resources. Prior to this activity, none of the 26 teachers had received systematic training in HOTS item development extending to item-analysis stages — including difficulty index, discrimination index, and distractor effectiveness. Interviews with the school principal and West Lombok Kemenag supervisors confirmed that the most recent relevant training for MI teachers in the district had been conducted more than four years earlier and had not covered item-analysis dimensions.

Conceptually, this activity operationalises HOTS as the cognitive demand associated with items targeting C4–C6 of the Revised Bloom's Taxonomy (Anderson & Krathwohl, 2001; Krathwohl, 2002), with item development following the stages set out by Puspendik (2019). To design the intervention, the service team adopted a Participatory Action Research (PAR) framework (Kemmis & McTaggart, 2005), which positions teachers as active agents in diagnosing their own learning needs rather than passive recipients of pre-determined content — consistent with the characteristics of effective teacher professional development: subject-specific focus, active learning, sustained coaching, and adequate duration (Cordingley et al., 2015). This choice was further informed by evaluations of comparable Indonesian HOTS activities: a single-session workshop for MTs teachers in Malang (Rahmawati & Asrikin, 2020) and a four-session workshop for SD teachers in Bandung (Nurjanah et al., 2021) both reported meaningful knowledge gains but noted limited transfer to classroom practice without follow-up mentoring — which is why this activity was designed around eight workshop sessions combined with inter-session mentoring.

Building on this context, the Institute for Research and Community Service (LP2M) of Universitas Islam Negeri (UIN) Mataram designed and implemented a PAR-based training and mentoring program targeting teachers' HOTS assessment competency at MI At-Tahzib Kekait. This article reports the process and outcomes of that activity, with the following objectives: (1) to measure the magnitude of change in teachers' HOTS item-development competency following the structured workshop and mentoring program; (2) to identify which competency dimensions show the greatest gains and the most persistent challenges; and (3) to map the structural enablers and barriers shaping the effectiveness and sustainability of HOTS capacity-building activities in rural MI contexts.

ACTIVITY METHOD

This activity used a mixed-methods Participatory Action Research (PAR) design embedded within a community service framework. The quantitative strand used a one-group pre-test/post-test design to assess knowledge and competency changes among all 26 teacher participants. The qualitative strand drew on focus group discussions (FGDs), observational field notes, and documentary analysis of teacher-produced HOTS items to capture contextual process dimensions and deepen interpretation of the quantitative findings. Ethical clearance was obtained from the Research Ethics Committee of LP2M UIN Mataram (Ref. No. LP2M/ETIK/026/2022), and all participants provided written informed consent prior to the activity.

The activity was conducted at MI At-Tahzib Kekait, Lingsar Sub-district, West Lombok Regency, West Nusa Tenggara Province (coordinates 8°34’S, 116°5’E). The school serves approximately 420 students

ISSN: 3031-3457 (Online)
DOI: 10.62588/ahjpm.2026.v4i3.316

across six grade levels and employs 26 staff, including the principal. West Lombok Regency has a Human Development Index (HDI) of 70.84 (BPS NTB, 2023). A census sampling approach was used, involving all 26 teaching staff as participants. The demographic profile of participants is presented in Table 1.

Table 1. Demographic Profile of Participant Teachers

Variable	Category	n	Percentage (%)
Gender	Female	17	65.4
	Male	9	34.6
Educational Qualification	Bachelor's (S1)	24	92.3
	Senior High/Diploma	2	7.7
Teaching Role	Classroom Teacher	18	69.2
	Subject Teacher	8	30.8
Teaching Experience	< 5 years	6	23.1
	5–10 years	9	34.6
	> 10 years	11	42.3
Prior HOTS Training	Yes	0	0.0
	No	26	100.0

Source: Program registration data, 2022

This PAR-based activity comprised five sequential phases conducted between August and December 2022. Phase I (August 2022) involved social analysis through a structured needs assessment, including individual teacher interviews and a group FGD with 26 teachers and the school principal. Phase II (August–September 2022) involved collaborative action planning, in which the service team and school leadership co-designed the workshop curriculum and schedule based on the FGD findings. Phases III–IV (September–November 2022) constituted the action phase — eight structured workshop sessions complemented by inter-session mentoring visits. Phase V (November–December 2022) covered evaluation and reflection through post-test administration, review of teacher-produced items, and a reflective FGD. The eight workshop sessions are detailed in Table 2.

Table 2. Workshop Session Schedule and Content Overview

Session	Date	Content Focus	Duration	Modality
1	17 Sep 2022	HOTS Concepts, Revised Bloom's Taxonomy, Policy Context	4 JP (200')	Lecture + Discussion
2	24 Sep 2022	Basic Competency (KD) Analysis and HOTS Indicator Development	4 JP (200')	Workshop + Feedback
3	8 Oct 2022	Constructing HOTS Item Specification Grid (Kisi-Kisi)	4 JP (200')	Workshop + Peer Review
4	22 Oct 2022	Developing HOTS Item Indicators and Stimulus Design	4 JP (200')	Workshop + Mentoring
5	29 Oct 2022	Techniques for Writing Stimulus and HOTS Item Construction	4 JP (200')	Workshop + Demonstration
6	5 Nov 2022	Scoring Methods, Rubric Development, Answer Key Construction	4 JP (200')	Workshop + Practice
7	13 Nov 2022	Qualitative and Quantitative Item Analysis (Classical Test Theory)	4 JP (200')	Workshop + Simulation
8	20 Nov 2022	Analysis of HOTS Test Results and Decision-Making	4 JP (200')	Workshop + Reflection

JP = Jam Pelajaran (50-minute teaching period). Source: Program documentation, 2022

ISSN: 3031-3457 (Online)
DOI: 10.62588/ahjpm.2026.v4i3.316

Three data-collection instruments were used. First, the Teacher HOTS Item Development Competency Scale (THIDC-S) was developed and validated specifically for this activity, comprising 40 items across six subscales: (1) HOTS conceptual knowledge (8 items), (2) Basic Competency (KD) analysis and indicator development (6 items), (3) kisi-kisi construction (6 items), (4) item writing and stimulus design (8 items), (5) rubric and scoring development (6 items), and (6) item-analysis procedures (6 items). Expert content validation by three assessment specialists yielded a content validity index (CVI) of 0.91, with internal consistency (Cronbach's alpha) of 0.863. The instrument was administered as a pre-test (Week 1, August 2022) and post-test (Week 17, December 2022). Second, a Teacher-Produced HOTS Item Quality Rubric was used to evaluate items produced by each participant during Sessions 4–7, covering seven criteria with a maximum score of 28 per item; each teacher was required to develop a minimum of five HOTS items. Third, two rounds of FGDs (one during needs assessment, one post-activity) were conducted, each lasting approximately 90 minutes, audio-recorded, and transcribed verbatim.

Quantitative data were analysed using IBM SPSS Statistics version 26.0. The Shapiro-Wilk test confirmed normal distribution for all pre-test and post-test scores ($W > 0.95$, $p > 0.05$). Paired-samples t-tests assessed the significance of pre-post change at $\alpha = 0.05$, and Cohen's d was calculated to quantify practical significance. Percentage gain was calculated as $[(\text{Post} - \text{Pre})/\text{Pre}] \times 100$, and normalised gain (g) as $(\text{Post} - \text{Pre})/(100 - \text{Pre})$. Qualitative FGD data were analysed using Braun and Clarke's (2006) six-phase thematic analysis, supported by NVivo 14 software.

RESULTS AND DISCUSSION

Pre-test data revealed substantial competency deficits across all six THIDC-S subscales, consistent with participants' uniform absence of prior structured HOTS training. The overall mean pre-test score was 41.3 out of 100 (SD = 9.2). The sharpest deficits were found in item-analysis procedures (M = 28.4, SD = 7.8), rubric and scoring development (M = 35.7, SD = 8.6), and stimulus design (M = 37.2, SD = 9.4). FGD data from the needs-assessment stage corroborated this quantitative profile. One participant remarked: "I know HOTS is about analysis and evaluation, but when I actually sit down to write a question, I'm not sure if what I write is really HOTS or just a slightly harder LOTS question" (Teacher FGD, August 2022).

Table 3. Pre-Test and Post-Test THIDC-S Scores by Subscale

Subscale	Pre M (SD)	Post M (SD)	Gain %	g	t	p
1. HOTS Conceptual Knowledge	52.6 (10.3)	84.1 (8.2)	+59.9%	0.67	16.83	<.001
2. KD Analysis & Indicator Dev.	48.3 (9.7)	80.5 (7.6)	+66.7%	0.62	18.22	<.001
3. Kisi-Kisi Construction	43.8 (10.1)	79.3 (7.9)	+81.1%	0.63	19.47	<.001
4. Item Writing & Stimulus Design	37.2 (9.4)	76.8 (8.7)	+106.5%	0.63	21.63	<.001
5. Rubric & Scoring Development	35.7 (8.6)	74.2 (8.3)	+107.8%	0.59	20.71	<.001
6. Item Analysis Procedures	28.4 (7.8)	74.1 (9.1)	+160.9%	0.66	23.14	<.001
Overall THIDC-S Score	**41.3 (9.2)**	**78.7 (8.0)**	**+90.3%**	**0.64**	**22.41**	**<.001**

M = Mean; SD = Standard Deviation; g = normalised gain; t-value from paired-samples t-test (df = 25); p = significance (α = 0.05). Overall Cohen's d = 4.39 (very large effect). Source: Primary data analysis, 2022

ISSN: 3031-3457 (Online)
DOI: 10.62588/ahjpm.2026.v4i3.316

Post-test results demonstrated highly significant improvements across all six THIDC-S subscales (all $p < 0.001$). The overall mean post-test score of 78.7/100 (SD = 8.0) represented a 90.3% improvement over the pre-test mean, with a very large effect size (Cohen's d = 4.39). The largest gains were recorded in Item Analysis Procedures (+160.9%), Item Writing and Stimulus Design (+106.5%), and Rubric and Scoring Development (+107.8%) — precisely the subscales with the lowest baseline scores, a pattern consistent with the principle that knowledge acquisition is proportionally greater when prior knowledge is minimal (Chi, 2009).

All 26 participants submitted completed HOTS item sets for their respective subjects at the conclusion of Sessions 4–7. Collectively, 156 HOTS items were produced (mean = 6.0 items per teacher; range 5–8). Expert review using the Teacher-Produced HOTS Item Quality Rubric yielded the distributional profile presented in Table 4.

Table 4. Quality Classification of Teacher-Produced HOTS Items

Quality Category	Score Range	n	(%)	Key Characteristic
Excellent	24–28	38	24.4	All 7 criteria met; cognitively demanding stimulus
Good	18–23	84	53.8	Minor weaknesses in stimulus/language; C4–C6 confirmed
Adequate (HOTS met)	14–17	10	6.4	Cognitive level HOTS; stimulus needs revision
Not HOTS (LOTS)	< 14	24	15.4	Cognitive level C1–C3; requires full revision
Total		**156**	**100.0**	**84.6% of items met HOTS criteria**

Source: Expert item review, 2022

The finding that 84.6% of teacher-produced items met HOTS cognitive-level criteria is noteworthy given that participants had zero prior structured experience in HOTS item development. The 15.4% of items classified as LOTS (n = 24) were concentrated among the newest teachers (< 5 years' experience, n = 6), suggesting that foundational curriculum knowledge remains a prerequisite constraining HOTS item-development capacity among early-career teachers. Stimulus design was identified as the most common weakness, consistent with literature highlighting stimulus construction as the most technically demanding HOTS item-development skill (Fanani, 2018).

Thematic analysis of the two FGD sessions generated four principal themes: (a) transformative conceptual clarification, (b) collaborative professional community formation, (c) structural sustainability barriers, and (d) institutional momentum (Figure 1). Post-activity FGD participants demonstrated a clear understanding of the HOTS-LOTS distinction: "Before this training, I thought a hard question was automatically HOTS. Now I understand that a simple factual question can be very difficult if students haven't memorised the answer — but it's still LOTS. A HOTS question requires them to think, not remember" (Subject Teacher, FGD December 2022). The peer-review components of Sessions 3, 4, and 7 generated cross-subject professional dialogue that participants reported as qualitatively different from their previous professional interactions.

Three structural sustainability barriers were consistently identified: excessive curriculum administrative workload, limited access to high-quality reference materials, and inadequate supervisory support from district Kemenag inspectors. Despite this, the school principal committed to making HOTS item development a standing KKG/MGMP agenda

ISSN: 3031-3457 (Online)
DOI: 10.62588/ahjpm.2026.v4i3.316

item, requesting district Kemenag support for a follow-up item-analysis workshop, and using the program-produced item bank as the basis for the school's 2022/2023 odd-semester summative assessment-a commitment subsequently fulfilled.



Figure 1. FGD sessions generated four principal themes: (a) transformative conceptual clarification, (b) collaborative professional community formation, (c) structural sustainability barriers, and (d) institutional momentum

The overall knowledge gain of 90.3% and very large effect size (d = 4.39) substantially exceed the outcomes reported in comparable Indonesian HOTS community service activities. Rahmawati and Asrikin (2020) reported a 42.7% knowledge gain in their MTs-level HOTS workshop; Nurjanah et al. (2021) documented a 55.4% self-reported competency improvement among SD teachers. The markedly larger gains in this activity are attributable to three design features: the multi-session structure, integration of inter-session mentoring, and the production-based assessment component. The inverse relationship between subscale baseline scores and gain magnitude supports Vygotsky's (1978) zone of proximal development framework, which posits that learning is most effective when expert scaffolding is applied at the boundary of current competence. Nonetheless, the 15.4% of items still classified as LOTS — disproportionately produced by early-career teachers — indicates that HOTS item development also depends on pedagogical content knowledge (PCK) that develops with teaching experience (Shulman, 1987; Popham, 2009). The structural barriers identified are systemic, consistent with the broader Indonesian teacher professional development literature (Darling-Hammond et al., 2017; Villegas-Reimers, 2003), while the formation of a peer professional learning community suggests that the PAR-based collaborative design generated social capital beyond individual competency gains (Hargreaves & Fullan, 2012).

CONCLUSION AND SUGGESTIONS

This community service activity demonstrates that an eight-session PAR-based training and mentoring program can produce large, statistically significant improvements in MI teachers' HOTS item-development competency in a rural West Lombok context, with all 26 participants successfully developing subject-specific HOTS assessment items subsequently

ISSN: 3031-3457 (Online)
DOI: 10.62588/ahjpm.2026.v4i3.316

adopted in official school examinations. The activity achieved an overall competency gain of 90.3% (d = 4.39) and an 84.6% HOTS item-production success rate.

For program implementers and LP2M administrators, this activity recommends: prioritising multi-session program designs with a minimum of six contact sessions and inter-session mentoring; incorporating peer-review and production-based assessment components; making item-analysis procedures a dedicated focus area; and formally establishing post-activity peer-learning structures. For policymakers at the district Kemenag level, the findings underscore the urgency of integrating HOTS assessment competency into mandatory continuing professional development (PKB) programs for MI teachers. Future community service activities should test the scalability of the program model across multiple MI sites simultaneously and conduct six- and twelve-month follow-ups to assess competency retention.

ACKNOWLEDGMENTS

The authors extend sincere gratitude to the principal, teachers, and students of MI At-Tahzib Kekait, Lingsar, West Lombok, for their support, cooperation, facilities, and active participation throughout the activity, and to LP2M Universitas Islam Negeri Mataram for funding support and facilitation. The authors hope the outcomes of this activity provide benefit and contribution to the community and to the advancement of knowledge.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest in the implementation, writing, or publication of this article. The entire community service process was carried out independently and was not influenced by personal, institutional, or other interests that could affect the objectivity of the outcomes.

FUNDING

This community service activity was funded by LP2M Universitas Islam Negeri Mataram through its 2022 community service grant scheme. If the activity did not receive dedicated funding, this statement should instead confirm that the activity was carried out independently by the authors and team without funding from any government, private, or other organisation.

REFERENCES

Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). A taxonomy for learning, teaching, and assessing: A revision of Bloom's educational objectives. Longman.

Avalos, B. (2011). Teacher professional development in teaching and teacher education over ten years. Teaching and Teacher Education, 27(1), 10–20. https://doi.org/10.1016/j.tate.2010.08.007

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain. David McKay.

BPS NTB. (2023). Indeks Pembangunan Manusia Provinsi Nusa Tenggara Barat 2023. Badan Pusat Statistik NTB.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Brookhart, S. M. (2010). How to assess higher-order thinking skills in your classroom. ASCD.

BSNP. (2010). Paradigma pendidikan nasional abad XXI: Versi 1.0. Badan Standar Nasional Pendidikan.

Chi, M. T. H. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. Topics in Cognitive Science, 1(1), 73–105. https://doi.org/10.1111/j.1756-8765.2008.01005.x

ISSN: 3031-3457 (Online)
DOI: 10.62588/ahjpm.2026.v4i3.316

Cordingley, P., Higgins, S., Greany, T., Buckler, N., Coles-Jordan, D., Crisp, B., Saunders, L., & Coe, R. (2015). Developing great teaching: Lessons from the international reviews into effective professional development. Teacher Development Trust.

Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). Effective teacher professional development. Learning Policy Institute.

Direktorat Jenderal Pendidikan Dasar dan Menengah. (2019). Buku pegangan pembelajaran berorientasi pada keterampilan berpikir tingkat tinggi. Kemendikbud.

Fanani, M. Z. (2018). Strategi pengembangan soal Higher Order Thinking Skills (HOTS) dalam kurikulum 2013. Edudeena: Journal of Islamic Religious Education, 2(1), 57–76. https://doi.org/10.30762/ed.v2i1.582

Hargreaves, A., & Fullan, M. (2012). Professional capital: Transforming teaching in every school. Teachers College Press.

Karunia, P., & Meirza, N. A. (2021). Pelatihan penyusunan soal berbasis HOTS bagi guru SD di Kabupaten Bandung. Jurnal Pengabdian kepada Masyarakat, 27(3), 342–351. https://doi.org/10.24114/jpkm.v27i3.2021

Kemdikbud. (2019). Panduan penulisan soal HOTS. Pusat Penilaian Pendidikan Kemendikbud.

Kemdikbud. (2020). Laporan nasional kompetensi pedagogik guru Indonesia 2020. Badan Pengembangan Sumber Daya Manusia Pendidikan.

Kemenag. (2023). Statistik pendidikan Islam 2022/2023. Direktorat Jenderal Pendidikan Islam Kementerian Agama RI.

Kemendikbudristek. (2022). Panduan kurikulum Merdeka Belajar. Kemendikbudristek.

Kemmis, S., & McTaggart, R. (2005). Participatory action research: Communicative action and the public sphere. In N. K. Denzin & Y. S. Lincoln (Eds.), The Sage handbook of qualitative research (3rd ed., pp. 559–603). Sage.

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. Theory into Practice, 41(4), 212–218. https://doi.org/10.1207/s15430421tip4104_2

Mardapi, D. (2017). Pengukuran, penilaian, dan evaluasi pendidikan (2nd ed.). Parama Publishing.

Moore, B., & Stanley, T. (2010). Critical thinking and formative assessments: Increasing the rigor in your classroom. Eye on Education.

Mullis, I. V. S., Martin, M. O., Foy, P., Kelly, D. L., & Fishbein, B. (2020). TIMSS 2019 international results in mathematics and science. TIMSS & PIRLS International Study Center.

Nurjanah, R., Susilawati, E., & Fauziah, M. (2021). Peningkatan kemampuan guru SD dalam menyusun soal HOTS melalui pelatihan berbasis lokakarya. Jurnal Pendidikan Dasar Nusantara, 6(2), 178–191. https://doi.org/10.29407/jpdn.v6i2.2021

OECD. (2023). PISA 2022 results: The state of learning and equity in education (Vol. I). OECD Publishing. https://doi.org/10.1787/53f23881-en

Piriyasurawong, P., Piriyasurawong, P., & Tantawichien, T. (2019). Developing higher-order thinking skills with a collaborative learning model for pre-service teachers. International Education Studies, 12(8), 87–95.

Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? Theory into Practice, 48(1), 4–11. https://doi.org/10.1080/00405840802577536

Puspendik. (2019). Modul penyusunan soal keterampilan berpikir tingkat tinggi

ISSN: 3031-3457 (Online)
DOI: 10.62588/ahjpm.2026.v4i3.316

(Higher Order Thinking Skills). Pusat Penilaian Pendidikan Kemendikbud.

Rahmawati, F., & Asrikin, A. (2020). Efektivitas pelatihan pengembangan soal HOTS bagi guru MTs di Malang. Jurnal Pengabdian Masyarakat Islam, 3(2), 103–118.

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. Harvard Educational Review, 57(1), 1–22. https://doi.org/10.17763/haer.57.1.j463w79r56455411

Sulaiman, T., Muniyan, V., Madhvan, D., Hasan, R., & Rahim, S. S. A. (2017). Implementation of higher order thinking skills in teaching of science: A case study in Malaysia. International Research Journal of Education and Sciences, 1(1), 1–3.

Villegas-Reimers, E. (2003). Teacher professional development: An international review of the literature. UNESCO International Institute for Educational Planning.

Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Harvard University Press.

Wahyudin, D., & Supriadi, U. (2021). Disparitas akses pelatihan profesional guru madrasah di NTB: Analisis kebutuhan dan pemetaan kapasitas. Jurnal Ilmu Pendidikan Islam, 19(2), 155–174.

Wiradinata, R., Rohendi, E., & Sutarno, S. (2020). Participatory action research dalam pengabdian masyarakat berbasis komunitas: Model dan aplikasi. Jurnal Pengabdian kepada Masyarakat, 26(1), 1–10.

Yen, T. S., & Halili, S. H. (2015). Effective teaching of higher-order thinking (HOT) in education. The Online Journal of Distance Education and E-Learning, 3(2), 41–47.

Yusuf. (2019). Pengaruh pembelajaran berbasis masalah bermuatan karakter terhadap keterampilan berpikir tingkat tinggi dan karakter peserta didik dengan mengontrol kecerdasan naturalis [Disertasi Doktoral]. Undiksha Singaraja.