

P-ISSN: 2338-8617

E-ISSN: 2443-2067

Jurnal Ilmiah

PEURADEUN

Vol. 13, No. 3, September 2025



JIP
The Indonesian Journal of the Social Sciences
www.journal.scadIndependent.org
DOI Prefix Number: 10.26811

INDEX  COPERNICUS
INTERNATIONAL

 **sinta** **1**
Science and Technology Index

Accredited "Sinta 1" by Decree No. 72/E/KPT/2024
Valid Until the May 2027 Edition

 **Scopus**[®]
ELSEVIER

 **Clarivate**
Analytics
WEB OF SCIENCE[™]

**Assessment of Teachers' Ability to Develop Hots-Based Test Items:
A GRM Analysis**

**Lian Gafar Otaya¹; Kasim Yahiji²; Rasimin Rasimin³;
Rahmawati Rahmawati⁴; Nurul Imtihan⁵; Norhisham bin Muhamad⁶**

^{1,2,4}*Institut Agama Islam Negeri Sultan Amai Gorontalo, Indonesia*

³*Universitas Islam Negeri Salatiga, Indonesia*

⁵*Universitas Islam Negeri Mataram, Indonesia*

⁶*University of Malaya, Malaysia*

Article in Jurnal Ilmiah Peuradeun

Available at : <https://journal.scadindependent.org/index.php/jipeuradeun/article/view/1555>

DOI : <https://doi.org/10.26811/peuradeun.v13i3.1555>

How to Cite this Article

APA : Otaya, L.G., Yahiji, K., Rasimin, R., Rahmawati, R., Imtihan, N., & Muhamad, N.B. (2025). Assessment of Teachers' Ability to Develop Hots-Based Test Items: A Grm Analysis. *Jurnal Ilmiah Peuradeun*, 13(3), 1995-2018. <https://doi.org/10.26811/peuradeun.v13i3.1555>

Others Visit : <https://journal.scadindependent.org/index.php/jipeuradeun>

Jurnal Ilmiah Peuradeun (JIP), *the Indonesian Journal of the Social Sciences*, is a leading peer-reviewed and open-access journal, which publishes scholarly works, and specializes in the Social Sciences that emphasize contemporary Asian issues with interdisciplinary and multidisciplinary approaches. JIP is published by SCAD Independent and published 3 times a year (January, May, and September) with p-ISSN: 2338-8617 and e-ISSN: 2443-2067. JIP has become a CrossRef member. Therefore, all articles published will have a unique DOI number. JIP has been accredited Rank 1 (Sinta 1) by the Ministry of Education, Culture, Research, and Technology, the Republic of Indonesia, through the Decree of the Director-General of Higher Education, Research, and Technology No. 72/E/KPT/2024, dated April 1, 2024. This accreditation is valid until the May 2027 edition.

All articles published in this journal are protected by copyright, licensed under a Creative Commons 4.0 International License (CC-BY-SA) or an equivalent license as the optimal license for the publication, distribution, use, and reuse of scholarly works.

JIP indexed/included in Web of Science, Scopus, Sinta, MAS, Index Copernicus International, Erih Plus, Garuda, Moreref, Scilit, Sherpa/Romeo, Google Scholar, OAJL, PKP, Index, Crossref, BASE, ROAD, GIF, Advanced Science Index, JournalTOCs, ISI, SIS, ESJI, SSRN, ResearchGate, Mendeley and [others](#).



ASSESSMENT OF TEACHERS' ABILITY TO DEVELOP HOTS-BASED TEST ITEMS: A GRM ANALYSIS

Lian Gafar Otaya¹; Kasim Yahiji²; Rasimin Rasimin³;
Rahmawati Rahmawati⁴; Nurul Imtihan⁵; Norhisham bin Muhamad⁶

^{1,2,4}Institut Agama Islam Negeri Sultan Amai Gorontalo, Indonesia

³Universitas Islam Negeri Salatiga, Indonesia

⁵Universitas Islam Negeri Mataram, Indonesia

⁶University of Malaya, Malaysia

¹Correspondence Email: lianotaya82@iaingorontalo.ac.id

Received: May 27, 2024	Accepted: September 5, 2025	Published: September 30, 2025
Article Url: https://journal.scadindependent.org/index.php/jipeuradeun/article/view/1555		

Abstract

Developing Higher-Order Thinking Skills (HOTS)-based test items is a critical competency for teachers in 21st-century education. However, its implementation remains limited, particularly among Islamic education teachers. This study explores teachers' ability to participate in the Professional Teacher Education Program under the Indonesian Ministry of Religious Affairs at the State Institute for Islamic Studies (IAIN) Sultan Amai Gorontalo to construct HOTS-based assessments. Utilizing an exploratory, cross-sectional design, the study involved 495 madrasah and Islamic education teachers across Indonesia. Data were collected through document analysis of teacher-developed test items and classroom observations, then analyzed using the Graded Response Model (GRM) in the R program. The findings reveal that most participants demonstrated moderate ability levels in developing HOTS-based test items, with ability scores ranging between -1.00 and 1.00. The study highlights significant challenges in item construction, particularly in aligning cognitive levels with stimulus and question indicators. These results emphasize the need for improved pedagogical training and targeted support to strengthen teachers' competence in designing cognitively demanding assessments. The study contributes to the global discourse on HOTS-based assessment in religious education and offers practical implications for curriculum design, teacher training, and assessment policy in culturally rooted educational contexts.

Keywords: HOTS; Assessment Design; Teacher Competence; Graded Response Model; Islamic Education.



A. Introduction

In the evolving landscape of 21st-century education, the integration of Higher-Order Thinking Skills (HOTS) encompassing analysis, evaluation, and creation has become a cornerstone for preparing learners to navigate increasingly complex and unpredictable global challenges. A growing body of scholarship underscores that the development of these cognitive skills must be embedded not only in instructional strategies but also in assessment practices, as assessments profoundly shape classroom teaching and learning priorities (Albab et al., 2020; Osborne et al., 2016; Solissa et al., 2023; Tanujaya et al., 2017). Central to this effort is teachers' capacity to design HOTS-oriented assessments that move beyond rote memorization and factual recall toward fostering critical, analytical, and creative thinking.

Yet, in many contexts, including Indonesia, this ideal remains aspirational. In the Ministry of Religious Affairs' Professional Teacher Education program, hereinafter referred to as Professional Teacher Program or PPG, HOTS is promoted as a strategic objective for pedagogical reform. However, empirical evidence consistently reveals a persistent gap between policy prescriptions and classroom realities. Research consistently shows that teachers across educational levels and disciplines tend to design assessment items that focus on lower-order cognitive domains, particularly C1 (remember) and C2 (understand), while questions aligned with higher-order thinking skills (HOTS) at C3 (apply), C4 (analyze), C5 (evaluate), and C6 (create) remain limited in both frequency and quality (Hakim et al., 2021; Nugraha, 2020; Tibrani et al., 2023; Winarti et al., 2021). This predominance of lower-order questioning not only limits students' opportunities to engage in higher cognitive processing but also constrains the development of transferable problem-solving skills essential for life and work in the 21st century.

While prior studies have documented this overreliance on recall-based assessment, the majority have been situated in general or secular education contexts (Sagala & Andriani, 2019; Seman et al., 2017; Suhady et al., 2020). As a result, there is limited empirical understanding of the distinct challenges faced by teachers in faith-based education systems, where curriculum design

and assessment practices must balance cognitive rigor with theological and moral imperatives. Meanwhile, in Islamic educational institutions, such as the State Islamic Institute of Sultan Amai Gorontalo, teachers operate within a distinct epistemological framework that informs how Higher Order Thinking Skills (HOTS) are interpreted, adapted, and implemented in classroom practice. This context requires a careful alignment between theological foundations and contemporary pedagogical demands. Although initiatives to provide training in HOTS-aligned assessment design are currently being developed, the extent to which these efforts effectively bridge conceptual understanding with practical application remains a critical area for further inquiry and evaluation (Destiniar et al., 2020; Poerwanti & Tribudhiarto, 2020), yet findings indicate that many religious educators still struggle to translate conceptual understanding of HOTS into authentic, cognitively demanding tasks. These challenges are further intensified by structural constraints, such as limited access to teaching resources, inadequate mentoring support, overloaded curricula, and restrictive instructional time allocations. Such conditions hinder teachers' capacity to effectively design and implement HOTS-oriented assessments, thereby limiting opportunities for meaningful cognitive engagement in the classroom (Suardipa, 2018).

Existing scholarship has identified critical factors affecting HOTS implementation, such as mastery of content knowledge, clarity of assessment objectives, and prior exposure to higher-order question construction (Ahmad et al., 2018; Hakim et al., 2021; Saido et al., 2018; Syafitri & Anggraini, 2023), but has not sufficiently examined how these interact with the religious and cultural specificities of Islamic education. Furthermore, although existing studies highlight that teachers' capacity to develop valid and reliable HOTS-oriented assessments remains limited (Harjanti et al., 2022; Mahmudati, 2023; Yudha, 2023), there is a notable lack of systematic inquiry into the extent to which the Ministry of Religious Affairs' Professional Teacher Programs effectively address these deficiencies.

The research gap is identified into twofold. First, a contextual gap exists due to the limited empirical investigations on HOTS-based assessment



practices within faith-based educational settings, particularly in Islamic higher education institutions, where integrating theological frameworks with contemporary pedagogical demands presents unique challenges. Second, a programmatic gap emerges from the limited evidence on the effectiveness of the Ministry of Religious Affairs' Professional Teacher Program in enhancing religious educators' assessment literacy and, more critically, on whether such advancements are successfully translated into the sustained application of HOTS-oriented practices within actual classroom contexts.

This study addresses these gaps by employing Bloom's Revised Taxonomy as an analytical lens for evaluating the cognitive orientation of assessment items produced by Professional Teacher Program (PPG) students. Bloom's taxonomy serves here as a diagnostic tool for measuring the depth of cognitive engagement in teacher-generated items and as a benchmark for professional competency in HOTS-oriented assessment design (Krathwohl & Krathwohl, 2017). In parallel, the study is informed by the Assessment for Learning (AfL) framework, which emphasizes formative and diagnostic assessment practices that actively support student thinking and metacognition.

The present study positions itself at the critical juncture of teacher professional development, assessment theory, and religious education reform. It offers an empirical and context-sensitive analysis of how higher-order thinking skills (HOTS) are conceptualized and enacted within Islamic teacher education. In contrast to previous studies that tend to address HOTS in broad or generalist terms, this research probes more deeply into how national policy directives and the Ministry of Religious Affairs' Professional Teacher Programs are translated into classroom practices in faith-based educational settings. By situating the investigation within the epistemological traditions and institutional logics that shape Islamic teacher education, the study seeks to uncover how these frameworks mediate teachers' capacity to design cognitively demanding assessments that genuinely foster advanced problem-solving and critical thinking skills.

The novelty and urgency of this inquiry lie in its dual contribution. Locally, it responds to pressing concerns over the effectiveness of teacher



professionalization programs in Indonesia, particularly in their ability to bridge policy aspirations with classroom realities in religious institutions. Globally, it contributes to ongoing debates about the contextual adaptability of 21st-century learning frameworks, illustrating how culturally and epistemologically distinct settings engage with and sometimes challenge the presumed universality of HOTS pedagogies. In doing so, the study raises important questions about whether professional development programs can meaningfully enhance assessment literacy in contextually grounded and globally relevant ways. The findings are expected to sharpen theoretical understandings of how belief systems and institutional traditions shape assessment practices. They also provide practical insights for curriculum reform, targeted professional training, and policy interventions to align religious teacher education with global competencies without eroding its distinctive epistemological foundations.

B. Method

This study employed an exploratory cross-sectional design (Köhler et al., 2023) to capture a snapshot of teachers' competencies in constructing Higher-Order Thinking Skills (HOTS)-based test items within a defined period. This approach aligns with the research objective of evaluating the effectiveness of the Ministry of Religious Affairs' Professional Teacher Program (PPG) at the State Islamic Institute of Sultan Amai Gorontalo in enhancing teachers' capacity to apply HOTS in both formative and summative assessments (Sagala & Andriani, 2019). A total of 495 *Madrasahs* and Islamic education teachers from various regions of Indonesia participated in the study. Participants were selected through simple random sampling to ensure representativeness across teaching experience, with many having more than five years in the profession.

Data collection involved two complementary techniques: direct observation and documentation review. Observations were conducted over a one-day training session for each participant group, focusing on the real-time process of formulating HOTS-based questions during practical workshops. Trained observers used a standardized scoring rubric to evaluate teachers'



work according to four core indicators proposed by Stedman (2011): (1) understanding of HOTS question concepts, (2) alignment with subject material, (3) creation of items at the HOTS cognitive level (C4–C6), and (4) clarity of question presentation (Na et al., 2021). Each indicator was measured through five specific items, yielding a total of 20 items. Scoring used a 5-point Likert scale, with descriptors for each score level to ensure consistency.

The documentation review analyzed written test items and assessment tools produced by participants during the training. Each product was assessed against the same HOTS criteria used in observations, allowing triangulation between observed competencies and actual outputs. The instrument underwent expert content validation using Aiken's *V* method (Merino-Soto, 2023), confirming that all 20 items met the validity threshold. Reliability analysis yielded a coefficient of 0.96, indicating excellent internal consistency.

The data was analyzed in three sequential stages. First, instructors scored all assessment products using the validated rubric. Second, structured feedback sessions were held to refine teachers' HOTS-based question design skills. Third, the Graded Response Model (GRM) from Item Response Theory (IRT) was applied to the collected data using R software. GRM was chosen for its capacity to handle ordinal Likert-type data and to estimate both item difficulty and discrimination parameters with precision (Ackerman et al., 2022; An & Yung, 2014; Fayyaz et al., 2024).

So, this study was conducted in accordance with ethical research standards. All participants were informed about the study's objectives, and their voluntary participation was ensured. Written informed consent was obtained prior to data collection, and participants were assured of the confidentiality and anonymity of their responses. The research protocol was reviewed and approved by the Institutional Research Ethics Committee of the State Islamic Institute of Sultan Amai Gorontalo.

C. Results and Discussion

After the data collection procedures were completed and analyzed using the Graded Response Model (GRM) approach, this section presents the research



findings that reflect the level of teachers' ability to develop HOTS-based test items during their participation in the Ministry of Religious Affairs' Professional Teacher Program (PPG). The results are then analyzed and discussed in depth by referring to the theoretical framework and previous studies, to gain a more comprehensive understanding of the teachers' position and challenges in developing assessments that promote higher-order thinking skills.

1. Results

a. Evaluation of HOTS question design

This study assesses pre-service teacher education students' proficiency in designing Higher-Order Thinking Skills (HOTS) questions, emphasizing their ability to apply HOTS principles in assessments. The analysis provides insights into key evaluation indicators, item difficulty levels, and the overall effectiveness of the designed questions.

Teachers' abilities in designing HOTS-based questions were evaluated using four core indicators: (1) understanding of HOTS concepts, (2) alignment of questions with instructional content, (3) cognitive level of questions (C4–C6), and (4) clarity of question formulation. Across all indicators, most participants demonstrated moderate levels of proficiency. In terms of conceptual understanding, many teachers correctly identified that HOTS items require analytical, evaluative, or creative responses rather than simple recall. For example, a participant teaching Islamic history developed the following question: *"Analyze the influence of the Abbasid dynasty's educational policies on the development of science in the Islamic Golden Age. Provide at least two supporting arguments"*.

This question was correctly classified at the C4 (analysis) level and aligned precisely with curricular objectives. However, some items fell short of HOTS criteria despite being thematically relevant. For instance: *"Mention the names of three prominent scientists during the Abbasids' period"*.

While relevant to the topic, this item only requires factual recall (C1), thus not meeting HOTS standards. Clarity of formulation varied. While some



items were well-structured, others contained ambiguous phrasing or multiple cognitive demands within a single question, potentially reducing validity.

b. Item characteristic analysis

Item analysis using the Graded Response Model (GRM) provided discrimination and difficulty parameters for each question and four representative items. Items 1, 10, 16, and 20 – are presented in Table 1.

Table 1. Item characteristic analysis

Item	Discrimination (a)	Difficulty (b1)	Difficulty (b2)	Difficulty (b3)
1	1.31	-2.26	-1.04	0.60
10	1.55	-1.74	-0.79	0.51
16	1.72	-1.81	-0.74	0.55
20	1.61	-1.92	-0.88	0.73

In Table 1, item 1 “Evaluate the effectiveness of zakat management in your local community and suggest two improvements”. It is shown to have relatively low difficulty thresholds, indicating accessibility for participants with moderate ability levels. Conversely, Item 20 “Design a community-based program integrating Islamic economic principles with modern cooperative business models” exhibited higher difficulty, requiring integration of conceptual understanding and application in novel contexts. In other words, all four items had discrimination values above 1.3, indicating a strong capacity to differentiate between participants with varying proficiency in HOTS question design.

To sum up, these findings demonstrate the effectiveness of the instrument in capturing nuanced differences in teacher competency. The GRM analysis supports the conclusion that the assessment tool can differentiate performance levels and reliably identify strengths and weaknesses in HOTS-based assessment design. This reinforces the instrument’s suitability for future use in Teacher Training Institution (LPTK) contexts, especially where a fine-grained understanding of assessment literacy is needed. Moreover, it is supported by the Test Information Function (TIF) analysis below.

Meanwhile, the test information function indicated the instrument’s measurement precision across ability levels. Figure 1 shows that peak information

was obtained at a theta level of -1.7 , with an information value of 11.5 and a measurement error of 0.29 . This indicates high measurement reliability within the moderate to lower-high ability range.

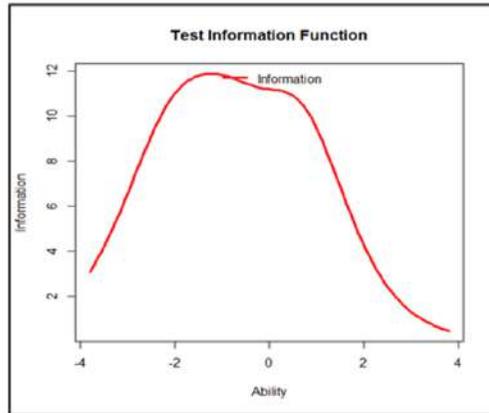


Figure 1. Information function for HOTS question creation ability

In conclusion, based on the explanation, the test instrument offers high measurement accuracy and reliability in capturing teachers' abilities to construct HOTS-based questions, especially among those with moderate proficiency, thus providing robust evidence to support the study's goal of evaluating the effectiveness of Teacher Training Institution (*LPTK*) in HOTS implementation.

c. Interpretation of Ability Levels

Ability levels were categorized into five intervals ranging from very low to very high (Table 2).

Table 2. Ability Level Categories

Ability Interval	Ability Interpretation
3.00 - 2.00	Very High
2.00 - 1.00	High
1.00 - -1.00	Average
-1.00 - -2.00	Low
-2.00 - -3.00	Very Low

Most participants scored between -1.00 and 1.00 , falling into the average proficiency category. For instance, one participant in this range designed a



question that partially met HOTS criteria: “Compare the leadership approaches of Caliph Umar bin Khattab and Caliph Ali bin Abi Talib in managing societal justice”. While the comparison element demonstrated analytical intent (C4), the absence of specific evaluative prompts or synthesis tasks limited its cognitive depth.

Importantly, these findings directly support the research objective by highlighting a clear performance gap in the ability of pre-service and in-service teachers to design assessments that target higher-order cognitive domains. Although foundational HOTS knowledge exists, there is a pressing need for structured, targeted instructional support within the Professional Teacher Program. Without such intervention, many educators may continue to rely on lower-order thinking tasks that do not fully support the development of students’ critical and creative thinking skills.

In summary, the interpretation of ability levels underscores that most teachers are at a moderate level of competence, revealing both the impact and the limitations of current training efforts. This insight provides a compelling rationale for enhancing the *PPG* curriculum with more practice-oriented, skill-based interventions to elevate teachers’ proficiency in constructing effective HOTS-based assessments, thus addressing the core aim of this study.

2. Discussion

The findings of this study reveal that Indonesian pre-service teachers demonstrate moderate proficiency in constructing Higher-Order Thinking Skills (HOTS) based questions at a level suggesting basic conceptual familiarity but insufficient mastery for designing cognitively demanding assessments. This aligns with prior research by Ishartono et al., which emphasizes that robust subject-matter knowledge is a prerequisite for effective HOTS implementation. (Ishartono et al., 2021; Sidiq et al., 2021; Tyas et al., 2019). However, unlike those studies that implicitly assume a linear correlation between knowledge acquisition and HOTS capability, our data challenge this assumption. The evidence demonstrates a persistent inability to translate theoretical understanding into practice, particularly in differentiating between lower-order and higher-order cognitive demands, suggesting that knowledge alone does not guarantee application.

This mismatch between theoretical exposure and applied competence reveals a structural weakness in teacher education programs. Despite formal instruction, many participants default to conventional, recall-based questioning patterns, indicating that HOTS frameworks in current training remain procedural rather than transformative. From a constructivist and inquiry-based learning perspective, this suggests that higher-order cognition cannot be cultivated through transmissive instruction alone; it requires iterative, reflective, and contextually embedded practice. The lack of sustained engagement with authentic assessment scenarios limits the development of metacognitive and analytical capacities in pre-service teachers, which are essential for fostering deep learning in students.

Furthermore, the findings disrupt another prevailing assumption in the literature: that teaching experience naturally enhances HOTS assessment proficiency. The analysis found no significant correlation between years of teaching experience and competence in designing HOTS-oriented questions. This supports critiques in teacher development literature, e.g., Miharja et al. (2022) and Saraswati et al. (2023) that experience, without targeted skill-specific training, does not necessarily evolve into expertise (Miharja et al., 2022; Saraswati et al., 2023). So, it is implied that teacher development programs cannot rely on experience alone to strengthen HOTS assessment proficiency. Instead, targeted, skill-specific training is essential. Policies should prioritize structured professional development and continuous support to ensure teachers acquire and sustain the expertise needed for effective HOTS-oriented assessment practices.

In addition, academic background emerged as a stronger predictor of HOTS proficiency than professional tenure. Teachers with specialized training in pedagogy or deep disciplinary knowledge outperformed those with general teaching qualifications. This resonates with the argument that disciplinary thinking frameworks in science, humanities, or religious education equip teachers with the cognitive tools necessary for constructing higher-order assessments. However, in many faith-based or rural contexts, teachers' own educational histories lack exposure to inquiry-based learning methods, perpetuating what Abdullah et al. (2021) and Suanto et al. (2023) identify as a cycle of assessment



conservatism (Abdullah et al., 2021; Suanto et al., 2023). Without authentic engagement in higher-order learning, teachers are less confident and less effective in promoting it for their students.

However, the moderate proficiency levels observed suggest that HOTS training in its current form is overly theoretical and insufficiently contextualized. Consistent with Setyarini et al., (2023), Suryani (2024), and Yani & Mulyadi (2022), teacher education must move beyond “*teaching teachers how to write HOTS questions*” toward a deeper epistemological engagement with why higher-order thinking matters for learning (Setyarini et al., 2023; Suryani, 2024; Yani & Mulyadi, 2022). This requires embedding HOTS principles into broader pedagogical theories of cognition, metacognition, and knowledge construction.

Applied and iterative training through collaborative design tasks, peer critique, and structured feedback can bridge the gap between conceptual understanding and practical competence (Abella-Peña & García-Martínez, 2023; Gu et al., 2022; Lawrence et al., 2024). Without such sustained, reflective engagement, teachers risk internalizing HOTS as a checklist requirement rather than a pedagogical orientation.

Moreover, HOTS training should be continuous and embedded within teacher competency frameworks, rather than delivered as isolated workshops. As Aleo et al., (2024) noted, lifelong professional learning, supported by mentorship and reflective practice, is vital for developing durable assessment literacy (Aleo et al., 2024). Integrating AI-driven training platforms offers a promising route to provide adaptive, scalable, and context-sensitive practice opportunities, particularly valuable in geographically dispersed or resource-limited educational systems (Brovina & Arifi, 2023; Abd Ghani et al., 2025).

From a global standpoint, this study reinforces the position of HOTS-oriented assessment as a core competency in 21st-century teacher standards, echoing the OECD’s TALIS emphasis on critical thinking and problem-solving as universal teacher benchmarks (Niu et al., 2021; Ranta et al., 2022). The Indonesian case contributes to international debates by illustrating that moderate competence

in HOTS assessment is not unique to one country but reflects a broader pattern across education systems struggling to reconcile traditional pedagogies with global learning goals (Muchson et al., 2024; Rekan et al., 2025).

Countries with religious or traditional educational systems often face tensions between preserving epistemological heritage and adopting cognitively progressive pedagogies (Anita et al., 2023; Nurtawab, 2019). This study demonstrates that these tensions can be navigated in Indonesia's HOTS framework, which integrates cultural relevance with cognitive rigor, supporting Paul & Alqraini's (2019) claim that assessment literacy must blend technical skill with reflective capacity (Paul & Alqraini, 2019). It also confirms Botetano and Abrahamson's (2022) argument that teacher expertise in assessment must be deliberately cultivated rather than assumed (Botetano & Abrahamson, 2022).

Importantly, the findings resonate with UNESCO's (2021) global equity agenda (Hadjeris, 2021), showing that HOTS integration is not merely an instructional upgrade but a pathway toward equitable, student-centered education (Hadjeris, 2021). Education systems can foster critical, analytical, and creative thinking across socio-cultural boundaries by equipping teachers in diverse contexts with the capacity to design cognitively challenging assessments.

Yet, as Basterfield et al., (2023), Jerebine et al., (2024), and Ulvik et al., (2022) have found, policy endorsement of HOTS does not guarantee classroom implementation (Basterfield et al., 2023; Jerebine et al., 2024; Ulvik et al., 2022). This study confirms that policy-practice gaps remain significant, and bridging them will require long-term professional support, conceptual alignment, and coherent policy frameworks.

Theoretically, this study challenges linear development models of teacher competency by showing that knowledge acquisition, experience, and proficiency in HOTS assessment are not automatically correlated. It advances a practice-integrated model where theoretical frameworks must be reinforced through sustained, reflective, and applied learning cycles. Internationally, it



positions HOTS-based assessment design as a cross-cultural competency with universal applicability but requiring context-sensitive adaptation.

By offering a replicable, evidence-based training model, this study contributes to global teacher education discourse, particularly for balancing traditional systems values with modern learning imperatives. The findings underscore the need for a shift from procedural compliance to cognitive transformation in teacher development, which is essential for preparing educators capable of cultivating the higher-order thinking that 21st-century learners require, as shown in Figure 2 below.

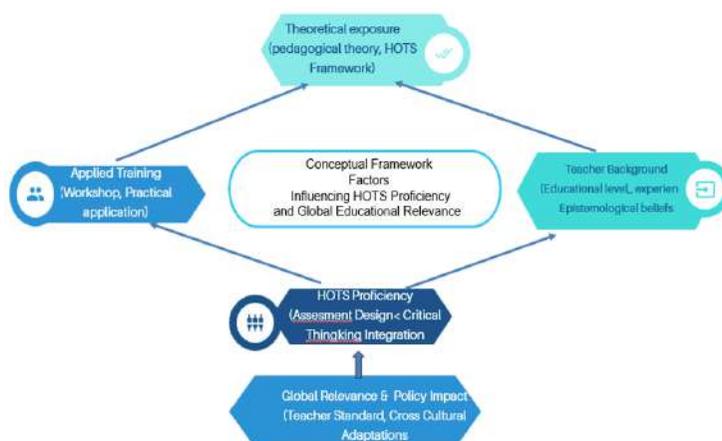


Figure 2. Factors influencing HOTS proficiency and global educational relevance

The conceptual framework in Figure 2 illustrates the interrelated factors that shape teachers' proficiency in Higher Order Thinking Skills (HOTS) and their broader educational relevance. At the core, HOTS proficiency, defined as the ability to design assessments that integrate critical thinking, emerges from three key dimensions: theoretical exposure, applied training, and teacher background. First, theoretical exposure provides teachers with the foundational understanding of pedagogical theories and HOTS frameworks (Krathwohl & Krathwohl, 2017; Tanujaya et al., 2017; Ahmad et al., 2018). This dimension emphasizes why HOTS matters for cognitive development, moving beyond technical item writing to a deeper comprehension of learning processes (Na et

al., 2021; Suryani, 2024). Second, applied training bridges theory and practice through workshops, iterative design exercises, and formative feedback, ensuring that teachers internalize and apply the principles of HOTS in authentic classroom contexts (Setyarini et al., 2023; Abella-Peña & García-Martínez, 2023; Gu et al., 2022). Third, teacher background, including educational level, prior experience, and epistemological beliefs, is a filter through which theory and training are interpreted and enacted in practice (Abdullah et al., 2021; Suanto et al., 2023).

These factors influence HOTS proficiency, which is positioned as a driver of global educational reform. At the policy level, enhanced proficiency translates into global relevance and systemic impact, informing teacher standards, curriculum development, and cross-cultural adaptations. This framework underscores that improving HOTS is not only a technical matter but also a systemic and globally significant educational reform.

Furthermore, the integration of HOTS within teacher education programs reflects the international movement toward developing learners with critical, creative, and ethical reasoning abilities that align with the demands of the twenty-first century. Such integration connects local educational contexts, including religious and cultural frameworks, to global learning standards, thereby promoting inclusive and context-responsive pedagogy (Hadjeris, 2021; Basterfield et al., 2023; Ulvik et al., 2022). Strengthening teachers' HOTS proficiency thus contributes not only to national educational quality but also to the broader agenda of equitable, human-centered, and sustainable global education envisioned by UNESCO's Sustainable Development Goal 4.

The findings of this study contribute to the global discourse on teacher competence in designing higher-order thinking assessments, particularly in culturally and religiously grounded education systems. By situating HOTS within the epistemological framework of Islamic education, this research demonstrates that global educational standards can be localized without undermining cultural authenticity. The study highlights that the challenge of balancing tradition and innovation in pedagogy is shared across diverse educational systems worldwide. Therefore, the Indonesian experience offers a



relevant model for other Muslim-majority and developing countries seeking to harmonize international learning standards with their local educational values. This contribution reinforces the universality of HOTS as a 21st-century learning competency while acknowledging its necessary contextual adaptation.

Despite its valuable contributions, this study has several limitations. It employed an exploratory cross-sectional design, which only captured teachers' abilities at a single point in time and did not measure long-term improvement. The research context was limited to participants of the Professional Teacher Program at the State Islamic Institute of Sultan Amai Gorontalo, so the findings may not be fully generalizable to other teacher training institutions. Data were collected from one-day training sessions and document reviews, reflecting short-term performance rather than sustained pedagogical competence. Moreover, the study did not examine contextual or affective factors—such as institutional support or teaching workload—that may influence teachers' ability to design HOTS-based assessments. Future research should therefore employ longitudinal or mixed-method approaches to provide a more comprehensive understanding.

D. Conclusion

This study reveals that teachers participating in the Professional Teacher Education Program (PPG) under the Indonesian Ministry of Religious Affairs demonstrate a moderate level of competence in developing Higher-Order Thinking Skills (HOTS)-based test items. Through the application of the Graded Response Model (GRM), the study identified a consistent pattern in which teachers possess adequate conceptual understanding but encounter significant challenges in translating theoretical knowledge into practice, particularly in designing assessment items that reflect analytical, evaluative, and creative dimensions of thinking. These findings collectively highlight that while the conceptual integration of HOTS within teacher training programs has begun, its practical realization remains limited. The synthesis of evidence thus confirms that enhancing teachers' assessment literacy requires sustained, contextualized, and practice-based interventions to ensure that HOTS

principles are effectively embedded in real classroom assessment design.

Theoretically, this study contributes to the development of a more integrated understanding of teachers' cognitive assessment competence by combining Bloom's Revised Taxonomy with the Graded Response Model (GRM) as a diagnostic analytical framework. This integration advances the empirical application of psychometric theory in assessing teachers' ability to construct cognitively demanding test items, particularly within religious education contexts. Practically, the study contributes to the improvement of teacher education policy by emphasizing the importance of targeted and continuous pedagogical support within the PPG framework. The research novelty lies in its focus on religious education teachers within Islamic higher education settings—an area that has received limited attention in global discussions on HOTS implementation. By situating HOTS within the epistemological and pedagogical traditions of Islamic education, this study offers a new lens through which local pedagogical practices can be aligned with global 21st-century educational standards.

Building on the research limitations identified earlier, future studies should adopt longitudinal or mixed-method approaches to capture long-term developments in teachers' ability to design HOTS-based assessments. Expanding the context beyond a single institution would also allow for comparative insights between religious and general teacher training institutions, thereby enhancing the generalizability of findings. Further investigation into contextual and affective factors—such as institutional support, teaching workload, and mentoring quality—would deepen the understanding of how these dimensions influence teachers' assessment design skills. Cross-country comparative research, particularly among Muslim-majority education systems, could further explore how cultural and theological frameworks shape the adaptation and sustainability of HOTS-based assessment practices in diverse learning environments.

This study reinforces that fostering higher-order thinking in education is not merely a curricular aspiration but a transformative process that demands



capable, reflective, and culturally grounded teachers. Strengthening teachers' capacity to design cognitively challenging assessments is therefore fundamental to building an education system that cultivates critical, creative, and ethical thinkers. By bridging the gap between national pedagogical reform and global educational goals, this study underscores that the pursuit of HOTS is both an intellectual and moral endeavor – essential for shaping a generation of learners prepared to engage thoughtfully and responsibly with the complexities of the modern world.

Acknowledgment

We extend our deepest gratitude to all those who contributed to the success of this study. First and foremost, we sincerely thank the teachers who participated in study, especially those from the Indonesian Ministry of Religious Affairs' Professional Teacher Program at the State Islamic Institute of Sultan Amai Gorontalo. Their dedication and willingness to contribute to this study have been invaluable.

Bibliography

- Abd Ghani, A. R., Huda, M., Sholihah, M., Tabrani, Z. A., & Hussin, H. (2025). Tahfiz (Memorization Sciences) Curriculum Practice: An Empirical Study from Private School in Johor, Malaysia. *Nazhruna: Jurnal Pendidikan Islam*, 8(1), 233–245. <https://doi.org/10.31538/nzh.v8i1.265>
- Abdullah, A. H., Soh, H. M., Mokhtar, M., Hamzah, M. H., Ashari, Z. M., Ali, D. F., Samah, N. A., Jumaat, N. F., Ibrahim, N. H., Surif, J., & Rahman, S. N. S. A. (2021). Does the Use of Smart Board Increase Students' Higher Order Thinking Skills (HOTS)? *IEEE Access*, 9, 1833-1854. <https://doi.org/10.1109/ACCESS.2020.3042832>
- Abella-Peña, L. E., & García-Martínez, Á. (2023). Conceptual and Methodological Bases of a Training Proposal Supported by ICT for Preservice Chemistry Teachers. *Investigacoes Em Ensino de Ciencias*, 28(1), 23-38. <https://doi.org/10.22600/1518-8795.ienci2023v28n1p23>
- Ackerman, T., Ma, Y., Ma, M., Pacico, J. C., Wang, Y., Xu, G., Ye, T., Zhang,

- J., & Zheng, M. (2022). Item Response Theory. In *International Encyclopedia of Education: Fourth Edition*. <https://doi.org/10.1016/B978-0-12-818630-5.10010-7>
- Ahmad, S., Prahmana, R. C. I., Kenedi, A. K., Helsa, Y., Arianil, Y., & Zainil, M. (2018). The instruments of Higher Order Thinking Skills. *Journal of Physics: Conference Series*, 943(1), 012053. <https://doi.org/10.1088/1742-6596/943/1/012053>
- Albab, U., Budiyo, B., & Indriati, D. (2020). Metacognition Skills and Higher Order Thinking Skills (HOTS) in Mathematics. *Journal of Physics: Conference Series*, 1613(1), p.012017. <https://doi.org/10.1088/1742-6596/1613/1/012017>
- Aleo, G., Pagnucci, N., Walsh, N., Watson, R., Lang, D., Kearns, T., White, M., & Fitzgerald, C. (2024). The Effectiveness of Continuing Professional Development for the Residential Long-Term Care Workforce: A Systematic Review. In *Nurse Education Today*, 137, 106161. <https://doi.org/10.1016/j.nedt.2024.106161>
- An, X., & Yung, Y. F. (2014). Item Response Theory: What It Is and How You Can Use the IRT Procedure to Apply It. *SAS Institute Inc*, 10(4), 364-377. <https://support.sas.com/resources/papers/proceedings14/SAS364-2014.pdf>
- Anita, A., Hasan, M., Warisno, A., Anshori, M. A., & Andari, A. A. (2023). Pesantren, Kepemimpinan Kiai, dan Ajaran Tarekat sebagai Potret Dinamika Lembaga Pendidikan Islam di Indonesia. *Scaffolding: Jurnal Pendidikan Islam Dan Multikulturalisme*, 4(3), 509-524. <https://doi.org/10.37680/scaffolding.v4i3.1955>
- Basterfield, L., Machaira, T., Jones, D., Rapley, T., Araujo-Soares, V., Cameron, N., & Azevedo, L. B. (2023). Early Years Physical Activity and Motor Skills Intervention – A Feasibility Study to Evaluate an Existing Training Programme for Early Years Educators. *Children*, 10(1), 145-. <https://doi.org/10.3390/children10010145>
- Botetano, C., & Abrahamson, D. (2022). The Botetano Arithmetic Method: Introduction and Early Evidence. In *International Journal of Mathematical Education in Science and Technology*, 53(2), 516-534. <https://doi.org/10.1080/0020739X.2020.1867916>
- Brovina, N., & Arifi, D. (2023). Institutional And Non-Institutional Actors In Policy-Making Processes: A Case Study. *Journal of Governance and*



- Regulation, 12(2), 147-155. <https://doi.org/10.22495/jgrv12i2art13>
- Destiniar, D., Mulbasari, A. S., Fuadiah, N. F., Octaria, D., Ningsih, Y. L., Retta, A. M., & Isroqmi, A. (2020). Pelatihan Penyusunan Soal HOTS untuk Mengembangkan Kemampuan Pedagogik Guru. *J-ABDIPAMAS (Jurnal Pengabdian Kepada Masyarakat)*, 4(1), 163-169. <https://doi.org/10.30734/j-abdipamas.v4i1.585>
- Fayyaz, S., Moeinaddini, M., Pourebrahim, S., Khoshnevisan, B., Kazemi, A., Toufighi, S. P., Schjønberg, M. S., & Birkved, M. (2024). Assessing Environmental Enhancement Scenarios in a Petrochemical Port: A Comprehensive Comparison using a Hybrid LCA-GRM Model. *Journal of Cleaner Production*, 445, 141079. <https://doi.org/10.1016/j.jclepro.2024.141079>
- Gu, J., Tang, L., Liu, X., & Xu, J. (2022). Promoting Pre-service Teacher Students' Learning Engagement: Design-Based Research in a Flipped Classroom. *Frontiers in Psychology*, 13, 810275. <https://doi.org/10.3389/fpsyg.2022.810275>
- Hadjeris, F. (2021). Revisiting Sustainable Development Goal 4 in the Context of COVID-19 Pandemic: A Case Study of Online Teaching in Algerian Higher Education Institutions. *Human Behavior and Emerging Technologies*, 3(1), 160-168. <https://doi.org/10.1002/hbe2.245>
- Hakim, B. R., Ardiyanto, A., & Wijayanti, A. (2021). Analisis Hots pada Instrumen Penilaian Siswa Kelas IV Sekolah Dasar. *Wawasan Pendidikan*, 1(2), 246-254. <https://doi.org/10.26877/wp.v1i2.9028>
- Harjanti, F. D., Anam, F., & Kirana, A. (2022). Pengembangan Kompetensi Pedagogi Guru Pamong Sekolah Mitra Fakultas Bahasa dan Sains Universitas Wijaya Kusuma Surabaya melalui Pembelajaran HOTS. *Jurnal Pengabdian Kepada Masyarakat*, 2(1), 3741-3748. <https://doi.org/10.53625/jabdi.v2i1.2271>
- Ishartono, N., Dessty, A., Prayitno, H. J., & Sidiq, Y. (2021). The Quality of HOTS-Based Science Questions Developed by Indonesian Elementary School Teachers. *Journal of Education Technology*, 5(2), 236-245. <https://doi.org/10.23887/jet.v5i2.33813>
- Jerebine, A., Heering, T., & Barnett, L. M. (2024). Educator-Perceived Barriers and Facilitators to Structured-Physical Activity in Early Childhood Centres: A Systematic Review. *Research Quarterly for Exercise and Sport*, 95(1), 243-262. <https://doi.org/10.1080/02701367.2023.2193243>

- Köhler, T., Rumyantseva, M., & Welch, C. (2023). Qualitative Restudies: Research Designs for Retheorizing. *Organizational Research Methods*, 28(1), 32-57. <https://doi.org/10.1177/10944281231216323>
- Krathwohl, D. R., & Krathwohl, D. R. (2017). A Revision of Bloom's Taxonomy: *Routledge*.
- Lawrence, L. E. M., Echeverria, V., Yang, K., Aleven, V., & Rummel, N. (2024). How Teachers Conceptualise Shared Control with an AI Co-Orchestration Tool: A Multiyear Teacher-Centred Design Process. *British Journal of Educational Technology*, 55(3), 823-844. <https://doi.org/10.1111/bjet.13372>
- Mahmudati, R. (2023). The Characteristics and Developing an Assessment Instrument of High Order Thinking Skill (HOTS) Items With 4-D Models. *Jurnal Pendidikan Matematika Dan IPA*, 14(1), 107-117. <https://doi.org/10.26418/jpmipa.v14i1.55497>
- Merino-Soto, C. (2023). Aiken's V Coefficient: Differences in Content Validity Judgments. *MHSalud*, 20(1), 23-32. <https://doi.org/10.15359/mhs.20-1.3>
- Miharja, F. J., Fauzi, A., & Hudha, A. M. (2022). Pelatihan Pengembangan Soal Berbasis HOTS bagi Guru Sekolah Dasar. *Sasambo: Jurnal Abdimas (Journal of Community Service)*. <https://doi.org/10.36312/sasambo.v4i2.677>
- Muchson, M., Cobern, W. W., & Saefi, M. (2024). The Science Education Research trends (SERT) in Indonesian secondary schools: A Systematic Review and Bibliometrics Study. *Cogent Education*, 11(1), 2308407. <https://doi.org/10.1080/2331186X.2024.2308407>
- Na, S. J., Ji, Y. G., & Lee, D. H. (2021). Application of Bloom's Taxonomy to Formative Assessment in real-Time Online Classes in Korea. *Korean Journal of Medical Education*, 33(3), 191-201. <https://doi.org/10.3946/KJME.2021.199>
- Niu, S. J., Niemi, H., Harju, V., & Pehkonen, L. (2021). Finnish Student Teachers' Perceptions of Their Development of 21st-Century Competencies. *Journal of Education for Teaching*, 47(5), 638-653. <https://doi.org/10.1080/02607476.2021.1951602>
- Nugraha, E. (2020). Implementasi Program Tahfizh Qur'an di PAUD Inklusif dengan Model HOTS. *As-Sibyan: Jurnal Pendidikan Anak Usia Dini*, 5(2), 95-106. <https://ftk.uinbanten.ac.id/journals/index.php/assibyan/article/view/10551>

- Nurtawab, E. (2019). The Decline of Traditional Learning Methods in Changing Indonesia: Trends of Bandungan-Kitab Readings Pesantrens. In *Studia Islamika*, 26(3), 511-541. <https://doi.org/10.36712/sdi.v26i3.11026>
- Osborne, J. F., Henderson, J. B., MacPherson, A., Szu, E., Wild, A., & Yao, S. Y. (2016). The Development and Validation of a Learning Progression for Argumentation in Science. *Journal of Research in Science Teaching*, 53(6), 821-846. <https://doi.org/10.1002/tea.21316>
- Paul, P. V., & Alqraini, F. (2019). Conclusion: Perspectives on Language, Literacy, and Deafness. *Education Sciences*, 9(4), 286. <https://doi.org/10.3390/educsci9040286>
- Poerwanti, J. I. S., & Tribudhiarto. (2020). Pelatihan Merancang Instrumen Asesmen High Order Thinking Skills pada Guru-Guru SD di Kecamatan Laweyan Surakarta. *Jurnal Widya Laksana*, 9(1), 66-73. <https://doi.org/10.23887/jwl.v9i1.16660>
- Ranta, M., Kruskopf, M., Kortessalmi, M., Kalmi, P., & Lonka, K. (2022). Entrepreneurship as a Neglected Pitfall in Future Finnish Teachers' Readiness to Teach 21st Century Competencies and Financial Literacy: Expectancies, Values, and Capability. *Education Sciences*, 12(7), 463. <https://doi.org/10.3390/educsci12070463>
- Rekan, A. A., Howell, H. R., Salleh, M. N. M., Rosin, N., Tabrani, Z. A., Ma`arif, M. A., & Adnan, M. A. M. (2025). Arabic Language Curriculum as a Foundation for Strengthening Religious Education in Public Higher Education. *Jurnal Pendidikan Agama Islam*, 22(1), 97-121. <https://doi.org/10.14421/jpai.v22i1.11340>
- Sagala, P. N., & Andriani, A. (2019). Development of Higher-Order Thinking Skills (HOTS) Questions of Probability Theory Subject Based on Bloom's Taxonomy. *Journal of Physics: Conference Series*, 1188(1), p.012025. <https://doi.org/10.1088/1742-6596/1188/1/012025>
- Saido, G. M., Siraj, S., Nordin, B. A. B., & Al_Amedy, O. S. (2018). Higher Order Thinking Skills among Secondary School Students in Science Learning. *IEEE Transactions on Antennas and Propagation*, 66(9), 4936-4941. <https://doi.org/10.1109/TAP.2018.2845550>
- Saraswati, N., Inderawati, R., & Vianty, M. (2023). Applying Higher Order Thinking Skills (Hots) Pedagogy: A Case Study of an EFL Core Teacher. *English Review: Journal of English Education*, 11(3), 767-776. <https://doi.org/10.25134/erjee.v11i3.7818>

- Seman, S. C., Yusoff, W. M. W., & Embong, R. (2017). Teachers challenges in teaching and learning for Higher Order Thinking Skills (HOTS) in Primary School. *International Journal of Asian Social Science*, 7(7), 534–545. <https://doi.org/10.18488/journal.1.2017.77.534.545>
- Setyarini, S., Salim, H., & Purnawarman, P. (2023). Higher-Order Thinking Skills (HOTS)-Based Literacy Media: An Innovative Learning Strategy to Promote the Secondary Students' Social Awareness. *Forum for Linguistic Studies*, 5(2), 1706. <https://doi.org/10.59400/FLS.V5I2.1706>
- Sidiq, Y., Ishartono, N., Dessty, A., Prayitno, H. J., Anif, S., & Hidayat, M. L. (2021). Improving Elementary School Students' Critical Thinking Skill in Science Through Hots-Based Science Questions: A Quasi-Experimental Study. *Jurnal Pendidikan IPA Indonesia*, 10(3), 378-386. <https://doi.org/10.15294/JPII.V10I3.30891>
- Solissa, J., Hasbullah, B., Rumahlewang, E., & Latar, I. M. (2023). Pendampingan dalam Mengenal dan Menyelesaikan Soal Uji Pengetahuan (UP) PPG Berbasis HOTS pada guru. *German Für Gesellschaft (J-Gefüge)*, 2(2), 79–85. <https://doi.org/10.30598/jgefuege.2.2.79-85>
- Suanto, E., Maat, S. M., & Zakaria, E. (2023). The Effectiveness of the Implementation of Three Dimensions Geometry KARA Module on Higher Order Thinking Skills(HOTS) and Motivation. *International Journal of Instruction*, 16(3), 95-116. <https://doi.org/10.29333/iji.2023.1636a>
- Suardipa, I. P. (2018). Guru sebagai Agen Inovator Berbasis Higher Order Thinking Skills. *Purwadita*, 2(2), 73–83.
- Suhady, W., Roza, Y., & Maimunah, M. (2020). Pengembangan Soal untuk Mengukur Higher Order Thinking Skill (HOTS) Siswa. *Jurnal Gantang*, 5(2), 143–150. <https://doi.org/10.31629/jg.v5i2.2518>
- Suryani, H. (2024). Higher Order Thinking Skills in EFL Classrooms: Evidence in Five Secondary Schools. *Kasetsart Journal of Social Sciences*, 45(1), 103–112. <https://doi.org/10.34044/j.kjss.2024.45.1.11>
- Syafitri, A., & Anggraini, D. P. (2023). Workshop Pembuatan Soal HOTS bagi Guru di SMA Negeri 1 Secanggang. *Jurnal Pengabdian Kepada Masyarakat*, 4(1), 34–39. <https://jurnal.stkipalmaksum.ac.id/index.php/jpkm/article/view/377>
- Tanujaya, B., Mumu, J., & Margono, G. (2017). The Relationship between Higher Order Thinking Skills and Academic Performance of Student

- in Mathematics Instruction. *International Education Studies*, 10(11), 78–85. <https://doi.org/10.5539/ies.v10n11p78>
- Tibrani, M., Madang, K., & Anggraini, N. (2023). Pengembangan Asesment Higher Order Thinking Skills (HOTS) Berbasis Computer Based Test (CBT) pada Mata Kuliah Prinsip Pengajaran dan Asesmen yang Efektif. *Bioscientist: Jurnal Ilmiah Biologi*, 11(2), 1509–1519. <https://doi.org/10.33394/bioscientist.v11i2.9639>
- Tyas, M. A., Nurkamto, J., Marmanto, S., & Laksani, H. (2019). Developing Higher Order Thinking Skills (HOTS) – Based Questions: Indonesian EFL Teachers’ Challenges. In *Proceedings of the International Conference on Future of Education*, 2(1), 52-63. <https://doi.org/10.17501/26307413.2019.2106>
- Ulvik, M., Eide, H. M. K., Eide, L., Helleve, I., Jensen, V. S., Ludvigsen, K., Roness, D., & Torjussen, L. P. S. (2022). Teacher Educators Reflecting on Case-Based Teaching—A Collective Self-Study. *Professional Development in Education*, 48(4), 657-671. <https://doi.org/10.1080/19415257.2020.1712615>
- Winarti, W., Hairida, H., & Lestari, I. (2021). Deskripsi Kemampuan Guru Membuat Soal Berdasarkan pada Kurikulum 2013 di Sekolah Menengah Atas Kabupaten Landak. *Jurnal Ilmiah Wahana Pendidikan*, 7(2), 108–115. <https://doi.org/10.5281/zenodo.4659018>
- Yani, A., & Mulyadi, A. (2022). Higher Order Thinking Skills (HOTS) - Oriented Learning in Geography Learning. *Jurnal Pendidikan Ilmu Sosial*, 31(1), 1-5-118. <https://doi.org/10.17509/jpis.v31i1.47407>
- Yudha, R. P. (2023). Higher Order Thinking Skills (HOTS) Test Instrument: Validity and Reliability Analysis with the Rasch Model. *EduMa: Mathematics education learning and teaching*, 12(1), 21–38. <https://doi.org/10.24235/eduma.v12i1.9468.g5505>