



Vol. 14, No. 1, January 2026	Copyright © 2026, is licensed under a CC-BY-SA
Pages: 689-714	Publisher: SCAD Independent
DOI: <a href="https://doi.org/10.26811/peuradeun.v14i1.1459">https://doi.org/10.26811/peuradeun.v14i1.1459</a>	P-ISSN: 2338-8617 / E-ISSN: 2443-2067

## Affective Assessment within Madrasa Teachers: An Analysis of Iterative Thematic Inquiry

Zubaidah<sup>1</sup>; Suraiya<sup>2</sup>; Abdul Manan<sup>3</sup>; Kamarullah<sup>4</sup>;  
Rahimah Embong<sup>5</sup>; Nurrahmi<sup>6</sup>

<sup>1,2,3,6</sup>Universitas Islam Negeri Ar-Raniry, Banda Aceh, Indonesia

<sup>4</sup>Sekolah Menengah Atas Negeri 16 Banda Aceh, Indonesia

<sup>5</sup>Universiti Sultan Zainal Abidin, Malaysia

<sup>1</sup>Correspondence Email: [zubaidah@ar-raniry.ac.id](mailto:zubaidah@ar-raniry.ac.id)

Received: April 22, 2024	Accepted: January 23, 2026	Published: January 30, 2026
Article Url: <a href="https://journal.scadindependent.org/index.php/jipeuradeun/article/view/1459">https://journal.scadindependent.org/index.php/jipeuradeun/article/view/1459</a>		

### Abstract

*Effective assessment of students' attitudes, behaviors, and values is a key component of value-based education reforms, yet its implementation often remains uneven at the classroom level. This study examines how affective assessment policies are enacted in madrasa education under the KDJPI framework in Aceh Province, Indonesia. Using a qualitative research design grounded in Iterative Thematic Inquiry, the study analyzes data from questionnaires administered to 224 madrasa teachers, semi-structured interviews with six teachers (E1–E6), and relevant instructional documents. The findings show that the persistent gap between affective assessment policy and practice cannot be sufficiently explained by teachers' lack of commitment or competence. Although teachers generally demonstrate positive orientations toward affective assessment instruments such as reflective journaling and checklists, their implementation remains inconsistent due to structural constraints, including limited instructional time, large class sizes, administrative demands, and insufficient institutional support. In response, teachers adopt adaptive strategies that reflect professional agency but also reveal the limits of individual effort in the absence of coherent governance arrangements. By conceptualizing affective assessment as a negotiated and structurally constrained practice, this study contributes to debates on policy enactment and non-cognitive assessment in value-based education systems, with implications extending beyond the specific socio-religious context of Aceh.*

**Keywords:** *Affective assessment; Policy enactment; Teacher agency; Madrasa education; Value-based education.*

## A. Introduction

The evaluation of students' learning outcomes has long been dominated by cognitive measurement, yet contemporary educational discourse increasingly emphasizes that intellectual development alone is insufficient for fostering holistic learners. Emotional intelligence, resilience, reflective capacity, and social responsibility are now widely recognized as essential dimensions of meaningful education. As learning environments become more complex and socially embedded, affective assessment has emerged as a critical component for capturing students' attitudes, values, emotional regulation, and social conduct alongside cognitive achievement (Khoiruddin & Takhmid, 2024). Empirical evidence suggests that the systematic integration of affective assessment can support well-rounded student development and strengthen social competencies crucial to academic success and civic engagement (Jones et al., 2017). In this sense, affective assessment is no longer supplementary but integral to contemporary educational quality.

Despite growing awareness of its importance, the practical implementation of affective assessment remains deeply problematic. Numerous studies indicate that many schools and educators are not adequately prepared to conduct systematic affective evaluation due to limited instructional time, the absence of explicit and operational assessment tools, and insufficient understanding of observable affective indicators (Andriyati et al., 2025; Charismana et al., 2023). These challenges are further intensified by the expansion of digital-based administrative responsibilities, which increase teachers' workload and reduce opportunities for sustained observational assessment (Kamarullah et al., 2024; Khoiruddin et al., 2024). Consequently, although affective development is frequently articulated as a curricular objective, its assessment often remains implicit, fragmented, or symbolic in classroom practice.

This tension between normative expectations and instructional realities becomes particularly pronounced in educational contexts that place a strong formal emphasis on moral and character education. In such settings, value formation is often assumed to occur naturally through teaching and school culture, while the procedural mechanisms for assessing affective outcomes remain underdeveloped. Early evidence from Aceh Province, Indonesia, clearly illustrates this dilemma. Madrasa teachers operating within a socio-religious educational environment shaped by the implementation of Islamic law frequently report constraints in

monitoring students' affective behavior during routine teaching activities (Manan et al., 2023). These constraints suggest that strong moral discourse does not automatically translate into structured affective assessment practices.

At the policy level, the Indonesian education system has historically prioritized cognitive testing, even though affective development has consistently been included in curricular objectives (Susmiyati et al., 2023; Wijaya et al., 2023). The introduction of the *Kurikulum Merdeka* represents a significant shift by explicitly mandating the assessment of students' spiritual and social attitudes alongside cognitive and psychomotor domains (Wasehudin et al., 2023). Within this framework, teachers are encouraged to employ reflective journals, observation checklists, rubrics, and self- or peer-assessment to capture affective learning outcomes. In the context of Islamic education, this policy direction is further reinforced by the Decree of the Director General of Islamic Education (KDJPI No. 3751/2018), which provides technical guidance for affective assessment in madrasas, including the use of affective notebooks, observation journals, and structured checklists.

Theoretical and empirical scholarship has consistently demonstrated the importance of affective assessment for balanced educational achievement. Studies show that integrating affective dimensions into instructional practices can enhance students' empathy, self-regulation, and social understanding (Green & Batool, 2017; Monita & Ferdi Hasan, 2024). Specific assessment instruments have also been shown to be effective in academic settings. Checklists facilitate the early identification of behavioral and emotional issues and support timely interventions (Birgin & Adnan, 2007; Eldik et al., 2004; Merrell et al., 2008), while rubrics offer systematic standards for assessing self-regulation and social-emotional learning (Panadero & Romero, 2014; Pancorbo et al., 2021). To ensure validity and reliability, affective assessment instruments must be grounded in observable behavior and applied consistently (Glennon et al., 2015). Moreover, affective assessment has been widely employed within program evaluation frameworks, such as the Context, Input, Process, and Product (CIPP) model, to examine the effectiveness of educational policies and interventions (Manan et al., 2020).

Nevertheless, despite the availability of policy frameworks, assessment instruments, and theoretical justification, a critical gap remains insufficiently addressed in the literature. While previous studies have examined teachers' challenges, conceptual models, and policy intentions related to affective assessment, there is a notable lack of empirical research investigating how official affective



assessment guidelines—particularly those articulated in the KDJPI—are actually perceived, interpreted, and enacted by madrasa teachers in everyday classroom practice. This gap is especially salient in Aceh Province, where the formal authority of Islamic law in educational governance creates a distinctive socio-religious environment that may shape teachers' professional judgment, instructional priorities, and assessment practices. Existing studies have yet to adequately explain how teachers negotiate affective assessment demands in this context, or how these practices align with the broader objectives of the *Kurikulum Merdeka*.

To address this gap, the present study offers two key contributions. Substantively, it provides an in-depth empirical examination of affective assessment practices among madrasa teachers in Aceh Province, focusing on how KDJPI guidelines are interpreted, adapted, and negotiated in practice. Methodologically, this study employs Iterative Thematic Inquiry (ITI) as its analytical approach. Unlike conventional thematic analysis, ITI integrates cyclical reflection with continuous refinement of themes, enabling a dynamic understanding of how meanings, beliefs, and practices evolve throughout the research process (Morgan & Nica, 2020). This approach is particularly well-suited to madrasa education, where instructional practices are deeply embedded in cultural, religious, and moral values. The application of ITI thus represents a methodological innovation that enhances analytical rigor while remaining sensitive to local context.

Accordingly, this study aims to identify and explain the dominant themes emerging from madrasa teachers' perceptions and practices regarding affective assessment under the KDJPI framework in Aceh Province, Indonesia. The analysis focuses on how KDJPI guidelines are interpreted, adapted, and enacted in everyday classroom contexts, as well as on the structural and professional conditions shaping these practices. By foregrounding teachers' lived experiences and institutional constraints through iterative thematic inquiry, this study offers empirical insight into the dynamics of affective assessment implementation within a Sharia-based educational system. In doing so, it contributes conceptually to discussions on affective assessment and policy enactment, and empirically to the literature on value-based education in Islamic and broader Asian educational contexts.

## B. Method

This study employed a qualitative research design, using Iterative Thematic Inquiry (ITI) as proposed by Morgan and Nica (2020), to examine madrasa teachers'

perceptions and practices regarding affective assessment within the KDJPI framework in Aceh Province, Indonesia. The study was conducted through a cyclical and iterative process in which data collection and analysis proceeded in tandem, allowing emerging themes to be progressively refined through repeated engagement with empirical material. The overall operational flow of the research followed the ABLE cycle embedded within the ITI approach, guiding the movement between data gathering, analytic reflection, and thematic consolidation throughout the study.

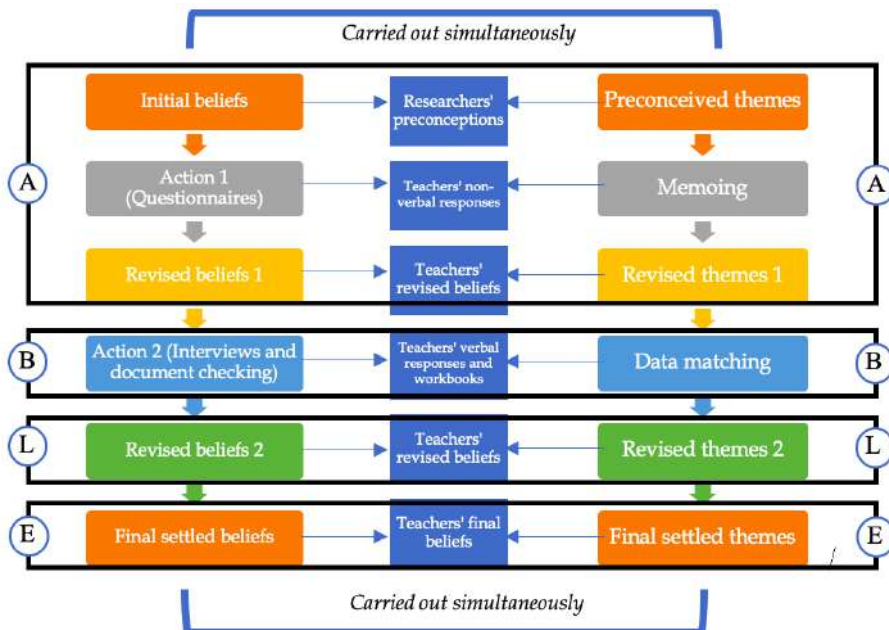


Figure 1. ITI Processes (Source from the authors)

Participants were recruited through purposive sampling to ensure that all respondents had direct and relevant experience with affective assessment practices in madrasas. In total, 224 madrasa teachers participated, comprising 128 teachers from Banda Aceh City and 96 teachers from Aceh Besar Regency. Data collection was conducted during routine *Musyawaharah Guru Mata Pelajaran* (MGMP) meetings, which provided access to teachers actively engaged in subject-based teaching and assessment activities. To deepen understanding of classroom-level practices, six teachers from *Madrasah Aliyah Negeri* (MAN) 3 Kota Banda Aceh were selected for in-depth interviews based on their demonstrated involvement in the implementation of affective assessment. Teachers with less than one year of teaching experience or without prior engagement in affective assessment activities were excluded to ensure the relevance and credibility of the

data. The interviewed teachers represented a range of subject areas, including *akidah-akhlak*, civic education, mathematics, Indonesian language, and social sciences, as summarized in Table 1.

Table 1. Demographics of the informants for interviews

No.	Initials	Subjects	Quantity
1.	HM, HR	<i>Akidah-akhlak</i>	2
2.	RF	Civic Education	1
3.	NM	Mathematics	1
4.	ST	Indonesian	1
5.	FP	Social Sciences	1

Data were collected through questionnaires, semi-structured interviews, and document analysis to enable methodological triangulation. A closed-ended questionnaire was administered to all participating teachers to capture general perceptions of affective assessment practices. The questionnaire items were developed from existing literature, policy documents, and preliminary observations, and were pilot tested with 30 non-sample teachers to ensure clarity and content validity. Semi-structured interviews were conducted with six selected teachers to obtain in-depth insights into classroom-level implementation. Each interview lasted approximately 45–60 minutes and was audio-recorded with participants' consent. Document analysis of lesson plans, assessment records, and reflective journals was undertaken to examine the practical use of affective assessment instruments and to corroborate data across sources.

Data analysis followed an iterative process based on the ITI framework. Questionnaire data were analyzed descriptively to identify general trends, while qualitative data from interviews and documents were coded through successive analytic cycles, beginning with broad thematic identification and followed by focused refinement. Memoing and data matching were employed throughout the analysis to compare emerging themes across data sources and against initial assumptions, in line with the ABLE stages illustrated in Figure 1. The analytic process continued until thematic saturation and internal consistency were achieved.

Methodological rigor was ensured through data triangulation, member checking with selected participants, and peer debriefing among research team members to minimize bias. Ethical clearance was obtained prior to data collection,

and all participants provided informed consent. Anonymity and confidentiality were maintained through de-identification of all records, and the data were securely stored and used exclusively for research purposes.

### **C. Results and Discussion**

This section presents and examines the empirical findings of the study concerning affective assessment practices among madrasa teachers within the KDJPI framework. The presentation is organized to reflect the logical progression from descriptive empirical evidence to interpretive discussion. Accordingly, the section is divided into two parts. The Results subsection reports the key findings derived from questionnaire data, interview narratives, and document analysis, focusing on observable patterns and tendencies without introducing conceptual or theoretical interpretation. The subsequent Discussion subsection situates these findings within broader scholarly debates, critically interpreting their implications for affective assessment, policy enactment, and institutional governance in value-based education contexts.

#### **1. Results**

This is the part where the empirical results from questionnaires, interviews, and document analysis are discussed regarding the practices of affective assessment of madrasa teachers within the framework of the KDJPI. The findings are structured around the main tendencies that emerged from the data, without any conceptual or theoretical explanation.

##### *a. Teachers' initial perceptions of affective assessment instruments*

A questionnaire administered to 224 teachers at madrasas shows that they have fairly positive views of instruments of affective assessment, especially reflective journals and checklists. Table 2 shows that a significant percentage of respondents agreed or strongly agreed that reflective journaling can be used to assess students' spiritual and social attitudes (Q1: 204 respondents) and to affect how teachers judge students' behavior (Q2: 202 respondents). Nevertheless, there was greater splitting in responses to the levels of how reflective journaling increases student engagement in learning (Q4), and a significant percentage of respondents chose the neutral and disagreeing options.

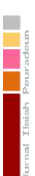


Table 2. Questionnaire Results

Items	Strongly Agree (SA)	Agree (A)	Neutral (N)	Disagree (D)	Strongly Disagree (SD)
Q1	66	138	1	19	0
Q2	61	141	16	6	0
Q3	44	133	31	16	0
Q4	28	80	82	24	10
Q5	36	122	45	20	1
Q6	47	118	37	18	4

Checklist responses were more varied. Whereas a majority of respondents (more than half) agreed that checklists assist in identifying emotional and behavioral problems (Q5) and in providing intervention support (Q6), strong agreement was less common, and the number of neutral responses was also comparatively high. These trends suggest unequal levels of confidence among teachers in the prospective efficacy of the checklist-based affective evaluation. These distributions of responses in reflective journaling and checklists are depicted in Figure 4(a) and 4(b), respectively.



Figure 2. Questionnaire results for (a) reflective journaling; (b) checklists

Figure 2 (a) indicates that responses concerning reflective journaling (Q1-Q3) have a high concentration in the agree and strongly agree categories. In case of Q1, there were 66 strongly agree and 138 agree respondents; there were 19 disagree and 0 strongly disagree respondents. The same pattern of response is seen in Q2 (61 strongly

agree; 141 agree) and Q3 (44 strongly agree; 133 agree), both of which have relatively small proportions of neutral and disagreeing responses. Neutral responses were more pronounced in Q3, where 31 respondents chose neutral, indicating some uncertainty among teachers about this facet of reflective journaling. In all three items, no respondents chose strongly disagree, indicating a generally positive attitude towards reflective journaling as an affective assessment tool.

Also, the distribution of teachers' responses on the checklists shows greater variety across categories, as shown in Figure 4(b). In Q4, the distribution of answers was as follows: agree (80), neutral (82), and disagree (24); however, the number of people who chose strongly agree (28) was lower. Q5 showed a higher level of agreement (36 strongly agree; 122 agree), although neutral (45) and disagree (20) answers were also reflected. The same was observed with Q6: the majority of respondents (118) gave an agree or strongly agree (47) response, with significant numbers neutral (37) and disagree (18). In all the items related to checklists, a few respondents used strongly disagree to indicate that they had a limited outright rejection, rather than different confidence in checklist-based affective assessment.

*b. Documented practices of affective assessment*

The analysis of documents examined how affective assessment practices were formalized and operationalized in teachers' classroom documentation. The documents consulted, as shown in Figure 5(a), consisted of lesson plans, assessment rubrics, affective assessment notebooks, and recapitulation sheets adopted during one academic semester, all compiled into a document known as *buku kerja* (workbook). These documents were examined to assess consistency between the questionnaire and interview reports of teachers regarding the reported practices and the subsequent documented assessment activities.

Generally, the documents showed a skewed and biased record keeping of affective assessment. In some instances, teachers' lesson plans used allusions to the affective areas outlined in the KDJPI guidelines, especially in disciplines like *akidah-akhlak*. These sources, however, were usually general and lacked specific indicators of operation. Affective goals were often stated in general terms (e.g., discipline, responsibility, respect) and were not specifically linked to any assessment tool or observation rubric. As Figure 5(b) shows, affective assessment documentation in teachers' physical records often consists of partial columns or sparse entries,

suggesting that teachers were not guided in the procedures. This trend aligns with other concerns raised by teachers in the interview data about the lack of systematic documentation of affective assessment activities (see E5).

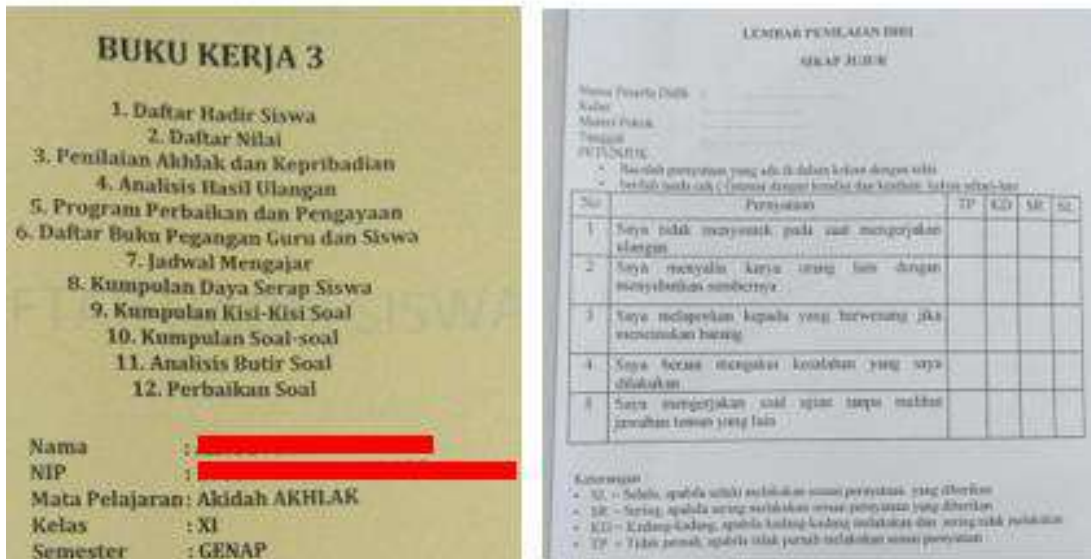


Figure 3. Example of incomplete affective assessment documentation from teachers' physical records (a) workbook cover; (b) students' self-assessment

The records of assessment also showed that they relied on summary judgments rather than continuous records. The brief rubrics or checklists were widely used by teachers, who noted positive or negative behavior without providing a narrative explanation. Figure 6 represents the application of checklist-based affective assessment in practice, in which categorical marks were noted without descriptive observational notes. The yellow arrow denotes good traits of a student, and the red one depicts the bad disposition of a pupil.

Nama Guru : Aqida Aqila  
 Mata Pelajaran : Akidah Akhlak

No.	NIS	Nama Siswa													
			2019	2020	2021	2022	2023	2024	2025	2026	2027				
1	3192	Andi Maulana	1	1	1	1	1	1	1	1	1	1	1	1	1
2	3193	Anggi Humaira	1	1	1	1	1	1	1	1	1	1	1	1	1
3	3194	Asyraf +	1	1	1	1	1	1	1	1	1	1	1	1	1
4	3195	Aulia Saputra +	1	1	1	1	1	1	1	1	1	1	1	1	1
5	3230	Auzan Rachmadani -	1	1	1	1	1	1	1	1	1	1	1	1	1
6	3196	Dewita Rahmiati +	1	1	1	1	1	1	1	1	1	1	1	1	1
7	3197	Khairina Afiyah	1	1	1	1	1	1	1	1	1	1	1	1	1
8	3200	Mardhatillah +	1	1	1	1	1	1	1	1	1	1	1	1	1
9	3202	Muhammad Alif	1	1	1	1	1	1	1	1	1	1	1	1	1
10	3203	Muhammad Farhan M.	1	1	1	1	1	1	1	1	1	1	1	1	1
11	3204	Naina Ursima +	1	1	1	1	1	1	1	1	1	1	1	1	1
12	3205	Nazira	1	1	1	1	1	1	1	1	1	1	1	1	1
13	3206	Nisahul Humera +	1	1	1	1	1	1	1	1	1	1	1	1	1
14	3207	Noufal Ramadan	1	1	1	1	1	1	1	1	1	1	1	1	1
15	3208	Nurul Aini +	1	1	1	1	1	1	1	1	1	1	1	1	1
16	3209	Rike Sukma Wadhoni +	1	1	1	1	1	1	1	1	1	1	1	1	1
17	3210	Rismal Al-Munawar	1	1	1	1	1	1	1	1	1	1	1	1	1
18	3212	Selvia Rahmi +	1	1	1	1	1	1	1	1	1	1	1	1	1
19	3213	Siti Radhiska	1	1	1	1	1	1	1	1	1	1	1	1	1
20	3214	Siti Rukaiyah +	1	1	1	1	1	1	1	1	1	1	1	1	1
21	3215	Sri Hermalinda +	1	1	1	1	1	1	1	1	1	1	1	1	1
22	3216	Venny Arianti +	1	1	1	1	1	1	1	1	1	1	1	1	1
23	3217	Vicha Mauliza +	1	1	1	1	1	1	1	1	1	1	1	1	1

Figure 4. Checklist-based affective assessment with no narrative observation notes

Affective assessment in such documents was frequently transformed into final attitudinal scores keyed in by homeroom teachers and composed of cumulative impressions rather than cumulative observational data. This trend aligns with issues that later arose for teachers in the interview data, given the general and non-procedural nature of affective assessment documentation (see E1).

In several workbooks, parts of the affective assessment were either left blank or not completed, although the teachers indicated in the questionnaires that they are confident in their affective assessment practices. This was actually more evident in the documents on process level assessment since observation journals were either not consistently filled (or not filled at all) during the semester. As discussed in Figure 5, the documentation of affective processes varied widely among teachers, and there was a lack of documentation of students' ongoing behavioral development. These results later justify the interview excerpts E5 and E6, which pointed out discrepancies between the practices supposed to be followed and those recorded.

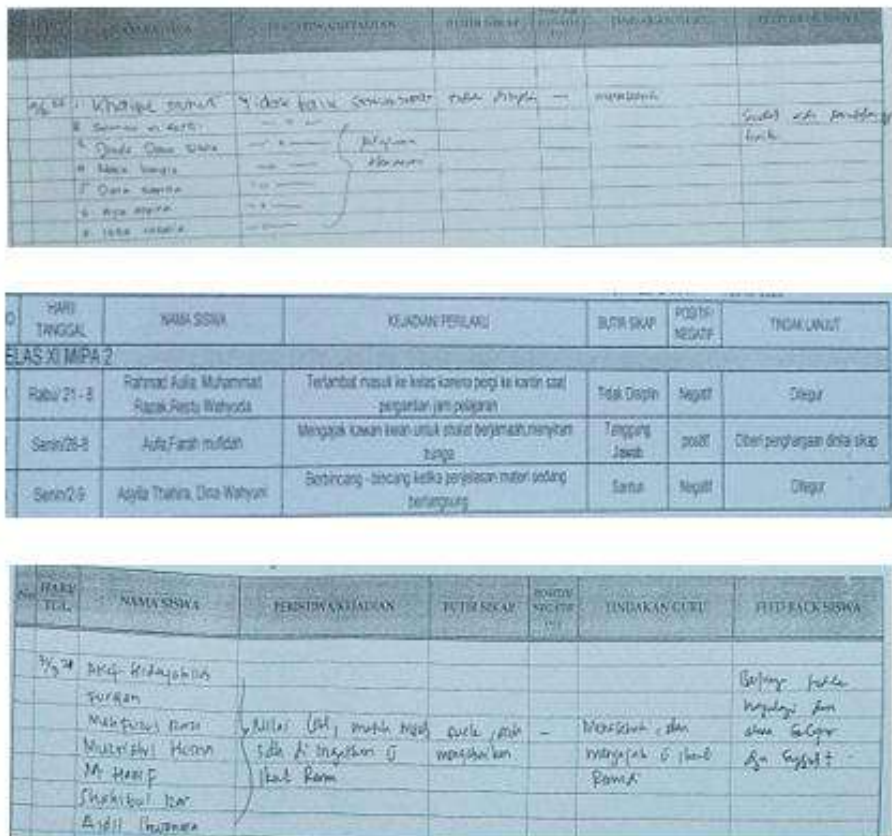


Figure 5. Fragmented documentation of affective assessment across assessment components within the teachers

The documents also varied in the assessment elements provided, aligned with the Context, Input, Process, and Product (CIPP) dimensions, as illustrated in KDJIPI. Although contextual and input-related factors (e.g., alignment with curriculum goals) were more commonly reported, process-oriented affective evaluation (e.g., continuous monitoring of student behavior) was less rigorously documented. The affective outcomes at the product level were usually summative at the end of the semester, which supported a summative and not a formative orientation.

Taken together, the document analysis shows that the practices of affective assessment were indeed present but only fragmentarily institutionalized. The documentation was more inclined toward administrative than pedagogical reflection, leading to disjointed records that reflected the students' affective development only partially. These results triangulate questionnaire patterns and interview materials,

bolstering arguments about the existence of a long-standing policy-practice discrepancy regarding the current state of affective assessment in classrooms.

*c. Interview findings: Beliefs and practical constraints*

The data obtained through interviews also provides more detail on teachers' perceptions and experiences. Some of the teachers admitted that they found affective assessment instruments useful but cited constraints on implementation. In the steps taken by the teachers in accruing the affective assessment of the students, as shown in Figure 6, they referred to it as “*marked generally*” due to time constraints and the lack of procedural guidelines. In the following excerpt 1 (E1), one of the teachers said:

*Reflective journaling is a valuable tool for affective assessments, facilitating students' self-awareness and emotional regulation. It allows us to gain insights into their thoughts and feelings, fostering deeper engagement with learning. However, implementing reflective journaling can be challenging due to limited class time and the large number of students we have to manage, so I sometimes marked generally for the assessed students. (Interview with HR, 2024).*

Similarly, another teacher described checklists as structured but challenging to maintain.

*In my opinion, checklists play a crucial role in affective assessments, enabling educators to identify, monitor, and support students' social-emotional well-being. They provide a structured framework for assessing behavioral and emotional issues, helping us track students' progress and intervene when necessary. However, implementing checklists can be challenging due to the lack of socialization of KDJPI. (Interview with HM, 2024).*

*d. Changes and revisions in teachers' beliefs*

Some of the teachers, during interviews and document review, indicated that they had changed their perceptions of the practice of affective assessment. In a few instances, the teachers initially thought they could use only reflective journaling, but then realized they needed to use a combination of tools. One teacher explained:

*Initially, I believed that reflective journaling was sufficient for assessing students' social-emotional development, but upon reflecting on the data and discussions, I realized the value of incorporating checklists into our practices. Checklists of course provide a systematic approach to identifying and monitoring students' social-emotional well-being, complementing the insights gained from reflective journaling. However, implementing both reflective journaling and checklists is challenging due to limited class time. I will say it is an impossible job because we have other jobs to prepare, mainly for the administrative documents. (Interview with NM, 2024).*

Another teacher expressed a similar shift in perspective:

*I used to think that checklists were secondary to reflective journaling in affective assessments, but after reviewing the findings, I see the importance of integrating both approaches. Reflective journaling allows students to express their thoughts and emotions, while checklists provide a structured framework for assessing their social-emotional well-being. However, implementing both reflective journaling and checklists is challenging due to the limited class sizes or the large number of students we have to handle. (Interview with FP, 2024).*

e. *Alignment and divergence between beliefs and practice*

The data from interviews and documents showed both conformity and nonconformity between the teacher's beliefs and actual practices. Other teachers said they were confident in implementing affective assessment, but their records of application did not indicate much organized implementation. As hinted in the analysis of the documents above and shown in Figure 5, a few teachers admitted that affective assessment activities were regularly undertaken but were not systematically documented. One teacher noted:

*I was under the impression that we were effectively using reflective journaling in our assessments, but the discovery of empty workbooks and unstructured assessments revealed significant gaps in our practices. While we intended to facilitate self-awareness and emotional regulation through reflective journaling, the lack of accountability and oversight has led to inconsistencies in implementation. Moreover, implementing reflective journaling and checklists is challenging due to limited class time. (Interview with ST, 2024).*

Another teacher highlighted discrepancies between intention and execution:

*I was surprised to find discrepancies between what we believed we were doing and what was documented. There's room for improvement in our implementation of both reflective journaling and checklists. Despite our intentions to integrate these tools into our lessons, the absence of clear protocols and training has hindered our efforts. Additionally, implementing both reflective journaling and checklists is challenging due to the lack of socialization of KDJPI. (Interview with RF, 2024).*

f. *Summary of empirical themes*

In data sources, there were three generalizing empirical patterns, the first one is (1) usually positive perceptions of affective assessment tools with varied degrees of confidence, the second one is (2) inconsistent recordings and application of the affective assessment practices, and the last one (3) is a perceived limitation of time allocation, class size and the lack of procedural guidelines. These trends were

also found unanimously in questionnaire responses, interview statements, and document analysis responses.

## **2. Discussion**

This study offers an empirically grounded and analytically consequential understanding of affective assessment practices in madrasa education under the KDJPI framework in Aceh Province. Drawing systematically on questionnaire data, interview excerpts (E1–E6), and documentary analysis, the findings reveal that the persistent gap between policy expectations and classroom practices cannot be sufficiently explained by individual teacher shortcomings, lack of awareness, or resistance to reform. Instead, the evidence indicates that affective assessment operates as a negotiated and structurally constrained practice, shaped by the interaction between policy design, institutional capacity, and teachers' bounded professional agency. Through the application of Iterative Thematic Inquiry, affective assessment is repositioned from a technical pedagogical task into a governance issue embedded within value-based educational reform.

The generally positive orientation of madrasa teachers toward affective assessment instruments, particularly reflective journaling and checklists, reflects a strong normative alignment with the ideals promoted by both KDJPI and *Kurikulum Merdeka*. Questionnaire results demonstrate that teachers broadly accept the pedagogical importance of assessing students' affective development and recognize its relevance to holistic education. This finding is consistent with previous studies showing that Indonesian teachers largely endorse affective and character education as legitimate educational goals (Andriyati et al., 2025; Susmiyati et al., 2023; Wijaya et al., 2023). Crucially, however, this normative alignment does not translate into systematic and sustained classroom practice (Wasehudin et al., 2023). The discrepancy between positive perceptions and inconsistent implementation challenges deficit-oriented explanations that attribute policy failure to teachers' limited competence or motivation. Instead, it suggests that affective assessment policies generate expectations that are not matched by adequate procedural guidance and institutional support.

Interview data illuminate how structural constraints mediate teachers' enactment of affective assessment. Teachers consistently reported limited instructional time, large class sizes, administrative workload, and insufficient policy socialization as key obstacles to systematic implementation. These constraints are not incidental but structural,

reflecting broader conditions of schooling in which teachers are required to meet expanding accountability demands while operating within limited temporal and organizational resources. Prior research has demonstrated that affective assessment requires continuous observation, detailed documentation, and reflective judgment, all of which are difficult to sustain in environments characterized by high workloads and performance-driven evaluation systems (Glennon et al., 2015; Worrell et al., 2002). In this context, teachers' adaptive strategies—such as general marking, selective documentation, or reliance on summative impressions—should be understood not as pedagogical negligence but as pragmatic responses to structurally misaligned policy demands.

The Aceh context intensifies these dynamics in a distinctive way. As a region governed by Sharia-based educational norms, Aceh places strong moral, spiritual, and social expectations on schooling (Manan et al., 2023). In principle, such a value-intensive environment might be expected to strengthen affective assessment practices. Paradoxically, the findings suggest the opposite. Document analysis reveals that affective assessment is frequently treated as an assumed moral outcome rather than as a systematically evaluated domain (Dharin et al., 2025). Affective goals are often articulated in general terms within lesson plans, yet they are rarely linked to explicit indicators, continuous observation, or structured documentation. This pattern underscores a critical distinction between moral education as cultural and normative transmission and affective assessment as a formal evaluative practice. Even in contexts where moral discourse is institutionally dominant, affective development risks becoming informalized and symbolic when procedural mechanisms are weak (Green & Batool, 2017; Dharin et al., 2025).

The triangulation of questionnaire, interview, and document data reinforces this interpretation. Survey responses indicate broad acceptance of affective assessment principles, interviews reveal practical constraints and adaptive strategies, and documents expose inconsistencies and fragmentation in recorded practices. Taken together, these data sources show that affective assessment in madrasas operates within a space of negotiated enactment rather than straightforward implementation. Teachers act as policy interpreters who selectively adapt assessment practices to align with institutional priorities and classroom realities. This finding aligns with policy enactment perspectives that conceptualize implementation as a situated and contingent process, shaped by local conditions and professional judgment rather than linear compliance with regulatory mandates (Rosdiana et al., 2024; Supriyanto et al., 2025).

Differences in teachers' responses to specific affective assessment instruments further illustrate this negotiation process. Reflective journaling was widely perceived as

pedagogically meaningful, particularly for fostering students' self-awareness and emotional regulation. Interview excerpts from teachers such as HR and NM indicate that reflective journaling is valued for its depth and formative potential, echoing findings that reflective practices can enhance affective learning when implemented consistently (Deepa & Mansurali, 2022; Harrison & Fopma-Loy, 2010). At the same time, teachers repeatedly emphasized the time-intensive nature of reflective journaling and its incompatibility with large class sizes and administrative demands. As a result, reflective journaling was often implemented selectively or assessed in a generalized manner, limiting its formative potential.

Checklist-based assessment elicited more ambivalent responses. Questionnaire data show moderate levels of agreement accompanied by relatively high neutrality, suggesting uncertainty rather than rejection. Interview evidence indicates that teachers recognize the practical advantages of checklists for identifying behavioral and emotional issues, consistent with previous research on structured affective assessment tools (Buckley & Epstein, 2004; Merrell et al., 2008; Pancorbo et al., 2021). However, teachers also expressed concern that checklists, when used without narrative observation or reflective follow-up, risk reducing affective assessment to a procedural formality. These differentiated responses highlight that teachers' preferences are shaped less by pedagogical ideology than by assessments of feasibility and institutional support under existing conditions.

A central contribution of this study lies in its reconceptualization of teacher agency within affective assessment reform. Rather than portraying teachers as either compliant implementers or resistant actors, the findings reveal a form of bounded professional agency. Teachers exercise judgment and adapt practices within the constraints imposed by policy design, institutional capacity, and workload demands. They are held responsible for achieving affective outcomes while lacking the procedural resources, time allocation, and sustained training necessary to document and evaluate those outcomes systematically. This asymmetry between responsibility and capacity produces assessment practices that are episodic, fragmented, and often symbolic (Nurwahyuni & Tabrani ZA, 2025; Arif et al., 2024). Similar dynamics have been observed in other educational reforms where accountability expectations exceed institutional support (Carberry et al., 2012; Snead & Freiberg, 2019).

From this perspective, the policy-practice gap identified in this study should not be understood as a failure of teacher professionalism, but as a structural feature of affective assessment governance. Affective assessment policies articulate ambitious

moral and educational objectives without adequately specifying how these objectives can be operationalized within everyday classroom practice. As a result, teachers are compelled to negotiate between normative expectations and practical feasibility, often prioritizing administrative compliance over pedagogical depth. This finding resonates with broader critiques of curriculum and assessment reform in Indonesia, which emphasize that policy effectiveness depends on institutional coherence and capacity building rather than normative aspiration alone (Andriyati et al., 2025; Susmiyati et al., 2023; Wijaya et al., 2023).

To integrate these empirical patterns into a coherent analytical account, this study advances a conceptual framework that positions affective assessment as a negotiated practice shaped by three interrelated dimensions: structural constraints, teachers' orientations toward assessment instruments, and adaptive enactment strategies. Structural constraints, including time limitations, class size, and administrative demands, delimit the range of feasible assessment practices. Teachers' beliefs and orientations influence how they interpret policy expectations and select assessment tools. Adaptive enactment strategies mediate between these dimensions, producing practices that partially align with policy ideals while remaining responsive to contextual realities. By emphasizing interaction and contingency, this framework moves beyond linear models of implementation and offers a transferable lens for examining affective assessment in other value-based education systems.

Figure 8 presents the conceptual framework of this study, illustrating how affective assessment practices in madrasa education are shaped by the interaction of structural constraints, teachers' orientations toward assessment instruments, and adaptive enactment strategies. Structural constraints—such as time limitations, class size, and administrative demands—define the conditions under which affective assessment is conducted, while teachers' perspectives on instruments like reflective journaling and checklists influence how assessment is interpreted and applied in practice. The framework further highlights adaptive enactment as a negotiated response to these conditions, where teachers align with or diverge from policy expectations to maintain practical feasibility. Rather than reflecting resistance, this adaptation represents bounded professional agency within existing governance arrangements. Overall, the figure emphasizes that affective assessment operates as a dynamic and context-dependent practice shaped by institutional design and policy enactment, rather than as a purely technical pedagogical activity.

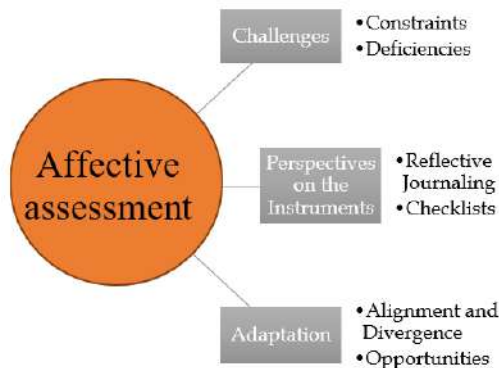


Figure 8. The final themes of affective assessments

The implications of this framework extend beyond the Indonesian context. Comparable challenges have been documented in Islamic education systems in Saudi Arabia and other Muslim-majority contexts, where large class sizes, limited professional development, and administratively driven reforms constrain the implementation of value-oriented assessment (Alhaggass, 2015; Alshamri, 2020; Bahanshal, 2013). The Aceh case contributes to this literature by demonstrating that strong moral governance does not automatically produce robust affective assessment practices. Instead, without procedural clarity, sustained training, and workload alignment, affective assessment risks remaining symbolic regardless of ideological commitment. This insight challenges assumptions that moral or religious contexts inherently facilitate affective education and underscores the importance of institutional design.

At the policy level, the findings highlight the limitations of regulatory mandates that prioritize outcomes without addressing processes. While KDJPI and *Kurikulum Merdeka* articulate comprehensive affective goals, their implementation is undermined by insufficient guidance on assessment procedures, inconsistent professional development, and limited monitoring mechanisms. Interview data repeatedly point to the lack of systematic policy socialization, suggesting that teachers are often left to interpret and operationalize affective assessment requirements independently. This condition reinforces the tendency toward fragmented and summative assessment practices, aligning with prior research on curriculum reform in Indonesia that emphasizes the need for sustained institutional investment (Anshori et al., 2024; Karim et al., 2018; Manan et al., 2020).

Beyond national and regional considerations, this study contributes to global debates on the assessment of non-cognitive learning outcomes (Verger & Skedsmo, 2021).

Across diverse education systems, policymakers increasingly emphasize character education, social-emotional learning, and holistic development, yet struggle to translate these aspirations into sustainable classroom practices (Nurwahyuni & Tabrani ZA, 2025). The Aceh case demonstrates that the central challenge is not cultural resistance or lack of moral commitment, but organizational and governance design (Hay, 2025; Hay, 2025). Even in contexts where affective and moral education are deeply embedded in educational ideology, assessment practices falter when institutional infrastructures lag behind policy ambition (Verger & Skedsmo, 2021). By reframing affective assessment as a governance issue rather than a purely pedagogical technique, this study offers a conceptual intervention relevant to scholars and policymakers concerned with value-based education across global contexts (Hay, 2025).

Methodologically, the use of Iterative Thematic Inquiry strengthens this contribution by enabling systematic cross-referencing of questionnaire trends, interview narratives, and documentary evidence. ITI facilitated iterative engagement with empirical material, capturing shifts in teachers' beliefs and tensions between perception and practice over time (Morgan & Nica, 2020; Supriyanto et al., 2025). Rather than producing a static list of themes, the approach supported the development of an explanatory framework that integrates structural conditions, professional agency, and adaptive practice. This methodological orientation enhances the study's analytical depth and positions it within broader qualitative policy research in education.

Taken together, the findings demonstrate that the primary challenge in implementing affective assessment in madrasa education lies not in teachers' rejection of its values, but in the limited translation of policy mandates into operational and sustainable institutional practices. Affective assessment is enacted through negotiated, adaptive, and often symbolic processes shaped by structural constraints and bounded agency. Without coherent alignment between policy design, institutional capacity, and classroom realities, affective assessment risks remaining an administrative requirement rather than a meaningful pedagogical tool. This conclusion reinforces the need to reconceptualize affective assessment reform as an issue of governance and institutional design, offering implications that extend beyond Aceh to Islamic and value-based education systems globally.

Despite its empirical and analytical contributions, this study has several limitations that should be acknowledged. First, the research is situated within a specific socio-religious and institutional context—namely, madrasa education in Aceh Province under the KDJPI framework—which may limit the direct transferability of the findings to other educational

systems with different governance structures and cultural configurations. Second, while this study employed methodological triangulation through questionnaires, interviews, and document analysis, it relied primarily on teachers' self-reported perceptions and institutional records, which may be influenced by social desirability and administrative compliance pressures. Third, the study focused on teachers' interpretations and enactment of affective assessment policies and did not directly examine students' affective development or learning outcomes, leaving the relationship between assessment practices and student-level impact beyond the scope of this analysis.

Finally, although Iterative Thematic Inquiry enabled a nuanced understanding of negotiated policy enactment, the qualitative design does not allow for causal generalization. These limitations suggest that future research would benefit from comparative, longitudinal, and multi-stakeholder approaches—including student perspectives and classroom observations—to further refine context-sensitive models of affective assessment in value-based education systems.

#### **D. Conclusion**

This study synthesizes evidence on how madrasa teachers in Aceh Province perceive and implement the KDJPI guidelines on affective assessment, revealing a persistent disjunction between policy endorsement and classroom implementation. While teachers broadly recognize the importance of affective assessment as part of holistic education and express favorable orientations toward tools such as reflective journaling and checklists, these commitments are not consistently translated into systematic practice. Affective assessment, as shown in this study, is enacted through negotiated and adaptive processes shaped by structural constraints, including limited instructional time, administrative pressures, and uneven policy socialization. Teachers' adaptive strategies signal professional agency, yet they also underscore that individual initiative alone is insufficient to sustain regular and reliable affective assessment without institutional reinforcement.

In advancing this synthesis, the study contributes academically by empirically mapping the policy–practice gap within a Sharia-based madrasa education system and by positioning affective assessment as a dynamic, context-sensitive practice rather than a purely technical requirement. Methodologically, the use of Iterative Thematic Inquiry strengthens research in Islamic education by enabling an iterative integration of beliefs, practices, and contextual constraints. Substantively, the empirically grounded framework developed in this study offers a conceptual lens to explain how structural



challenges, teachers' orientations toward assessment instruments, and adaptive practices interact to shape assessment outcomes in value-based learning environments, moving beyond fragmented accounts of tools or implementation barriers.

Given the limitations acknowledged in the discussion, future research should broaden the scope of inquiry beyond the specific regional and institutional context examined in this study to test the transferability of the findings across different madrasa systems and governance settings. Given the reliance on self-reported data, subsequent studies would benefit from incorporating additional observational and institutional data sources to reduce potential bias further. Moreover, as this research focused primarily on teachers' perceptions and practices, future investigations should directly examine student affective outcomes to clarify how assessment practices relate to students' moral, emotional, and social development. Comparative, multi-stakeholder, and longitudinal research designs are particularly needed to refine context-sensitive and sustainable models of affective assessment in Islamic education.

Ultimately, this study demonstrates that the primary challenge in implementing affective assessment in madrasas lies not in teachers' rejection of its values and purposes, but in the limited translation of policy directives into operational and sustainable institutional practices. For affective assessment to move beyond a symbolic mandate and function as a meaningful instructional tool, tangible alignment is required between policy design, sustained professional support, and the everyday realities of classroom practice. These findings are relevant not only to the context of Aceh but also to Islamic and value-based education systems more broadly that face similar tensions between normative ideals and educational practice.

### **Declaration of Competing Interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### **Declaration of Generative AI**

The authors declare that artificial intelligence-based tools (Grammarly and QuillBot) were used in a limited, supportive manner during the preparation of this manuscript, primarily to assist with language refinement, clarity, and coherence in academic writing. All substantive content, including research design, data collection, data analysis, interpretation of findings, and scholarly arguments, was conceived,

developed, and validated by the authors. The authors take full responsibility for the accuracy, originality, and integrity of the manuscript.

## **Bibliography**

- Alhaggass, Y. S. (2015). *An investigation of the effectiveness of professional learning activities for physics teachers in Saudi Arabia* [PhD thesis, Victoria University]. <https://vuir.vu.edu.au/32215/>
- Alshamri, K. H. (2020). Challenges and experiences of high school teachers with students having intellectual disabilities in inclusive classrooms in Saudi Arabia. *Universal Journal of Educational Research*, 8(5), 2191–2196. <https://doi.org/10.13189/ujer.2020.080559>
- Andriyati, N., Amin, A., & Li, Y. (2025). Voices from the classroom: Perceptions and obstacles in Indonesian curriculum reform. *Oxford Review of Education*. <https://doi.org/10.1080/03054985.2025.2539397>
- Anshori, I., Majid, M. R., Fitria, S., & Haq, G. F. (2024). The Role of the Cooperative Script Method in Developing Self-Confidence and Public Speaking Skills among Madrasah Ibtidaiyyah Student. *Jurnal Pendidikan Islam*, 10(2), 280–293. <https://doi.org/10.15575/jpi.v10i2.39169>
- Arif, M., Kartiko, A., Rusydi, I., Zamroni, M. A., & Hasan, Moch. S. (2024). The Existence of Madrasah Ibtidaiyyah Based on Pesantren: Challenges and Opportunities in The Digital Era. *Munaddhomah*, 5(4), 367–382. <https://doi.org/10.31538/munaddhomah.v5i4.1401>
- Bahanshal, D. A. (2013). The effect of large classes on English teaching and learning in Saudi secondary schools. *English Language Teaching*, 6(11), 49–59. <https://doi.org/10.5539/elt.v6n11p49>
- Birgin, O., & Adnan, B. (2007). The use of a portfolio to assess student's performance. *Journal of Turkish Science Education*, 4(2), 75–90. <https://files.eric.ed.gov/fulltext/ed504219.pdf>
- Buckley, J. A., & Epstein, M. H. (2004). The behavioral and emotional rating scale-2 (BERS-2): Providing a comprehensive approach to strength-based assessment. *The California School Psychologist*, 9(1), 21–27. <https://doi.org/10.1007/bf03340904>
- Carberry, A. R., Siniawski, M. T., & Dionisio, J. D. N. (2012). Standards-based grading: Preliminary studies to quantify changes in affective and cognitive student behaviors. *2012 Frontiers in Education Conference Proceedings*, 1–5. <https://doi.org/10.1109/fie.2012.6462211>
- Charismana, D. S., Ratri, S. Y., & Firmansyah, F. (2023). A portrait of affective assessment implementation in junior high schools: Assessing its ideality.

Jurnal Civics: Media Kajian Kewarganegaraan, 20(2), 248–256.  
<https://doi.org/10.21831/jc.v20i2.61970>

- Daulay, H. P., & Tobroni, T. (2017). Islamic education in Indonesia: A historical analysis of development and dynamics. *British Journal of Education*, 5(13), 109–126. <https://ejournals.org/bje/vol-5-issue-13-december-2017/>
- Deepa, R., & Mansurali, A. (2022). Reflective journaling to assure learning in the affective domain. *South Asian Journal of Business and Management Cases*, 11(2), 148–166. <https://doi.org/10.1177/22779779221104148>
- Dharin, A., Agus Setiawan, Nugroho, I. S., Sahabuddin, E. S., & Priyanto, D. (2025). Analysis of TPACK-Based Learning for Teachers of Madrasah Tsanawiyah in Banyumas Regency, Indonesia. *Nazhruna: Jurnal Pendidikan Islam*, 8(2), 282–303. <https://doi.org/10.31538/nzh.v8i2.118>
- Eldik, T. van, Treffers, Ph. D. A., Veerman, J. W., & Verhulst, F. C. (2004). Mental health problems of deaf Dutch children as indicated by parents' responses to the child behavior checklist. *American Annals of the Deaf*, 148(5), 390–395. <https://doi.org/10.1353/aad.2004.0002>
- Glennon, W., Hart, A., & Foley, J. T. (2015). Developing effective affective assessment practices. *Journal of Physical Education, Recreation & Dance*, 86(6), 40–44. <https://doi.org/10.1080/07303084.2015.1053635>
- Green, Z. A., & Batool, S. (2017). Emotionalized learning experiences: Tapping into the affective domain. *Evaluation and Program Planning*, 62, 35–48. <https://doi.org/10.1016/j.evalprogplan.2017.02.004>
- Harrison, P. A., & Fopma-Loy, J. L. (2010). Reflective journal prompts: A vehicle for stimulating emotional competence in nursing. *Journal of Nursing Education*, 49(11), 644–652. <https://doi.org/10.3928/01484834-20100730-07>
- Hay, A. (2025). What may be: policy enactment in education, a new conceptual framework with actor-network theory. *Journal of Education Policy*, 40(2), 179–198. <https://doi.org/10.1080/02680939.2024.2411989>
- Hay, A. (2025). What may be: policy enactment in education, a new conceptual framework with actor-network theory. *Journal of Education Policy*, 40(2), 179–198. <https://doi.org/10.1080/02680939.2024.2411989>
- Jones, S. M., Barnes, S. P., Bailey, R., & Doolittle, E. J. (2017). Promoting social and emotional competencies in elementary school. *The Future of Children*, 27(1), 49–72. <https://www.jstor.org/stable/44219021>
- Kamarullah, K., Istiarsyah, I., Maisura, M., Maulya, R., & Bahri, S. (2024). Analysis on Platform Merdeka Mengajar (PMM): Does it affect teachers' workload and well-being? *Innovative: Journal Of Social Science Research*, 4(2), 9407–9421. <https://doi.org/10.31004/innovative.v4i2.14294>

- Kamarullah, K., Nailah, N. A., Akbar, M., Alkhajaj, P., & Rafiqi, M. (2025). Unpacking the disconnect between educational infrastructure and institutional attractiveness. *Jurnal Manajemen Pendidikan (JMP)*, 14, 232–251. <https://doi.org/10.26877/jmp.v14i2.23293>
- Karim, A. A., Abduh, A., Manda, D., & Yunus, M. (2018). The effectivity of authentic assessment based character education evaluation model. *TEM Journal*, 7(3), 495–500. <https://doi.org/10.18421/tem73-04>
- Khoiruddin, M. A., & Takhmid. (2024). Learning Innovation based on seTARA daring in the Community Learning Activity Center. *At-Tadzkir: Islamic Education Journal*, 3(1), 59–69. <https://doi.org/10.59373/attadzkir.v3i1.61>
- Manan, A., Fadhilah, M. A., Kamarullah, & Habiburrahim. (2020). Evaluating paper-based TOEFL preparation program using the context, input, process, and product (CIPP) model. *Studies in English Language and Education*, 7(2), 457–471. <https://doi.org/10.24815/siele.v7i2.16467>
- Manan, A., Salasiyah, C. I., Kamarullah, K., & Tathahira, T. (2023). Educational and social constructions of the Sharia implementation in Aceh, Indonesia. *MIQOT: Jurnal Ilmu-Ilmu Keislaman*, 47(1). <https://doi.org/10.30821/miqot.v47i1.923>
- Merrell, K. W., Juskelis, M. P., Tran, O. K., & Buchanan, R. (2008). Social and emotional learning in the classroom: Evaluation of strong kids and strong teens on students' social-emotional knowledge and symptoms. *Journal of Applied School Psychology*, 24(2), 209–224. <https://doi.org/10.1080/15377900802089981>
- Monita, D., & Ferdi Hasan, M. (2024). The dynamics of affective assessment in MI schools: Challenges and solutions. *Ideguru: Jurnal Karya Ilmiah Guru*, 9(1), 112–119. <https://doi.org/10.51169/ideguru.v9i1.769>
- Morgan, D. L., & Nica, A. (2020). Iterative thematic inquiry: A new method for analyzing qualitative data. *International Journal of Qualitative Methods*, 19, 1. <https://doi.org/10.1177/1609406920955118>
- Muassomah, M., Yaacob, S. E., Khairiah, K., Yurisa, P. R., & Demina, D. (2025). Participatory-based character education: Indonesian school children's experiences. *Jurnal Ilmiah Peuradeun*, 13(3), 1615–1642. <https://doi.org/10.26811/peuradeun.v13i3.1969>
- Nurwahyuni, N., & Tabrani ZA. (2025). Enhancing Teacher Professionalism through the Development of a PositiveSelf-Concept Based on Communication Psychology. *Jurnal Nispatti*, 10(3), 265–284. <https://doi.org/10.26811/nispatti.v10i3.154>
- Panadero, E., & Romero, M. (2014). To rubric or not to rubric? The effects of self-assessment on self-regulation, performance and self-efficacy. *Assessment in Education: Principles, Policy & Practice*, 21(2), 133–148. <https://doi.org/10.1080/0969594x.2013.877872>

- Pancorbo, G., Primi, R., John, O. P., Santos, D., & De Fruyt, F. (2021). Formative assessment of social-emotional skills using rubrics: A review of knowns and unknowns. *Frontiers in Education*, 6. <https://doi.org/10.3389/feduc.2021.687661>
- Rahmiaty, R., & Kamarullah, K. (2024). How far a school program build students' character? A CIPP model evaluation. *Paedagogia: Jurnal Pendidikan*, 13(1), 23–50. <https://doi.org/10.24239/pdg.vol13.iss1.466>
- Rosdiana, A., Khuzaini, Zamrudi, Z., Zainul, M., & Zuana, M. M. M. (2024). The Role of Human Resources Management in Improving Teacher Innovation and Creativity in Madrasahs. *Munaddhomah*, 5(4), 409–420. <https://doi.org/10.31538/munaddhomah.v5i4.1391>
- Setiawan, A., Mardapi, D., Supriyoko, S., & Andrian, D. (2019). The development of instrument for assessing students' affective domain using self- and peer-assessment models. *International Journal of Instruction*, 12(3), 425–438. <https://doi.org/10.29333/iji.2019.12326a>
- Snead, L. O., & Freiberg, H. J. (2019). Rethinking student teacher feedback: Using a self-assessment resource with student teachers. *Journal of Teacher Education*, 70(2), 155–168. <https://doi.org/10.1177/0022487117734535>
- Supriyanto, S., Hadisi, L., Pairin, P., Dirman, D., & Raehang, R. (2025). Facing the Challenges of Globalization: Transforming Madrasah Education from Bilingual to International Classrooms. *Nazhruna: Jurnal Pendidikan Islam*, 8(1), 17–30. <https://doi.org/10.31538/nzh.v8i1.125>
- Susmiyati, S., Zurqoni, Z., Hasan Abdillah, M., & Saugi, W. (2023). Challenges of affective assessment of Islamic religious education learning in Merdeka Curriculum. *Al-Hayat: Journal of Islamic Education*, 7(2), 710–731. <https://doi.org/10.35723/ajie.v7i2.675>
- Verger, A., Skedsmo, G. (2021). Enacting accountabilities in education: exploring new policy contexts and theoretical elaborations. *Educ Asse Eval Acc* 33, 391–401. <https://doi.org/10.1007/s11092-021-09371-x>
- Wasehudin, W., Rohman, A., Nizarudin Wajdi, M. B., & Marwan, M. (2023). Transforming Islamic Education through Merdeka Curriculum in Pesantren. *Jurnal Pendidikan Islam*, 9(2), 255–266. <https://doi.org/10.15575/jpi.v9i2.28918>
- Wijaya, A., Andriani, R. T., Ismady, M. W., Suwarni, W. S. D., & Ahzim, R. (2023). The implementation of affective and psychomotor assessment in Islamic religious education learning. *TARBAWY: Indonesian Journal of Islamic Education*, 10(2), 195–206. <https://doi.org/10.17509/t.v10i2.65467>
- Worrell, V., Evans-Fletcher, C., & Kovar, S. (2002). Assessing the cognitive and affective progress of children. *Journal of Physical Education, Recreation & Dance*, 73(7), 29–34. <https://doi.org/10.1080/07303084.2002.10607844>