

Value-Based Assessment in Islamic Education: Pursuing High Objectivity in Measuring Spiritual and Social Affective Domains

Khadija An Nafisah, Didit Darmawan, Rahayu Mardikaningsih, Solchan Ghozali, Mila Hariani

Universitas Sunan Giri Surabaya, Indonesia

ARTICLE INFO

Article history:

Received 13 November 2024

Revised 23 November 2024

Accepted 05 December 2024

Key words:

Assessment,
Value,
Affective domain,
Spiritual,
Social,
Islamic education,
Objective measurement,
Evaluation.

ABSTRACT

The evaluation of Islamic religious education continues to grapple with the challenge of measuring spiritual and social attitudes in an objective manner. This qualitative literature study examines the effectiveness of value-based assessment approaches in capturing these affective domains within Islamic education settings. By synthesizing scholarly works on educational evaluation and character formation, this research identifies key conditions for successful implementation of value-based assessment. The findings reveal that objective measurement requires clear behavioral indicators, varied data collection methods, reduction of personal bias, and integration of self and peer assessments. Robust school culture, teacher professional development, and alignment among curriculum, instruction, and assessment emerge as essential supporting factors. Technology offers promising opportunities for data management and reporting, yet must be adopted thoughtfully within sound pedagogical frameworks. The study concludes that objective affective assessment is achievable through collaborative, systematic, and culturally sensitive practices. These insights provide theoretical foundations for improving evaluation practices and practical guidance for teachers and policymakers. The transformation of assessment from mere cognitive measurement toward meaningful character development depends on sustained commitment from all educational stakeholders.

INTRODUCTION

Islamic Religious Education (PAI) in public schools and universities in Indonesia faces a major challenge in ensuring that the teaching and learning process does not stop merely at the acquisition of cognitive knowledge. Over the past few decades, PAI learning evaluation systems have focused heavily on aspects of memorization, understanding religious texts, and students' ability to answer objective questions regarding the pillars of faith, the pillars of Islam, or the history of Islamic civilization (Fitri et al., 2022). Such evaluation practices yield quantitatively measurable data, yet the question of whether this knowledge genuinely shapes students' daily character and behavior frequently escapes attention. PAI teachers report that students with high religious exam scores do not necessarily exhibit better attitudes of tolerance, honesty, or social care compared to students with average scores (Mukmin & Nuraini, 2024). This indicates a gap between academic achievement and the reality of character

formation, which constitutes the primary objective of religious education. The evaluations taking place thus far appear to remain comfortably within the safe zone of cognitive measurement without daring to touch the complex and subjectivity-laden affective domain, with the result that PAI learning outcome reports often fail to provide a complete picture of achieving the goal of shaping Muslim individuals with noble morals.

PAI teachers encounter practical difficulties when assessing changes in students' spiritual and social attitudes because both domains are dynamic and heavily influenced by the students' living environments (Musthofa, 2023). The process of forming and observing student character within the school environment requires support from school culture as well as parental parenting styles (Dena & Darmawan, 2024). Spiritual behaviors such as solemnity in worship, persistence in reading the Quran, or the awareness of prayer represent personal expressions that are difficult to observe consistently

* Corresponding author, email address: dr.diditdarmawan@gmail.com

in the classroom. The success of forming and evaluating these spiritual attitudes depends heavily on the competence of PAI teachers and the active attention of parents (Firmansyah & Darmawan, 2023). Similarly, social attitudes such as the willingness to help peers, the ability to control anger, and sensitivity toward the conditions of those around require long observation periods to draw valid conclusions. The support of a conducive school environment further influences the formation of independent attitudes and student learning achievements (Haqiqi & Darmawan, 2023). The combination of curriculum demands requiring numerical assessments and field realities demanding fairness creates a distinct dilemma. Available measurement tools, such as simple observation sheets and teachers' daily journals, are considered insufficiently sensitive to capture gradations of behavioral change occurring gradually. The utilization of supporting facilities and the encouragement of independent learning constitute important strategies for enhancing evaluation effectiveness (Imanuddin & Darmawan, 2024). Teachers are also burdened by large student numbers, rendering individual observation ineffective and tending toward generalization. Meanwhile, parents and the community demand school accountability in shaping children's morals without realizing that the evaluation tools employed have not been designed to provide empirical evidence of affective changes objectively. These technical barriers are exacerbated by the minimal training teachers receive in developing adequate and scientifically valid affective assessment instruments.

Value-based assessment emerges as an alternative approach that offers a framework to measure learning achievement by referring to the fundamental values intended to be instilled in students (Ngwachio, 2024). This approach does not merely demand that students know these values textually, but rather encourages their internalization and manifestation in real behavior. In Islamic Religious Education (PAI), these values encompass theological aspects such as monotheism, worship, and morals toward Allah, as well as social aspects such as brotherhood, tolerance, and social responsibility. The application of Islamic values serves as an essential foundation in directing the development of students' potential and moral formation (Darmawan, 2021). This concept demands a paradigm shift in evaluation from merely measuring learning outcomes to a process that continuously monitors the development of values within students (Ajir et al., 2024). Every educational

interaction, learning task, and reflection moment can be utilized as an opportunity to assess the extent to which these values are embedded in students' consciousness. Directed and continuous education is capable of driving students' behavioral changes toward more valuable and responsible directions (Gautama & Mardikaningsih, 2022). PAI teachers need to design assessment strategies that emphasize the process of value development taking place over time, so that evaluation becomes more meaningful and capable of capturing the dynamics of students' behavioral changes. The challenge then shifts to the effort of formulating objective indicators to measure abstract and multidimensional affective domains.

Objectivity in measuring affective and social domains has become a central issue triggering debate among academics and religious education practitioners. Objectivity is frequently associated with measurability, repeatability, and freedom from evaluator bias, whereas the affective domain is deeply personal, situational, and value-laden (Green et al., 2019). Under such conditions, PAI teachers must find a middle ground between the demands of standardized measurement and the affective peculiarities that cannot be reduced merely to numbers. Attempting to measure piety through multiple-choice tests is clearly impossible; rather, it requires a more complex and multidimensional approach. Several practitioners utilize attitude scales, behavioral inventories, or peer assessments to approach objectivity, yet such instruments still possess validity and reliability weaknesses if employed hurriedly without systematic development. Teachers also frequently lack the time and expertise to develop instruments amid administrative burdens. It must be acknowledged that objectivity is not an ultimate goal, but rather a means to gain trust in evaluation results so they can be used as a basis for making appropriate educational decision-making.

The development of educational evaluation theory in recent years has shown a strong tendency toward authentic approaches capable of capturing the complexity of student learning more comprehensively (Barber, 2023). Authentic assessment emphasizes observing students' abilities to complete meaningful tasks relevant to real life, rather than merely answering constructed paper-and-pencil questions. This assessment can be manifested through observing student participation in religious activities within the community, written reflections on spiritual experiences, or evaluations of social projects involving collaboration among students. This approach opens up space to evaluate

affective and social domains more naturally and contextually, while still demanding clear criteria and standardized procedures. Student involvement through self-assessment and peer assessment is also viewed as effective in increasing students' awareness of the values being learned. Social education plays a strategic role in building students' social and moral value awareness more broadly (Hariani & Mardikaningsih, 2022). Teachers act as facilitators who help students reflect on their value development while providing constructive feedback. This authentic approach aligns with the principles of value-based evaluation that prioritize process and product in a balanced manner. However, its implementation is frequently hindered by limitations in time, cost, and teachers' ability to design and manage complex assessment processes.

The shift from conventional evaluation to value-based evaluation requires a fundamental change in teachers' perspectives regarding the essence of religious education and the purpose of its measurement (Maspuroh et al., 2024). Religious education has often been understood as the transmission of religious knowledge from teacher to student, leading evaluation to be directed toward measuring how much knowledge has been transferred. This understanding ignores the fact that religious education is essentially an effort to shape whole human beings, encompassing cognitive, affective, and psychomotor aspects in an integrated manner. The protection of children's rights to obtain quality value education serves as an important foundation in the learning process (Hariani et al., 2021). PAI teachers accustomed to multiple-choice questions will feel unfamiliar when asked to assess students' humility (tawadhu) or honesty in daily life. This discomfort is exacerbated by a learning outcome reporting system that still prioritizes numbers and letters without providing adequate formats to report affective achievements. Schools that have attempted to apply affective assessments often struggle to formulate qualitative descriptions that are accurate and meaningful for parents, rendering affective achievement reports generic and uninformative. These issues demand a reconceptualization of PAI evaluation capable of bridging accountability demands and the need for comprehensive information regarding student value development. PAI teachers must be encouraged to view evaluation not as a burdensome administrative activity, but rather as a reflective instrument that enriches teaching practices and student character development.

One of the primary issues hindering the

objectivity of affective assessment in Islamic Religious Education (PAI) is the lack of behavioral indicator formulations that can be observed and measured consistently. The indicators available in various assessment guides are frequently very general, such as "students can demonstrate an attitude of honesty" or "students can exhibit tolerant behavior," without being accompanied by concrete descriptions of the meanings of honesty and tolerance in various situations. This ambiguity leads to differing interpretations among teachers when assessing the exact same student behavior, rendering assessment reliability low and indefensible. One teacher might view silence as a form of tolerance, whereas another teacher sees it as an indication of apathy. This difference in perception becomes even more problematic when assessment results are used to determine graduation, scholarship awards, or recommendations for the next educational level (Charismana et al., 2023). The integration of technological facilities and teachers' digital skills is also required to support the effectiveness and accessibility of learning assessment processes (Arifin & Darmawan, 2021). The absence of clear and agreed-upon standards makes affective assessment highly vulnerable to subjectivity as well as personal and cultural biases. Students from specific cultural backgrounds may be judged as less religious simply because their religious expressions differ from teacher expectations, further complicating efforts to produce fair and objective assessments, which ultimately disadvantages students and diminishes the credibility of religious education.

Another issue is the mismatch between the objectives of PAI education, which emphasize character formation, and the evaluation system, which remains dominated by objective written tests. The PAI curriculum explicitly states that the ultimate goal of religious education is the realization of the *insan kamil*—the perfect human being in faith, knowledge, and deeds. In practice, however, evaluation is directed largely toward measuring the cognitive domain due to reasons of practicality and administrative ease. PAI exam questions tend to test the memorization of scriptural verses, knowledge of the pillars and procedures of worship, and comprehension of prophetic stories (Ajir et al., 2024). Students capable of answering such questions well are deemed successful, whereas students who are less proficient in memorization yet display good moral behavior end up receiving low scores. This imbalance creates a major distortion in the interpretation of PAI learning outcomes and causes the loss of critical information regarding students'

moral development. Such evaluation is clearly misaligned with the mandate of the national education system law, which positions character building as one of the primary pillars of educational administration. PAI teachers find themselves in a tight spot between ideal curriculum demands and the reality of practical assessment systems, frequently choosing the safe path by continuing to prioritize cognitive assessment. This condition threatens the existence of PAI as a subject that holds a significant contribution to shaping a national character with noble morals.

This study aims to analyze the effectiveness of value-based PAI learning evaluation in objectively measuring spiritual and social affective domains, with an emphasis on identifying the characteristics of assessment instruments that meet validity and reliability standards. This objective is elaborated through systematic efforts to examine various value-based assessment models and approaches that have been implemented in relation to religious education, as well as to evaluate the strengths and weaknesses of each approach based on formulated objectivity criteria. The theoretical contribution of this study lies in the development of a conceptual framework regarding objectivity in affective assessment, which has long remained a gray area in educational evaluation theory. Its practical contribution consists of providing a guide for PAI teachers in designing and implementing value-based assessments that are effective, valid, and scientifically accountable.

RESEARCH METHOD

This study employs a qualitative approach with a literature review design, as defined by Snyder (2019) as a systematic research methodology to collect, evaluate, and synthesize existing knowledge within a specific field. The selection of the literature review is based on the consideration that the topic of the effectiveness of value-based evaluation in measuring the affective domain requires a broad and deep exploration of various theories, concepts, and empirical findings scattered across diverse scholarly sources. Qualitative research allows researchers to understand social phenomena from a rich and complex perspective, as well as capture hidden meanings behind evaluation practices occurring in the field (Kasinath, 2016). Although this study does not collect primary data through interviews or observations, its qualitative nature is maintained through the process of interpretation and critical analysis of the scholarly texts serving as primary data sources. The literature review approach in qualitative research demands that the researcher act

as an active reader capable of synthesizing, evaluating, and critiquing the arguments put forward by experts in the reviewed literature. Consequently, this study does not merely summarize existing knowledge, but rather endeavors to construct a new understanding of the objectivity of affective assessment through the integration of various relevant theoretical perspectives and empirical findings.

The research procedure follows the guidelines proposed by Snyder (2019), which encompass four main stages: formulating research questions, systematically searching the literature, filtering and assessing source quality, and analyzing and synthesizing information. The research questions formulated in the preliminary stage served as a guide to determine the inclusion and exclusion criteria for the literature sources to be reviewed. The literature search was conducted through prominent academic databases such as Google Scholar using a combination of keywords relevant to the research topic. The selected sources were then filtered based on strict quality criteria, such as journal accreditation levels, author reputation, and the research methodology utilized in the encompassed empirical studies. Data analysis was carried out through thematic analysis techniques, allowing the researcher to identify patterns, themes, and relationships among concepts emerging from the reviewed literature (Ibrahim, 2019). The validity of the research findings was enhanced through theoretical triangulation, which involves comparing and contrasting arguments from various different sources to obtain a more comprehensive and reliable conclusion. By following a systematic and transparent procedure, this study is expected to produce a credible analysis regarding the effectiveness of value-based evaluation in measuring the affective domain objectively.

RESULT AND DISCUSSION

The effectiveness of value-based Islamic Religious Education (PAI) learning evaluation in objectively measuring spiritual and social affective domains depends on a number of key factors that interact with one another throughout the assessment process. The first factor is the clarity and operationalization of the values to be measured, as values such as honesty, tolerance, piety, and social care require specific, observable behavioral formulations (Syahriza et al., 2023). Without clear behavioral formulations, teachers will struggle to distinguish between students who genuinely internalize values and those who merely feign

specific behaviors during assessment. Research indicates that the most effective affective assessment instruments are those developed through a participatory process involving teachers, students, and even parents. The involvement of various parties in formulating behavioral indicators helps reduce cultural bias and enhances shared understanding regarding the meaning of the values being assessed. This participatory process enables assessment indicators to be tailored to the distinct social and cultural environments in which students reside, ensuring that measurement objectivity is not disrupted by differing perceptions of the evaluated behaviors. Consequently, the development of operational and participatory behavioral indicators serves as an absolute prerequisite for the effectiveness of value-based assessment.

Construct validity constitutes a crucial benchmark in determining whether a value-based assessment instrument genuinely measures the intended affective domain or instead assesses entirely different variables. An affective assessment instrument claimed to measure spiritual values must genuinely capture the essence of spirituality, rather than merely the frequency of worship activities or knowledge of religious rituals (Musthofa, 2023). Therefore, instrument validation requires careful pilot testing on representative populations, alongside adequate statistical analysis to ensure that questionnaire items perform in accordance with the underlying theoretical construct. Nonetheless, within the qualitative research underpinning this writing, construct validity is examined primarily through logical and conceptual approaches rather than empirical statistical testing. The logical approach requires researchers to critically examine the alignment between the operational definitions of values and the formulated behavioral indicators, ensuring that these indicators encompass all essential aspects of the measured construct. Guaranteed construct validity enhances confidence in assessment results and enables accurate interpretations of students' affective development.

The reliability of value-based affective assessment faces greater challenges compared to cognitive assessment due to the subjective involvement of human raters in the observation and evaluation process. Two teachers observing identical student behaviors can yield differing evaluations due to their respective background, experience, and expectation differences (Okur et al., 2018). To overcome this issue, various efforts are undertaken to enhance inter-rater reliability through intensive training for evaluators, the drafting of detailed

observation guidelines, and the use of audiovisual recordings as calibration material. Rater training aims to align perceptions and assessment standards among teachers, ensuring that any identical student behavior is evaluated in a relatively uniform manner. The use of detailed observation guidelines helps reduce the subjectivity factor originating from teachers' personal interpretations of student behavior. Calibration through audiovisual recordings allows teachers to discuss and agree upon evaluations of specific behavioral examples, minimizing assessment discrepancies. Consequently, the reliability of affective assessment is not impossible to achieve, provided there is commitment and systematic effort from all participating parties.

Value-based assessment demands an adequate variety of data collection methods to capture the complexity of multidimensional affective and social domains. A single method, such as an attitude scale questionnaire alone, is insufficient to provide a complete picture of student value development, as questionnaires capture only what students say about themselves rather than what they actually do in real life (Musthofa, 2023). A combination of participatory observation, reflective interviews, peer assessment, and behavioral portfolios can complement one another and overcome the individual weaknesses of each method. The use of task assignment methods and the improvement of social interactions in the classroom are proven capable of enhancing students' learning activity (Rahmawati & Darmawan, 2024). Participatory observation allows teachers to observe student behavior in natural settings without making students feel they are being evaluated. Reflective interviews provide opportunities for students to explain the reasoning and motivation behind their behaviors, which often remain invisible to external observation. Peer assessment offers perspectives from classmates who interact with the student daily and may observe behaviors unperceived by the teacher. Behavioral portfolios document concrete evidence of student actions reflecting specific values, thereby strengthening the evidentiary basis for teacher evaluations. This methodological variety collectively enhances the validity and reliability of affective assessment (Allal, 2002).

The measurement of the spiritual domain in value-based assessment presents unique challenges because spirituality is deeply personal, private, and difficult to express in directly observable behaviors. An individual's religious beliefs and internalization occur within an inner realm that can only be accessed through self-disclosure and self-reflection, which

naturally depends heavily on the individual's honesty and self-awareness (Büssing, 2019). The neurological understanding within Islamic education provides explanations regarding the process of spiritual value internalization that influences memorization and worship obedience (Samsiroh et al., 2024). PAI teachers cannot claim to know with certainty a student's level of piety merely from observing habitual prayer or voluntary fasting, as outward behaviors may be driven by motives outside of piety, such as parental compliance or social pressure (Alfian et al., 2022). Spiritual assessment requires a more careful and indirect approach that emphasizes reflection and dialogue rather than mere observational scoring. Regular spiritual reflection journals written by students can serve as a medium for them to express their inner struggles and the development of their appreciation for theological values. Students' moral development is also influenced by academic religious understanding and the level of social media usage in daily life (Al Mursyidi & Darmawan, 2023). Individual counseling sessions between teachers and students likewise provide a safe space for students to discuss their spiritual concerns without fear of being judged or condemned. Through such an approach, spiritual assessment is no longer understood as an act of judgment, but rather as a mentoring process that supports students' spiritual growth.

The measurement of the social domain in value-based assessment is equally complex because social behavior is heavily influenced by situational contexts, group norms, and constantly changing interpersonal dynamics (Civai & Sanfey, 2021). Social stereotypical views within society often limit opportunities and create inequalities in interactions among individuals (Sajjapong et al., 2022). Students who behave tolerantly inside the classroom do not necessarily exhibit the same tolerance when outside, especially if peer group pressure tends toward intolerant behavior. Social assessment must be conducted across various situations and over a sufficiently long period to obtain valid conclusions regarding students' social behavior tendencies. Strengthening civic education is essential for enhancing the legal awareness and social responsibility of school members (Rojak, 2021). Observations of students' participation in community social activities such as community service, social drives, or joint religious events provide more meaningful information about students' social values than mere classroom observation. Non-formal education holds a strategic role in building inclusive social progress within the

community environment (Warin, 2022). Project-based assessments that require students to collaborate with others in solving real social problems can also serve as effective instruments for evaluating cooperation, empathy, and social responsibility. Through such projects, students are placed in situations that demand them to manifest social values tangibly, while teachers can observe and assess their behavior in action. The social assessment results obtained through these various methods and sources will be more objective and accountable compared to evaluations relying on a single information source.

Student involvement in the evaluation process through self-assessment and peer assessment constitutes an essential characteristic of an effective value-based assessment approach (LeWitt, 2023). Self-assessment encourages students to reflect on their own behavior, evaluate its alignment with the taught values, and identify areas needing improvement. Learning discipline and self-regulation skills serve as determining factors in students' academic achievement (Maharani & Darmawan, 2024). This reflection process is invaluable because it helps students develop self-awareness and personal responsibility regarding their value development. Peer assessment trains students to provide fair and constructive evaluations of others while simultaneously fostering social sensitivity and the ability to accept feedback from peers. Both types of assessment reduce reliance on the teacher as the sole source of evaluation and distribute assessment responsibility across the entire learning community. Nevertheless, self-assessment and peer assessment require adequate training so that students understand evaluation criteria, avoid bias, and are capable of providing accurate evaluations. Without sufficient training, self-assessments tend to be overly high or low, while peer assessments can be influenced by popularity or friendship factors. Teachers must allocate sufficient time to train students in assessment skills while providing clear guidelines and concrete examples of proper and correct assessment.

Information and communication technology opens new opportunities to enhance the objectivity of value-based affective assessment through the development of digital platforms that support the efficient and transparent collection, processing, and reporting of assessment data (Carmo, 2023). Gadget-based assessment applications can be utilized to record student behavior in real time, reducing the risk of data loss and improving the accuracy of teachers' observation logs. Digital literacy skills and

the utilization of interactive learning media make a tangible contribution to student learning achievement (Rizal & Darmawan, 2024). Digital platforms can also provide automated analytical features that help teachers identify student behavioral patterns over time, ensuring continuous assessment throughout the year. The use of technology further enables collaboration among teachers in assessing the same student, allowing assessment data from various sources to be combined and analyzed collectively. Educational media innovations and varied teaching methods are proven capable of arousing students' learning interest in schools (Laili & Darmawan, 2024). Additionally, technology facilitates routine and transparent communication between schools and parents regarding students' affective development, which subsequently enhances public accountability and trust in school-based evaluations. Nonetheless, the adoption of technology in affective assessment must not be rushed without considering data security aspects and the ethics of utilizing technology in student character assessment. The utilization of technology must remain grounded in strong pedagogical principles and should not replace the vital role of human interaction in the affective assessment process.

The reporting system for value-based affective assessment requires a format different from cognitive assessment reporting so that the information conveyed remains meaningful and useful for all stakeholders. Numbers and letters, which serve as the standard for cognitive reporting, are insufficient to depict students' value development because the affective domain is qualitative and multidimensional (Ajjawi & Boud, 2023). Descriptive reporting formats that narratively describe student behavior, demonstrated value strengths, and areas still needing development are more appropriate for affective assessment. Such narratives provide a richer and more personal picture of students, enabling parents to clearly understand their children's character development. Digital portfolios containing collections of behavioral evidence, student reflections, and teacher notes can also be utilized to complement descriptive reports, providing concrete evidence supporting the given evaluations. Regular parent-teacher meetings also need to be utilized as a forum to discuss students' affective development in greater depth and build cooperation between the school and the family. Parental attention and school environment conditions are proven to have a close

correlation with student achievement (Umroh & Darmawan, 2024). Effective reporting transforms evaluation from a result-oriented activity into an activity oriented toward the continuous process of value development.

Effective value-based assessment cannot be separated from the construction of a school culture that supports the internalization of values intended to be instilled in students (Erdogan et al., 2018). A positive school culture, in which noble values are lived and practiced by all school members, creates an environment conducive to students' natural and continuous affective development. Institutional leadership styles play a crucial role in shaping the discipline culture of school members amidst social and technological changes (Al Laisty et al., 2024). In schools with a strong culture, evaluation is no longer a separate activity conducted at the end of learning, but rather becomes an integrated part of all school activities. Teachers do not need to struggle to find evidence of student behavior, as such behavior naturally emerges in daily interactions between students and the school environment. School environmental conditions and teacher competence play an important role in maintaining student learning motivation (Latif & Darmawan, 2024). Principals and administrative staff also demonstrate role modeling in executing the taught values, ensuring students see that these values are not merely theories taught in classrooms but are genuinely practiced by adults in their environment. A strong school culture also encourages students to remind and correct peers who exhibit deviant behavior, meaning supervision and evaluation responsibilities do not rest solely on teachers. Consequently, value-based assessments conducted in schools with conducive cultures will be easier, more accurate, and more meaningful for all parties.

The professional development of Islamic Religious Education (PAI) teachers in the field of affective assessment serves as a determining factor in the successful implementation of value-based assessment, as teachers are the primary actors responsible for the quality of evaluations conducted in the classroom (Erdogan et al., 2018). Teacher professionalism and discipline serve as an essential foundation in enhancing the quality and performance of educational institutions (Mubasysyir & Darmawan, 2024). Many PAI teachers have yet to receive adequate training on how to design, execute, and report valid and reliable affective assessments (Susmiyati et al., 2023). The curricula of religious education study programs in higher education institutions also frequently fail to allocate sufficient

weight to learning evaluation courses that specifically address the measurement of the affective domain. This condition is compounded by educational access disparities that remain a barrier to improving learning quality across various regions (Rojak & Khayru, 2022). Consequently, PAI teachers enter their profession with limited knowledge regarding affective assessment and tend to replicate the assessment practices they experienced as students. Teacher teaching styles and student learning motivation levels represent primary variables influencing learning outcomes (Mudzakkir & Darmawan, 2024). Enhancing student learning interest can also be stimulated through the alignment of teacher teaching styles and parental support (Safitri & Darmawan, 2023). Continuous professional development programs, such as affective assessment workshops, teacher learning communities, and expert mentoring, are urgently needed to elevate teacher competence in this field. The development of these technical and social competencies is required to prepare educators for an increasingly digitally based work environment (Mendonca et al., 2021). Professional development must also encompass training on qualitative data analysis and the use of assessment technology, as both skills become increasingly vital in modern value-based assessment. With competent and confident teachers in affective assessment, the quality of value-based evaluation will improve, and its results will gain greater trust from students, parents, and the community.

The objectivity of value-based affective assessment is frequently criticized due to the difficulty of freeing evaluation from the value loads and personal biases of evaluators, even though objectivity constitutes a crucial pillar in credible educational evaluation (Virk et al., 2020). Every teacher enters the classroom with a unique cultural background, life experience, and value preference that unconsciously influences how they perceive and assess student behavior. A teacher originating from a culture emphasizing politeness might judge a quiet student as more courteous, whereas a teacher from a culture valuing openness might view the exact same student as lacking self-confidence. To mitigate this bias, high awareness among teachers regarding their potential biases is required, along with systematic efforts to control them through the use of standardized instruments and clear assessment rubrics. Inter-teacher discussions regarding difficult assessment cases can also help neutralize individual bias and generate consensus regarding more objective evaluations. The use of multiple raters and the calculation of inter-rater

score averages have proven effective in enhancing reliability and reducing bias impact across various studies. Although perfect objectivity may never be attained in affective assessment, continuous efforts to approach objectivity will enhance assessment quality and fairness.

The synergy between formative and summative assessment in a value-based approach enables teachers to continuously monitor student progress and provide timely feedback for behavioral improvement. Formative assessment is conducted continuously throughout the learning process, allowing teachers to immediately identify when a student exhibits deviant behavior and promptly provide guidance. Formative assessment also provides opportunities for students to self-correct before summative assessment is carried out, prioritizing the purpose of assessment as a developmental tool over its purpose as an instrument of judgment (Khoirunnisak et al., 2024). Summative assessment is conducted at the end of a learning period to provide a comprehensive picture of students' final achievements and to determine decisions regarding graduation or grade promotion. The combination of these two assessment types yields a more complete overview of students' affective development compared to relying on a single type alone. Teachers also need to clearly communicate to students the differences between formative and summative assessments, explaining that formative assessment serves as an opportunity to learn and self-correct rather than to be judged. Students will be more motivated to develop positive values when they understand that the assessment process is designed to support their growth, not to punish or humiliate.

The alignment between PAI curriculum objectives, the learning process, and the value-based assessment system serves as a non-negotiable prerequisite for effectiveness, as all three must move in the same direction. A PAI curriculum emphasizing character formation becomes futile if the learning process and its assessment are not designed to achieve those objectives (Erihadiana & Zakiyyah, 2024). PAI teachers must ensure that every learning experience they design provides opportunities for students to develop and demonstrate the values that constitute the learning goals. Assessment must then capture evidence regarding the extent to which students have successfully internalized and expressed those values across various situations. When strong alignment occurs, students experience learning as a holistic and meaningful journey rather than merely a series of disconnected tasks. This

alignment also helps students understand the relevance between what they learn in the classroom and what they experience in daily life. To achieve this alignment, PAI teachers need to regularly review and revise their lesson and assessment plans based on reflection results and feedback from students and peers. This dynamic alignment process will ensure that value-based assessment remains effective despite changes in curricula, policies, or societal social conditions.

The effectiveness of value-based assessment is also determined by students' readiness to participate actively in the evaluation process and accept assessment results as an honest reflection of themselves (Bayimbetova, 2022). Students with low self-confidence may experience difficulties accepting negative feedback regarding their behavior and tend to become defensive or blame others. Students who are overly confident might ignore constructive feedback and feel that they do not need to self-correct. Teachers need to build a safe and supportive classroom environment where students feel comfortable acknowledging their weaknesses and committing to improvement. The role of the teacher as an empathetic and supportive mentor is crucial in creating such a climate, as students will be more open to assessment if they feel teachers genuinely care about their development. Teachers need to explicitly teach the importance of accepting criticism, adapting, and self-correcting as part of normal human growth. With good psychological readiness, students no longer view assessment as a threat, but rather as a useful tool for self-recognition and potential development. Consequently, the effectiveness of value-based assessment depends not only on the quality of instruments and procedures, but also on the readiness of the evaluated subjects to engage in the process.

The effective implementation of value-based assessment requires systemic support from various stakeholders, including the government, educational institution managers, teachers, parents, and the broader community (Gellel, 2023). The government needs to provide policies supporting the development of affective assessment, such as more detailed assessment standards and dedicated time allocations for affective assessment activities in schools. Educational institution managers need to ensure the availability of adequate resources, such as training funds, time for teacher collaboration, and technical support for assessment instrument development. Parents need to be actively involved in the assessment process and provided with an understanding of the importance of the developed

values as well as how they can support value development at home. The community needs to appreciate schools' efforts in developing student character rather than solely judging school success based on cognitive exam scores. Solid and comprehensive support from all parties will create a conducive assessment ecosystem, enabling teachers to execute value-based assessments more effectively. Without systemic support, teachers will remain fighting alone in developing adequate affective assessments and gaining societal recognition for the importance of such evaluation. The success of value-based assessment relies heavily on a collective commitment to position character building as the primary priority of education.

Critical reflection on current assessment practices indicates that value-based assessment is not an instant solution capable of overcoming all problems in affective evaluation within Islamic Religious Education (PAI). Each assessment approach possesses its respective strengths and limitations, and no single approach is perfect for all situations. Value-based assessment demands a significant investment of time, effort, and resources, meaning not all schools possess the same capacity to implement it. Schools in remote areas with limited teachers and facilities may encounter greater obstacles in adopting this approach compared to schools in urban areas with more abundant resources. Flexibility and adaptation to local conditions are key in integrating the principles of value-based assessment into daily practice. Teachers must be granted the freedom and trust to develop assessment approaches suited to the context and needs of their students, provided that the principles of validity, reliability, and objectivity are maintained. Future research needs to continuously examine the effectiveness of various value-based assessment models under different conditions and generate more specific and contextual recommendations.

CONCLUSION

The effectiveness of value-based Islamic Religious Education (PAI) learning evaluation in objectively measuring spiritual and social affective domains is heavily determined by the quality of assessment instruments, teacher competence, student involvement, and systemic support from the educational environment. The success of value-based assessment cannot be achieved without clear and operational behavioral indicator formulations, a combination of varied data collection methods, and adequate training for teachers to reduce bias in evaluation. Self-assessment and peer assessment

integrated into the evaluation process enhance students' awareness of the values being learned and reduce reliance on teacher evaluations. Information technology can be employed to support the collection and processing of assessment data more efficiently and transparently, although its implementation must remain grounded in strong pedagogical principles. Effective value-based assessment does not stand alone; rather, it forms an integral part of the school culture that supports the comprehensive internalization of values. A positive school culture positions assessment as an activity integrated with all aspects of school life, rather than merely an administrative task at the end of a learning period. Consequently, the objectivity of affective assessment is a goal attainable through collaborative efforts involving all parties, rather than merely the result of a single method or specific instrument.

The implications of this study for educational practice highlight the necessity of enhancing PAI teachers' capacity to design and implement value-based affective assessments through structured and continuous training programs. Educational institutions need to allocate adequate resources for the development of assessment instruments and facilitate collaboration among teachers in formulating consistent assessment standards. Recommendations for future research include employing quantitative and mixed-methods approaches to empirically test the effectiveness of various value-based assessment models across different contexts. Longitudinal research on students' affective development through value-based assessment is also highly necessary to understand the long-term impact of this approach on character building. National education policies should provide broader room for schools to develop innovative and creative affective assessment models, as well as provide more operational technical guidelines for teachers. Schools are also advised to involve parents and the community more actively in the assessment process, as character formation cannot take place effectively solely within the school environment. The development of inter-school networks to share best practices and affective assessment resources can accelerate the improvement of value-based assessment quality throughout Indonesia. By following these recommendations, PAI learning evaluation will shift from a mere cognitive measurement tool into a meaningful and sustainable character development instrument.

REFERENCES

- Ajir, I. C., Pahrudin, A., Jatmiko, A., & Koderi, K. (2024). Model pengembangan kurikulum pendidikan agama Islam dalam pendidikan dasar. *Al-Liqo: Jurnal Pendidikan Islam*, 9(2), 302–314. <https://doi.org/10.46963/alliqo.v9i2.2473>
- Ajjawi, R., & Boud, D. (2023). Changing representations of student achievement: The need for innovation. *Innovations In Education And Teaching International*, 1–11. <https://doi.org/10.1080/14703297.2023.2192513>
- Al Laisty, M. D., Darmawan, D., & Fajar, A. S. M. (2024). The Role of Leadership Style in Building a Discipline Culture in Pesantren: Facing the Challenges of Social and Technological Change. *Bulletin of Science, Technology and Society*, 3(3), 62–68.
- Al Mursyidi, B. M. & D. Darmawan. (2023). The Influence of Academic Success of Islamic Religious Education and Social Media Involvement on Student Morality. *Al-Fikru: Jurnal Ilmiah*, 17(2), 321–331.
- Alfian, A. F., Mujiburrahman, & Sukari. (2022). Upaya guru pendidikan agama islam dalam meningkatkan kesadaran beribadah siswa. 2(2), 227–227. <https://doi.org/10.54090/aujpai.v2i2.2>
- Allal, L. (2002). The Assessment of Learning Dispositions in the Classroom. *Assessment in Education: Principles, Policy & Practice*, 9(1), 55–58. <https://doi.org/10.1080/09695940220119184>
- Arifin, S., & Darmawan, D. (2021). Technology Access and Digital Skills: Bridging the Gaps in Education and Employment Opportunities in the Age of Technology 4.0. *Journal of Social Science Studies*, 1(1), 163–168.
- Barber, W. (2023). Authentic assessment and universal design for learning: exploring shifts in assessment and evaluation. *ICERI Proceedings*. <https://doi.org/10.21125/iceri.2023.1996>
- Bayimbetova, M. (2022). How can be put “learner-oriented assessment” in practice? *Ренессанс в Парадигме Новаций Образования и Технологий в XXI Веке*, 1, 198–199. <https://doi.org/10.47689/innovations-in-edu-vol-iss1-pp198-199>
- Büssing, A. (2019). *Measuring Spirituality and Religiosity in Health Research*. 11–31. https://doi.org/10.1007/978-3-030-21221-6_2
- Carmo, E. G. do. (2023). Information and Communication Technologies in Mental Health: Challenges and Opportunities. *Mental Health & Human Resilience International Journal*, 7(1), 1–2. <https://doi.org/10.23880/mhrij-16000219>
- 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*.
<https://doi.org/10.21831/jc.v20i2.61970>
- Civai, C., & Sanfey, A. G. (2021). *Mentalizing in Value-Based Social Decision-Making: Shaping Expectations and Social Norms*. 503–516.
https://doi.org/10.1007/978-3-030-51890-5_25
- Darmawan, D. (2021). Islamic Human Capital Implementation as Effective Framework in Islamic Human Capital Management. *Journal of Social Science Studies*, 1(1), 239-244.
- Dena, S. & D. Darmawan. (2024). Character Development of Students in Public High School 4 Surabaya Through the Role of School Culture and Parenting Style. *EduInovasi: Journal of Basic Educational Studies*, 4(1), 417–428.
- Erdogan, M., Turkkahraman, M., & Ozbek, G. (2018). *Teachers' assessment of values education implementation*. 8(3), 89–95.
<https://doi.org/10.18844/CERJ.V8I3.623>
- Erihadiana, M., & Zakiyyah, Q. N. (2024). *The Importance of Curriculum Components in PAI Learning*.
<https://doi.org/10.59613/armada.v2i3.2870>
- Firmansyah, B. & D. Darmawan. (2023). The Importance of Islamic Education Teacher Competence and Parental Attention in Enhancing Students' Character Formation at Nur Al-Jadid Excellent Islamic High School. *Jurnal Cahaya Mandalika*, 4(2), 1353-1363.
- Fitri, S. R. A., Majdi, A. L., komaruddin, koko, & Nida, S. F. (2022). *The Issues and Challenges of Teaching Islamic Religious Education at Darussalam High School, Ciamis, Indonesia*. 1(2), 220–233.
<https://doi.org/10.65345/anjis.v1i2.16>
- Gautama, E. C., & Mardikaningsih, R. (2022). Driving Sustainable Behavior Change Through Education and Public Awareness. *Journal of Social Science Studies*, 2(1), 259-264.
- Gellel, A.-M. (2023). Teachers as Key Players in Values Education. *Springer International Handbooks of Education*, 683–697.
https://doi.org/10.1007/978-3-031-24420-9_37
- Green, B., Cheng, A., & Smith, D. I. (2019). The Assessment of Faith and Learning. *Social Science Research Network*.
<https://doi.org/10.2139/SSRN.3486401>
- Haqiqi, M. F. & D. Darmawan. (2023). School Environment and Independence: Effects on Academic Achievement in MTs Nahdlatul Athfal Gersempal Omben Sampang Students. *Kabillah (Journal of Social Community)*, 8(2), 171-180.
- Hariani, M., & Mardikaningsih, R. (2022). The Social Education Role in Shaping Students' Global Awareness in Higher Education. *Journal of Social Science Studies*, 2(1), 55-60.
- Hariani, M., Aliyah, N. D., & Issalillah, F. (2021). Legal Guarantee of Children's Rights in Education and Health. *Journal of Social Science Studies*, 1(2), 177-180.
- Ibrahim, K. (2019). *Thematic analysis*.
<https://doi.org/10.13140/rg.2.2.36275.68649>
- Imanuddin, F. & D. Darmawan. (2024). Enhancing Learning Effectiveness Strategy: Self-Directed Learning and Learning Facilities at SMK Teknik Pal Surabaya. *Jurnal Al-Qayyimah*, 6(1), 99-105.
- Kasinath, H. M. (2016). Understanding and using qualitative methods in performance measurement. *MIER Journal of Educational Studies, Trends and Practices*, 3(1), 46–57.
<https://doi.org/10.52634/MIER/2013/V3/I1/1554>
- Khoirunnisak, A., Aufa, A. A., & Shofiyuddin, A. (2024). *Pengembangan asesmen formatif disertai feedback untuk mata pelajaran pendidikan agama islam*. 5(3), 269–279.
<https://doi.org/10.58401/salimiya.v5i3.1548>
- Laili, N. & D. Darmawan. (2024). Investigating the Impact of Educational Media and Teaching Methods on Student Interest at SMP Buana Waru Sidoarjo. *Jurnal Pendidikan Inovatif*, 6(2), 456-471.
- Latif, A. & D. Darmawan. (2024). Examining How School Environment and Teacher Competence Affect Student Learning Motivation at MA Al Fatih Tambak Osowilangun Surabaya. *Teaching and Learning Journal of Mandalika (Teacher)*, 5(1), 69-75.
- LeWitt, P. A. (2023). *Peer- and self-assessment*. 79–88.
<https://doi.org/10.1016/b978-0-12-818630-5.09021-7>
- Maharani, L. & D. Darmawan. (2024). Factors Affecting Learning Achievement: Learning Discipline and Self-regulation at MTs Wachid Hasyim Surabaya. *TA'DIBUNA: Jurnal Pendidikan Agama Islam*, 7(1), 12-20.
- Maspuroh, M., Khotimah, H., Nurhikmah, I. Y., Mutmainnah, S., & Mubarak, W. (2024). The Meaning of Evaluation of Islamic Religious Education. *Kasyafa*, 1(2), 118–124.
<https://doi.org/10.61166/kasyafa.v1i2.26>
- Mendonca, C. N., Wahyudi, Kabalmay, R. N. K., & Amri, M. W. (2021). Developing Technical and Social Competencies for Future-Ready Education in Digitally Mediated Labor Environments. *Journal of Social Science Studies*, 1(2), 259-266.
- Mohajan, H. K. (2018). Qualitative research

- methodology in social sciences and related subjects. *Journal of Economic Development, Environment and People*, 7(1), 23-48.
- Mubasysyir, M. M., & Darmawan, D. (2024). Improving Performance: The Role of Teacher Professionalism and Discipline at the Tahsinul Akhlaq Bahrul Ulum Foundation Surabaya. *Hikamatzu | Journal of Multidisciplinary*, 1(1), 337-354.
- Mudzakkir, M. & D. Darmawan. (2024). The Influence of Teacher Teaching Styles and Learning Motivation on The Learning Achievement. *Edu-Riligia: Jurnal Kajian Pendidikan Islam dan Keagamaan*, 8(1), 79-91.
- Mukmin, M., & Nuraini, N. (2024). Integrasi Penilaian Tes dan Non-Tes dalam Pendidikan Agama Islam: Menuju Evaluasi Holistik untuk Pembelajaran Berkelanjutan. *Journal of Instructional and Development Researches*, 4(5), 370-379. <https://doi.org/10.53621/jider.v4i5.384>
- Musthofa, M. (2023). *Development of Attitude Competency Assessment Instruments: Guidelines for the Integration of Spiritual and Social Attitudes*. <https://doi.org/10.4108/eai.12-11-2022.2327401>
- Ngwacho, G. A. (2024). *Value-Based Education Incorporation in Competency-Based Curriculum - Recipe for All-inclusive Education for Enhanced Global*. <https://doi.org/10.62049/jkncu.v4i1.56>
- Okur, E., Aslan, S., Alyuz, N., Esme, A. A., & Baker, R. S. (2018). *Role of Socio-cultural Differences in Labeling Students' Affective States*. 367-380. https://doi.org/10.1007/978-3-319-93843-1_27
- Rahmawati, D. & D. Darmawan. (2024). The Relationship Between Assignment Methods and Social Interaction with the Level of Student Learning Activeness at Madrasah Ibtidaiyah Darul Ulum Tandes. *International Journal of Islamic Thought and Humanities*, 3(1), 49-58.
- Rizal, M. I., & Darmawan, D. (2024). Digital Literacy and Utilization of Learning Media: Their Contribution to Academic Achievement in Intensif Taruna Pembangunan High School, Surabaya. *Jurnal Inovasi Pendidikan*, 7(3), 22-30.
- Rojak, J. A. (2021). The Importance of Civic Education to Increase Community Legal Awareness. *Journal of Social Science Studies*, 1(1), 44-50.
- Rojak, J. A., & Khayru, R. K. (2022). Disparities in Access to Education in Developing Countries: Determinants, Impacts, and Solution Strategies. *Journal of Social Science Studies*, 2(1), 31-38.
- Safitri, M. S. & D. Darmawan. (2023). Enhancing Students' Learning Interest: The Role of Teacher's Teaching Style and Parental Support at SD Negeri Wadungasri Waru Sidoarjo. *Jurnal Cahaya Mandalika*, 4(2), 1343-1352.
- Safitri, M. S. & D. Darmawan. (2023). Enhancing Students' Learning Interest: The Role of Teacher's Teaching Style and Parental Support at SD Negeri Wadungasri Waru Sidoarjo. *Jurnal Cahaya Mandalika*, 4(2), 1343-1352.
- Sajjapong, T., Darmawan, D., & Marsal, A. P. (2022). The Role of Social Stereotypes in Shaping Opportunities and Inequalities in Society: Their Impact on Education, Employment, and Intergroup Interactions. *Bulletin of Science, Technology and Society*, 1(1), 44-49.
- Samsiroh, S, N. D. Aliyah, & M. Hariani. (2024). Neuroscience in Islamic Education: A Systematic Review of Implications for Memorization, Devotion, and Faith, *Journal of Social Science Studies* 4(2), 195 - 206.
- Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, 104, 333-339.
- Susmiyati, S., Zurqoni, Z., Abdillah, M. H., & Saugi, W. (2023). *Challenges of Affective Assessment of Islamic Religious Education Learning in Merdeka Curriculum*. <https://doi.org/10.35723/ajie.v7i2.675>
- Syahriza, F. A., Chasanah, I., Kumala, S. D., & Asrori, M. (2023). Analysis of the Evaluation of the Affective Domain of Islamic Education in the Merdeka Curriculum. *Edumaspul : Jurnal Pendidikan*, 7(2), 3561-3568. <https://doi.org/10.33487/edumaspul.v7i2.6986>
- Umroh, U. & D. Darmawan. (2024). The Dynamics of School Environment and the Impact of Parental Attention on the Academic Achievement of Students at SMA Negeri 1 Ketapang Sampang. *Fondatia*, 8(1), 77-87.
- Virk, A., Joshi, A., Mahajan, R., & Singh, T. (2020). The power of subjectivity in competency-based assessment. *Journal of Postgraduate Medicine*, 66(4), 200-205. https://doi.org/10.4103/JPGM.JPGM_591_20
- Warin, A. K. (2022). Reconstructing Community Futures Through Non-Formal Education for Participatory and Inclusive Social Advancement. *Journal of Social Science Studies*, 2(1), 183-188.

* An Nafisah, K., D. Darmawan, R. Mardikaningsih, S. Ghozali, M. Hariani. (2025). Value-Based Assessment in Islamic Education: Pursuing High Objectivity in Measuring Spiritual and Social Affective Domains, *Journal of Social Science Studies*, 5(1), 15 - 26.