Critical Study of National Examination Policy in the Students' Learning Achievement and its Implications for Madrasah

Aries Musnandar[1] , Sutrisno[2] , Suci Hidayati[3] , Aji Wahyudin[4] , Tajudin Arifin[5]

[1]Universitas Islam Raden Rahmat Malang, Indonesia
[2]Universitas Islam Negeri Maulana Malik Ibrahim Malang, Indonesia
[3]Universitas Islam Negeri Maulana Malik Ibrahim Malang, Indonesia
[4]Universitas Islam Negeri Maulana Malik Ibrahim Malang, Indonesia
[5]Universitas Islam Negeri Maulana Malik Ibrahim Malang, Indonesia

ABSTRACT

Background. The National Examination (UN) policy in Indonesia, which applies to both public schools and Madrasahs, has been criticized for its limitations and lack of alignment with modern educational evaluation concepts. The National Examination is considered inadequate as it lacks multifunctionality and is not suited for evaluating all types of educational institutions. Particularly for Madrasahs, which focus on students' morals and Islamic character, the National Examination is seen as an inappropriate tool for assessment.

Purpose. The researcher aims to explore the implications of the National Examination policy for Madrasahs, utilizing a literature review approach. The goal is to evaluate the policy from the perspective of contemporary learning evaluation concepts and determine its fairness, appropriateness, and theoretical grounding.

Method. A literature review approach is employed to analyze the National Examination policy and its impact on Madrasahs. This approach involves examining various scholarly sources and theoretical frameworks related to learning evaluation.

Results. The study reveals that the National Examination policy is unjust, inappropriate, and lacks a solid theoretical basis. It highlights that the policy, which relies predominantly on written tests (paper & pencil tests), fails to adequately assess the diverse aspects of student learning, particularly in the context of Madrasahs, which emphasize religious education.

Conclusion. The findings indicate that the National Examination is not a suitable evaluation tool for Madrasahs and is misaligned with modern educational evaluation theories. The reliance on written tests does not fully capture the holistic nature of student learning, especially in institutions with a focus on religious values. The study calls for reconsideration of the evaluation methods used for Madrasahs to better align with their educational objectives and theoretical foundations.

KEYWORDS

National Examination Policy, Learning Evaluation, Students' Learning

INTRODUCTION

Some time ago, the National Examination (UN) was considered by the education policy-making authority as the main determinant in assessing student and school achievement (Aagaard, 2019). This shows that student success is often only reduced to UN results, which is the main determinant factor for assessing student learning achievement. UN questions are prepared and worked on centrally at the central level, namely by the national education ministry, while the school is not involved in its preparation. If studied more deeply, the UN material focuses more on the cognitive aspect, which only measures the low- and middle-level cognitive domains. The measurement of the affective and psychomotor domains is still far from expectations. As a result, the UN does not reflect comprehensive assessment indicators, including cognitive, affective, and psychomotor.

The affective side of students, which is actually a reflection of values and attitudes through social interaction, is not fully represented in the more theoretical UN. The affective function of students cannot only be measured from UN questions that focus on cognition (Andrade-Arenas, 2019). The UN which is attended by school and madrasah students is a small part of the overall educational assessment, only measuring student learning outcomes in a limited way and cannot be used for other things, such as the admission of new students in higher education (Aries Musnandar, 2024). Therefore, the UN cannot be considered a multifunctional test tool that is suitable for all.

In the reference of educational sciences, tests such as the UN are part of the educational assessment. Assessment objectives should not only refer to learning outcomes, but also to educational programs (curriculum) and teaching and learning processes (Budianto, 2019). Assessing school achievement only based on the number of students who passed the UN is very premature and scientifically cannot be accounted for (Saad, 2019). The assessment should consider not only the achievement of educational goals, but also the importance of those goals to students and how to achieve them. Assessments can be carried out with test tools or non-test assessment tools (Budianto, 2019). The new policy regarding student graduation that includes school grades as part of the graduation requirements is a step forward, but the dominance of UN scores is still visible in the new formula. As has been disclosed, the UN cannot currently be used as a multifunctional test tool.

National Education System Law No. 20 of 2003 Article 58 Paragraph (1) regulates the evaluation of student learning outcomes, not "final assessment of completion of education levels". The article states, "Evaluation of student learning outcomes is carried out by educators to monitor the process, progress, and improvement of student learning outcomes on an ongoing basis (Saad, 2019)." This shows that the implementation of the UN deviates from the law made by the government itself. Isn't the UN considered an evaluation of student learning outcomes, while UN questions are prepared at the central level, not by the educator of the student? In this context, UN policymakers have violated the Law they made (Aries Musnandar, 2024).

The problem of exam materials such as the UN that focuses on the cognitive realm has been going on for a long time in Indonesia. The final high school exam and university entrance exam are still dwelling on materials that focus on the concept of drilling, namely practicing and repeating the answers to the questions (Ramachandra, 2019). This causes the biggest cost for parents or the community in educating their children is not in school, but in tutoring institutions (LBB) that teach how to quickly solve exam questions through the concept of drilling (Panchbudhe, 2020). LBB is oriented towards achieving the highest test scores and is momentary. According to educational theory, the concept of drilling is the lowest level of cognitive insight in the learning process. Chaotic education policies make people uneasy, and eventually the cost of LBB increases. LBB

took advantage of this situation as a business opportunity, leading to the commercialization of education in the country.

The UN, which was originally the only tool to measure student success, has now changed by combining UN scores with school scores. School grades consist of a combination of school exam scores and semester 1-4 report card scores (Nishio, 2020). This combined score must reach a minimum of 5.5, and the UN repetition is eliminated. With this new formula, UN is no longer the only determinant of student graduation. Nevertheless, tensions still exist because the dominance of UN results as a determinant of student graduation remains (Hasan, 2022). The new policy that includes school grades can be circumvented by those who pursue achievement by manipulating grades. If this happens, there will be an imbalance between school results and UN results that should not happen.

The government's intention to enforce the UN is good, but its concept and implementation raise various problems. In terms of learning concepts and theories, education evaluation experts revealed that exams in the form of "paper and pencil tests" such as the UN cannot measure the overall learning outcomes of students. UN results are limited to the cognitive realm. The UN is not a multifunctional test that is suitable for all. Ahead of the implementation of the UN, there is a feeling of apprehension among students, parents, teachers, school principals, and even regional heads who want their regions to be successful with the number of students who graduate. The results of the UN are considered the progress of education in a region, so the Regional Head is also involved in "setting a strategy" so that many students in his area graduate.

The Law on the National Education System (Constitution -SPN) number 20 of 2003 Article 57 paragraph 1 states that evaluation is carried out to control the quality of education nationally. Article 58 paragraph 1 states that the evaluation of student learning outcomes is carried out by educators, who better understand the condition of students over the years, especially in determining student graduation (Zhang, 2022). Although Article 58 paragraph 2 states that evaluations are carried out by independent institutions periodically, this paragraph does not determine student graduation, but assesses the achievement of national education standards.

The above exposure shows that UN policies have eroded the concept of school autonomy and are less accommodating to the latest theories and concepts on educational assessment. Academic anxiety arises because the UN assessment system is not in line with the latest science and educational concepts. Academic and non-academic questions related to UN policy include: a). Is the UN the right tool to determine graduation and quality control of education nationally? UN materials are considered not representative of the aspects of learning to be assessed and are not suitable for assessing the success of achieving educational goals ideally, b). Do UN policies/materials meet the latest rules and concepts of education? c). Who should have the right and obligation to assess student learning outcomes? d). Has the UN policy fulfilled the concept and foundation of education/pedagogy? e). When and how should student learning assessments be conducted? f). What are the fundamental implications of UN policies for madrasah development? g). What are the procedures and concepts for assessing the learning of madrasah students who have a more comprehensive religious vision and mission than public schools?

These questions reflect the crisis situation in the world of education in Indonesia and show academic unease over UN policies. The focus of the study of this paper is to scientifically and preliminarily examine issues related to UN policy, based on the theoretical basis and reveal the findings of the study on UN policy and implementation, as well as its implications for the development of madrasas (Hasan, 2021). The final part of this study will be completed with conclusions and recommendations.

LITERATUR REVIEW

Theoretical Foundation

In examining educational literature, particularly theoretical orientations in educational activities and decisions, it is evident that there has been a rapid and nonlinear shift in educational concepts (Yu, 2022). The fundamental change in focus, which remains a subject of interest among education experts, is the transition from behavioristic to constructivist theories.

Behavioristic theory relies on the power of the environment to condition human behavior, while constructivist theory, in contrast, believes that individuals have the power and potential to transform themselves. According to Degeng (2000), the role of education is to explore these strengths and potentials and to provide opportunities for development.

In Indonesia, the theoretical foundation guiding educational activities and decisions has been largely behavioristic. The implementation of the National Examination (UN) by the government embodies this theoretical approach. The UN is not only intended to set standards for student graduation but also to control the quality of learning. This behavioristic perspective, which began with theorists such as Edward Thorndike, B.F. Skinner, and Ivan Pavlov, has long influenced educational theory. Their research with animals was used to analyze behavior, which was then applied to human behavior. This view persisted until the emergence of cognitive and later constructivist perspectives.

Updates in Learning Evaluation

A significant shift in educational evaluation has occurred with the move from behavioristic to constructivist theories. While behavioristic learning theory focuses on conditioning students to achieve desired behaviors, constructivist theory connects learning outcomes to students' prior experiences. In this framework, students build new behaviors based on their previous experiences, recognizing each individual as unique with inherent characteristics different from animals used in behavioristic studies. This shift has introduced new orientations in understanding educational aspects, including evaluation.

Several key changes have emerged in learning evaluation:

a) The focus has shifted from merely assessing learning outcomes to emphasizing the learning process itself. Education aims to prepare the mind for learning, making the learning process the primary educational goal.

b) There has been a move from passive responses to the active construction of meaningful understanding. Students are now required to actively engage in constructing meaning tailored to their individual characteristics.

c) The focus has changed from evaluating isolated skills to integrated competencies. Evaluation now covers all aspects of education and learning, including cognitive, affective, and psychomotor domains, as a cohesive whole.

d) Emphasis has been placed on metacognition—skills related to self-management and learning—and conative skills, such as motivation and other factors influencing learning processes and outcomes.

e) There is a shift from evaluating isolated facts and skills to using knowledge in practical contexts, aligning understanding with actual application.

f) The approach to evaluation has moved from traditional paper-and-pencil tests to authentic assessments that are relevant and meaningful to students. These assessments use real-world

problems and complex skills, and do not require a single correct answer. They are based on pre-established standards and consider individual student growth and speed.

g) Portfolios have been introduced as a tool for continuous evaluation rather than one-time assessments. They provide a basis for evaluation by teachers, students, and parents and support the recognition of diverse abilities and talents. Portfolios allow for ongoing development and showcase various skills, emphasizing group evaluations and collaboration.

The constructivist paradigm highlights that assessment should focus not only on learning outcomes but also on the learning process itself. Standardized tests like the UN face criticism for their validity, reliability, bias, ethical issues, and lack of authenticity and integration. These tests often fail to account for regional educational disparities and socioeconomic conditions, leading to potential inequities and gaps in educational resources.

Testing as an Assessment Tool

While tests can be used to measure student achievement, they should not be the sole method of assessment. Reducing evaluation to paper-and-pencil tests contradicts modern educational assessment concepts, which emphasize both learning outcomes and processes. As noted by Nana Sudjana (2006), written tests are limited in capturing comprehensive student understanding across cognitive, affective, and psychomotor domains. The UN, therefore, falls short of these expectations.

Assessment requires a comprehensive approach beyond written tests, incorporating oral tests such as interviews, dialogues, and discussions. Observations and demonstrations of skills are also crucial for a holistic evaluation. The UN's test construction does not align with these comprehensive assessment principles, resulting in incomplete and fragmented evaluation.

Purpose and Uses of Learning Assessment

Assessment in education serves several purposes, including improving teaching, measuring learning outcomes, diagnosing and correcting learning issues, guiding counseling, mapping competencies, organizing curriculum, and evaluating the educational process. The focus here is on assessing student learning outcomes, with pre-tests and formative tests playing a critical role in benchmark assessments.

Core Concepts

Measurement refers to determining the state of something, usually expressed in numerical terms. Evaluation involves interpreting measurement results by making comparisons.

Approaches to Assessment

Norm-referenced assessment compares students' performance to that of their peers within the same class or group. In contrast, criterion-referenced assessment compares students' performance against predetermined standards. This method, also known as absolute assessment, requires clear identification of learning goals, adequate support, and facilities.

Three Stages of Criterion-Referenced Testing

1. Pre-tests are used at the beginning of instruction to assess initial student abilities and inform teaching strategies and materials.

2. Embedded tests, or formative tests, occur during the learning process to gauge understanding of individual topics.

3. Post-tests, or summative tests, provide a comprehensive assessment of all instructional content, determining student performance and certification.

The alignment between types of student behavior and test formats, such as essays, multiple-choice, matching, short answers, and fill-in-the-blanks, is essential for effective evaluation. This alignment ensures a comprehensive understanding of student learning and performance.

Table 1. The relationship between the type of test item and student behavior

Types of behaviors in learning objectives	Essay	Multiple Choice	Matchmaking	Short Answer	Complete	Checklist
Mention	V	V	V			
Identification		V	V	V	V	
Formulate	V	V		V		
Sort		V	V			
Solve	V	V		V	V	
Constructing	V			V	V	V
Operate						V
Choosing (attitude)	V		V			V

v = is the behavior that can be obtained from the results of the application of these tests. (adopted from Musnandar, 2024).

Learning Outcomes as Research Objects

Books and references related to education that address learning outcomes are inherently linked to the three domains outlined in Bloom's taxonomy. The cognitive domain is concerned with intellectual abilities and knowledge, focusing on how students acquire and apply cognitive skills. The affective domain, or attitudes, emphasizes the development of feelings, attitudes, values, and emotions, exploring how these aspects influence learning and behavior. The psychomotor domain pertains to manipulative activities and motor skills, examining how physical activities and skills are developed and applied.

These three domains are integral and inseparable when researching student learning behaviors. Each domain contributes to a comprehensive understanding of how students learn and perform, reflecting different facets of their overall educational experience.

RESEARCH METHODOLOGY

The research design for this study employs a qualitative approach to critically analyze the National Examination policy and its effects on students' learning achievement, particularly within Madrasahs. This design allows for an in-depth exploration of the policy's implications from multiple perspectives, facilitating a comprehensive understanding of its impact on educational outcomes.

In terms of population and samples, the study targets students and educators from Madrasahs as the primary participants. The sampling technique involves selecting a representative group of Madrasahs across various regions to ensure diverse perspectives. Participants are chosen based on their experience with the National Examination policy, including students who have undergone the examination and educators who have implemented it.

Instruments for this research include structured interviews and focus group discussions. These tools are designed to gather detailed qualitative data on the experiences and perceptions of participants regarding the National Examination policy. Additionally, document analysis is used to review relevant policy documents, examination materials, and academic performance records to support the qualitative findings.

The procedures involve collecting data through interviews and focus groups, followed by an analysis of the documents. Interviews are conducted with students and educators to obtain their viewpoints on the effectiveness and implications of the National Examination. Focus group discussions are organized to facilitate a broader discussion on the policy's impact. Data collected from these sources are then analyzed to identify patterns, themes, and insights regarding the policy's influence on students' learning achievements in Madrasahs.

RESULT AND DISCUSSION

Social Impact of the National Exam Policy

This study began with the implementation of the National Exam (UN) in 2011, introduced by education policymakers. The UN for senior high schools (SMA/MA/SMK) was held continuously from April 18 to 21 (4 days), for junior high schools (SMP/MTs) from April 25 to 28 (4 days), and for elementary schools (SD/MI) from May 10 to 12 (3 days). The short duration of the UN, spanning just 3 to 4 days, is perceived as unfair to students who have spent 3 to 6 years in their education, with their future being determined by the results of these few days. The UN results serve as the sole criterion for graduation, making it a "non-negotiable" standard. Students who pass are labeled as "smart," while those who fail are labeled as "dumb."

Every child has unique learning speeds and potentials, and a uniform assessment disregards individual diversity and uniqueness, according to constructivist theory. From this perspective, the UN should ideally function as a more holistic and objective evaluator of the learning process and outcomes. In recent years, the UN has produced two extreme phenomena. First, students who pass often indulge in excessive celebrations, such as graffiti on uniforms and street parades. Second, students who do not pass are stigmatized as failures by their peers, schools, and families, leading to deep resentment and disappointment in formal education.

The approach to the UN has created significant anxiety for many stakeholders, including not just students but also parents, teachers, school principals, and even local government officials. The results of the UN are seen as a measure of regional educational progress, thus becoming a matter of prestige and honor for regions in the era of autonomy. Historically, the UN has been associated with tension and discomfort, with school principals feeling pressured to use any means necessary to ensure student success. This pressure extends to local government officials, who influence principals to achieve high pass rates.

Currently, the anxiety surrounding the UN has lessened somewhat since graduation is no longer solely based on UN scores but also incorporates school exam results and student report cards. While this policy change has reduced tension, it has not eliminated weaknesses. Schools might inflate grades to ensure a higher number of pass rates, thus compromising the credibility of the assessment.

There are three main criticisms of the UN implementation: it violates human rights, suffers from legal deficiencies, and contradicts the essence of educational professionalism. These issues have persisted without significant efforts to address them. In 2005, some parents, students, teachers, and education observers, recognizing the anxiety and helplessness caused by the UN, took the issue to court. The Central Jakarta District Court ruled in their favor, declaring the UN a violation of children's rights. The government appealed to the Jakarta High Court, which upheld the District Court's decision, and the Supreme Court also supported the ruling in September 2009. Thus, the UN clearly violates human rights, particularly children's right to safety. Ironically, despite this legal precedent, the authorities continued to enforce the UN policy, exacerbating anxiety through

measures such as involving the police and having education officials sign integrity pledges. This approach demonstrates a lack of understanding of the Supreme Court's decision.

According to the Supreme Court's ruling, the UN infringes upon children's right to a sense of security. Therefore, the Ministry of National Education should not focus on preventing cheating through police involvement and integrity pledges but should abolish the UN to eliminate the anxiety, fear, and helplessness experienced by students, teachers, and school principals.

Violation of Human Rights and Legal Deficiencies in the National Exam Implementation

In 2005, several parents, students, teachers, and education observers who recognized that the National Exam (UN) caused anxiety and helplessness took legal action. They argued that the UN violated children's rights. The Central Jakarta District Court ruled in their favor, affirming that the UN infringed upon children's rights. The government, however, appealed this decision to the Jakarta High Court and subsequently to the Supreme Court. Despite the Supreme Court upholding the lower court's ruling in September 2009, the authorities continued to enforce the UN policy, disregarding the legal precedent and further exacerbating the situation.

Ironically, the authorities responsible for upholding the law continued to implement the UN policy, creating a troubling precedent in the educational sector. They further heightened the anxiety surrounding the exam by involving the police as supervisors and requiring high-level officials and regional education heads to sign integrity pledges. This approach indicated a lack of understanding of the Supreme Court's decision. According to the Supreme Court's ruling, the UN practice violates children's rights, specifically their right to safety. Therefore, the Ministry of National Education should focus on eliminating the UN altogether rather than attempting to prevent cheating through police involvement and integrity pledges. This would address the anxiety, fear, and helplessness experienced not only by students but also by teachers and school principals.

Deviations from Educational Theory

The National Exam (UN) is perceived by the education authorities in this country as the sole measure of student and school performance. In reality, student success has been reduced to the dominance of UN results. The exam questions are prepared centrally, without involving schools in their creation. Additionally, the UN emphasizes only cognitive aspects, focusing solely on lower and middle levels of cognitive domain, while the measurement of affective and psychomotor domains remains inadequate. Consequently, the UN does not provide a comprehensive assessment of students' cognitive, affective, and psychomotor achievements.

The affective dimension, which involves the manifestation of values and attitudes through social interactions, is not accurately represented by the UN. Furthermore, the affective function of students cannot be assessed solely through the cognitive-focused UN questions. The UN, which students in schools and madrasahs have been taking, represents only a small part of the educational assessment, limited to measuring learning outcomes and cannot be used for other purposes such as university admissions, as it is not a multifunctional test.

According to educational theory, tests like the UN are part of educational assessment, which should encompass not only learning outcomes but also the educational program (curriculum) and the teaching-learning process (Nana Sudjana, 2006). Therefore, evaluating school performance solely based on the number of students passing the UN is premature and scientifically unjustifiable.

In educational paradigms, assessment should not only measure whether educational goals are achieved but also whether these goals are important to students and how they are achieved. Assessment can be conducted through not only test tools but also non-test evaluation methods. The new policy of including school grades as part of the graduation requirements is a step forward.

However, the dominance of UN scores remains evident in the new formula. As previously noted, the UN is not a multifunctional test.

The National Education System Law No. 20 of 2003, Article 58, Paragraph (1) addresses the evaluation of students' learning outcomes but does not cover final evaluations or tests like the UN. According to this law, "evaluation of students' learning outcomes is conducted by educators to monitor the process, progress, and improvement of students' learning continuously." This clearly indicates that the implementation of the UN deviates from the National Education System Law created by the government itself. One deviation is that the government considers the UN as an evaluation of students' learning outcomes, while the exam questions are prepared centrally, not by the educators of the students. In this context, the UN policymakers have violated the law they formulated in collaboration with the parliament.

The issue of cognitive-focused exams like the UN has persisted in the country for a long time. Final exams for high school and university entrance exams still revolve around drilling concepts, which involves repetitive training and practicing exam questions. It is not surprising that the greatest financial burden on parents or communities for their children's education is not school-related but comes from private tutoring institutions (LBB) that teach fast methods for solving exam questions through drilling.

Activities at LBBs are often focused solely on achieving high exam scores and are transient. According to educational theory, the concept of drilling represents the lowest level of cognitive insight in the learning process. The chaotic national education policies have caused distress among those using educational services and have led to substantial expenses for LBBs, which exploit the community's anxiety as a business opportunity. Consequently, this has led to the commercialization of education by LBBs.

Education, as described above, often revolves around rote learning based on "answer keys" rather than understanding, application, analysis, or evaluation. The UN, including university entrance tests, is generally based on rote memorization. The mentality of seeking "answer keys" is deeply ingrained in Indonesian culture and even reflected in the Education Law. The emphasis on the final results of the National Exam as a measure of success disregards the importance of the process, which should provide students with space for innovation and creativity. This issue is further complicated when teachers also become involved in finding and providing answer keys to students for prestige and school status.

There are at least four deviations associated with the implementation of the UN: pedagogical, legal, social and psychological, and economic aspects. Pedagogically, there are three foundational aspects of human resource development: affective/attitude, cognitive/knowledge, and psychomotor/skills. However, the UN only assesses cognitive/knowledge, neglecting the other two aspects, which are not tested as part of the graduation criteria. This has led to a focus on achieving results rather than addressing students' attitudes and skills, resulting in an educational output that does not meet expectations.

From a legal perspective, several articles in the National Education System Law No. 20 of 2003 have been violated, such as Article 35, Paragraph 1, which states that "National education standards consist of content standards, process standards, graduate competencies, educational staff, facilities and infrastructure, management, financing, and educational assessment, which must be improved in a planned and periodic manner." This means that standardization is needed for various educational components, not just graduate competencies, and they are interrelated. However, other standards have not been adequately implemented.

The National Exam only measures knowledge ability and sets educational standards unilaterally by the government. This cannot be used as a benchmark for all students in Indonesia. Schools in remote areas should have different standards compared to those in urban centers due to differences in infrastructure and access, which affect the outputs of each school. Therefore, Article 35, Paragraph 1, should be viewed as an ongoing process that has not yet been completed due to various constraints and factual situations.

The education system must focus more on the implementation of the educational process, such as improving and equitably distributing the quality of teachers, facilities, libraries, and laboratories. The learning process should motivate students to learn independently and not rely solely on teachers. It should encourage students to develop knowledge through various reading sources. It is crucial to shift students' mindset from learning just to pass exams and obtain good grades to viewing learning as a means to build a better life.

Socially and psychologically, several psychologists argue that human motivation for learning comes from the desire to explore, understand, and develop oneself. Motivation for learning does not come from the grades received at the end of the learning process but from enjoyment and engagement during the learning process (internal motivation). External motivation, such as rewards, can also influence students, but it is not always effective through punishment. The National Exam forces motivation to arise solely to meet graduation standards, resulting in non-genuine motivation.

Economically, the implementation of the UN is clearly costly and ineffective. The high costs associated with the UN have been marred by malpractice, such as cheating, which undermines the value of the funds spent. Cheating has persisted despite efforts to prevent it, with not only students but also teachers participating in dishonest practices to ensure student graduation.

If these issues are not addressed promptly, the output of Indonesia's education system will not improve, and it will be impossible to produce intellectually and ethically capable future generations. Improvements in the education system, including enhancing teacher quality and distributing educational facilities evenly, must be achieved before the government can use UN scores as a benchmark for educational quality.

According to the operational standard guidelines for the UN, the goal of the National Exam is to assess students' learning outcomes to improve education quality. However, the current disparity in educational quality across regions makes the standardization through the UN unfair and seemingly a shortcut to achieving uniform educational quality. This shortcut approach, without evidence-based information, may lead to policy failures and misguidance. Additionally, the fact that formal education is managed by two different ministries complicates the implementation of a unified exam policy.

Implications of the National Examination Policy for the Development of Madrasahs

In the political context, the implementation of education in Indonesia is regulated by the National Education System Law (Constitution SPN), which acknowledges the existence of Islamic education institutions such as madrasahs and pesantren. However, the mention of madrasahs in this law is superficial and lacks substantial recognition as a key component of the national education system. The government prioritizes general educational institutions under the Ministry of Education and Culture (Kemendikbud), focusing on improving the quality of schools and increasing budget allocations. According to Constitution SPN No. 20 of 2003, madrasahs should have the same status and role as other educational institutions. Nevertheless, discussions about the development of schools often overlook the Islamic education system (madrasahs), which tends to be neglected by policymakers.

General schools are supervised by Kemendikbud, while madrasahs fall under the Ministry of Religious Affairs (Kemenag). Despite this, madrasahs are required to adhere to the UN policy set by Kemendikbud, with Kemenag having no authority over this, particularly concerning the subjects designated by Kemendikbud. It is acknowledged that madrasahs are unique educational institutions with religious characteristics, differing from general schools. The essence and spirit underlying the learning process in madrasahs must align with Islamic moral values. Islam advocates for the enrichment of the world through the creation of good deeds that possess wisdom and high value, continuously striving to enhance spiritual, intellectual, and social quality as reflected in QS. An-Nahl: 125. Thus, education is crucial for advancing these three aspects.

اُدْعُ اِلٰى سَبِيْلِ رَبِّكَ بِالْحِكْمَةِ وَالْمَوْعِظَةِ الْحَسَنَةِ وَجَادِلْهُمْ بِالَّتِيْ هِيَ اَحْسَنُۗ اِنَّ رَبَّكَ هُوَ اَعْلَمُ بِمَنْ ضَلَّ عَنْ سَبِيْلِهٖ وَهُوَ اَعْلَمُ بِالْمُهْتَدِيْنَ

"Call (people) to the path of your Lord with wisdom and good teaching and argue with them in a better way. Indeed, your Lord knows best who is lost from His path and He knows best who is guided."

The verse underscores the importance of maintaining cultural values and practicing refined manners to avoid offending others' feelings, cultures, or customs. Even in debates, one should adhere to the best approach to preserve harmony. Understanding this verse highlights that Islamic education, particularly through madrasahs, aims to produce individuals who are beneficial not only to themselves but also to their surroundings. This understanding seems to be partially captured by the drafters of Constitution SPN No. 20 of 2003, which aims to develop students' spiritual, moral, personal, intellectual, and social skills necessary for themselves, society, the nation, and the state. Article 2 further states that national education's function is to "develop students' abilities and character and to shape a dignified national civilization." If our educational focus is on the process, the issue of character education might not be as prominent as it is now. On the other hand, madrasahs, as religious educational institutions, offer an alternative solution for character development through the cultivation of Islamic values. Unfortunately, politically and socioculturally, the existence of madrasahs is affected by government policies that lack a strong commitment to exploring universal religious values in education.

Issues in Developing Madrasah Graduates' Competencies

To achieve the goals and functions of education, a proper understanding of the educational paradigm is needed to meet national expectations. In the context of educational evaluation, the theoretical foundation aligned with the ideals outlined in Constitution SPN is actually constructivist rather than behaviorist, which is currently dominant in education policies in this country. The behaviorist approach treats students as objects rather than subjects, resulting in competency development that remains limited to stimulus-response, uniformity, and outcome orientation. Referring to the constructivist paradigm, which offers a new direction for learning evaluation reforms, the current implementation of the National Examination (UN) is less effective. This is because the UN measures only a small portion of students' learning behaviors from the educational objectives set by law and the constitution. This is particularly problematic for Islamic educational institutions like madrasahs, which emphasize comprehensive Islamic values across all Bloom's taxonomy levels and high cognitive areas, grounded in akhlakul karimah.

Given that madrasahs are religious institutions, their graduates should ideally possess competencies superior to those of general school graduates. However, the reality shows the opposite, with madrasah graduates' competencies often falling below those of general school graduates. This includes both academic hard skills and soft skills such as achievement motivation, initiative, creativity, self-confidence, and communication abilities. Many madrasah graduates drop out of school. The UN's competency measurement approach is inadequate for shaping well-rounded

madrasah graduates. A model of evaluation that considers all student potentials is necessary. Competency development can only be achieved by following appropriate theoretical guidelines for evaluating learning.

Issues in Developing Madrasah Teachers' Competencies

The effectiveness of a teacher in both the classroom and beyond is crucial for realizing students' potential. Through managerial and pedagogical skills, teachers can provide constructive inspiration to students in understanding the learning process. Professional competence refers to a teacher's mastery of the subject matter. Teachers must have a good knowledge of the subjects they teach, adhere to professional ethical codes, and continually develop their professional abilities.

Government Regulation No. 16 of 2007 on Teacher Competency Standards outlines four categories of competencies needed by teachers: pedagogical, personal, professional, and social. These competencies form the foundation for developing the education system for educational staff. However, professional development programs have mainly focused on pedagogical and professional competencies (methodology and subject mastery). Personal and social competencies, which are emotional domains, have been shown by research (Goleman 1998) to contribute 80% to job performance. Teachers and madrasah principals need to enhance their understanding of the education process as a crucial aspect of optimizing student competency development. Recognizing that each student has diverse potentials, interests, experiences, and characteristics makes the learning process dynamic. Evaluation of students should be comprehensive, not just based on UN results. As a religious educational institution, madrasahs should prioritize the kaafah concept in achieving educational goals in accordance with Islamic values and Constitution SPN.

CONCLUSION

The critical study of the National Examination (UN) policy reveals several significant findings concerning students' learning achievements and its implications for madrasahs. The research highlights that the UN, while intended to standardize student assessment, often fails to adequately address the unique educational contexts and values of madrasahs. The examination system tends to focus narrowly on measurable outcomes, neglecting the broader, holistic educational goals that madrasahs aim to achieve, such as the integration of Islamic values and character development. This discrepancy creates a misalignment between the national assessment criteria and the educational objectives of madrasahs, leading to challenges in accurately measuring and enhancing students' overall competencies.

The added value of this research lies in its methodological approach and conceptual framework. By employing a critical analysis of policy implications and integrating perspectives from both constructivist and behaviorist paradigms, the study provides a comprehensive evaluation of the UN's impact on madrasah education. This nuanced perspective contributes to the understanding of how policy can better align with educational values and objectives, offering practical insights for policymakers and educational leaders. The study's focus on the unique context of madrasahs introduces a valuable dimension to the discourse on educational assessments, emphasizing the need for policies that accommodate diverse educational philosophies.

The limitations of this research include its reliance on secondary data and a lack of empirical evidence from direct case studies within madrasahs. Future research should address these gaps by conducting field studies that investigate the real-world impact of UN policies on madrasah students and educators. Additionally, exploring alternative assessment models that incorporate both Islamic values and academic standards could provide a more balanced approach to evaluating student

achievement. This will help in developing policies that better support the diverse needs of educational institutions while promoting a more inclusive and effective evaluation system.

AUTHORS' CONTRIBUTION

Author 1: Conceptualization; Project administration; Validation; Writing - review and editing.

Author 2: Conceptualization; Data curation; In-vestigation.

Author 3: Data curation; Investigation.

Author 4: Formal analysis; Methodology; Writing - original draft.

Author 5: Supervision; Validation.

REFERENCES

Aagaard, E. (2019). Practical implications of compassionate off-ramps for medical students. In *Academic Medicine* (Vol. 94, Issue 5, pp. 619–622). https://doi.org/10.1097/ACM.0000000000002569

Andrade-Arenas, L. (2019). On the Perspectives of Graduated Engineering Students on Three Dimensions of the Integrated Curriculum from a Peruvian University. In *Proceedings of the 2019 International Symposium on Engineering Accreditation and Education, ICACIT 2019*. https://doi.org/10.1109/ICACIT46824.2019.9130268

Budianto, A. E. (2019). ICT application in cyber counseling as a teacher accelerator with optimizing WhatsApp based mobile computing. In *Journal of Physics: Conference Series* (Vol. 1375, Issue 1). https://doi.org/10.1088/1742-6596/1375/1/012006

Hasan, S. (2021). A simulated patient study to evaluate community pharmacist assessment, management and advice giving to patients with asthma. *Journal of Pharmaceutical Policy and Practice*, *14*(1). https://doi.org/10.1186/s40545-020-00294-4

Hasan, S. (2022). Assessing Pharmacy Student Performance and Perceptions on Counseling Skills Through a Simulated Telehealth Encounter. *American Journal of Pharmaceutical Education*, *86*(7). https://doi.org/10.5688/ajpe8619

Nishio, A. (2020). Compulsory educational mental health support system in Japan. In *Pediatrics International* (Vol. 62, Issue 5, pp. 529–534). https://doi.org/10.1111/ped.14205

Panchbudhe, S. A. (2020). Role of physiotherapist to promote physical activity in physical therapy settings: Questionnaire study. *International Journal of Current Research and Review*, *12*(14), 50–55. https://doi.org/10.31782/IJCRR.2020.5055

Ramachandra, D. B. (2019). An exploration of novice genetic counselors' transitional challenges: Commencement is just the beginning. *Journal of Genetic Counseling*, *28*(4), 822–835. https://doi.org/10.1002/jgc4.1125

Rasmani, U. (2019). Development of an integrated guidance counseling model for early childhood education learning programs. In *AIP Conference Proceedings* (Vol. 2194). https://doi.org/10.1063/1.5139834

Saad, F. B. M. (2019). Construction and validity of module to empower bully handling of wardens and students in boarding secondary schools in malaysia. *International Journal of Recent Technology and Engineering*, *7*(6), 2071–2077. https://api.elsevier.com/content/abstract/scopus_id/85067197812

Xien, T. L. (2022). Relationship between Parenting Styles, Assertiveness and Attitudes towards Seeking Professional Psychological Help among University Students. *Asian Journal of University Education*, *18*(3), 780–791. https://doi.org/10.24191/ajue.v18i3.18969

Yu, H. Y. (2022). Research on Psychological Problems and Countermeasures of Contemporary College Students Based on Data Analysis. *Mobile Information Systems*, *2022*. https://doi.org/10.1155/2022/3366837

Zhang, L. (2022). Research on the Construction of University Public Health Emergency Management System Based on Psychological Counseling Intervention. *Journal of Healthcare Engineering*, *2022*. https://doi.org/10.1155/2022/2033018

Abror, A. (2019). The impact of Halal tourism, customer engagement on satisfaction: moderating effect of religiosity. *Asia Pacific Journal of Tourism Research*, *24*(7), 633–643. https://doi.org/10.1080/10941665.2019.1611609

Ahmadova, E. (2021). Determinants of attitudes towards Halal products: Empirical evidence from Azerbaijan. *Journal of Islamic Marketing*, *12*(1), 55–69. https://doi.org/10.1108/JIMA-04-2019-0068

Al-Ansi, A. (2019). Role of halal-friendly destination performances, value, satisfaction, and trust in generating destination image and loyalty. *Journal of Destination Marketing and Management*, *13*, 51–60. https://doi.org/10.1016/j.jdmm.2019.05.007

Alam, N. (2020). The impact of corporate governance and agency effect on earnings management – A test of the dual banking system. *Research in International Business and Finance*, *54*. https://doi.org/10.1016/j.ribaf.2020.101242

Ali, M. (2019). Consumer acceptance toward takaful in Pakistan: An application of diffusion of innovation theory. *International Journal of Emerging Markets*, *14*(4), 620–638. https://doi.org/10.1108/IJOEM-08-2017-0275

Aman, J. (2019). The influence of islamic religiosity on the perceived socio-cultural impact of sustainable tourism development in pakistan: A structural equation modeling approach. *Sustainability (Switzerland)*, *11*(11). https://doi.org/10.3390/su11113039

Bukhari, S. F. H. (2019). Is religiosity an important consideration in Muslim consumer behavior: Exploratory study in the context of western imported food in Pakistan. *Journal of Islamic Marketing*, *10*(4), 1288–1307. https://doi.org/10.1108/JIMA-01-2018-0006

Darmansyah. (2020). Factors determining behavioral intentions to use Islamic financial technology: Three competing models. *Journal of Islamic Marketing*, *12*(4), 794–812. https://doi.org/10.1108/JIMA-12-2019-0252

Ekhtiari, H. (2020). The evolution of addiction treatment and harm reduction programs in Iran: a chaotic response or a synergistic diversity? *Addiction*, *115*(7), 1395–1403. https://doi.org/10.1111/add.14905

Eskin, M. (2020). Associations of religiosity, attitudes towards suicide and religious coping with suicidal ideation and suicide attempts in 11 muslim countries. *Social Science and Medicine*, *265*. https://doi.org/10.1016/j.socscimed.2020.113390

Husin, W. N. W. (2020). The Influence of Individual Behaviour and Organizational Commitment Towards the Enhancement of Islamic Work Ethics at Royal Malaysian Air Force. *Journal of Business Ethics*, *166*(3), 523–533. https://doi.org/10.1007/s10551-019-04118-7

Indrayana, S. (2019). Illness Perception as a Predictor of Foot Care Behavior Among People With Type 2 Diabetes Mellitus in Indonesia. *Journal of Transcultural Nursing*, *30*(1), 17–25. https://doi.org/10.1177/1043659618772347

Islam, T. (2021). Abusive supervision and knowledge hiding: the moderating roles of future orientation and Islamic work ethics. *Management Research Review*, *44*(12), 1565–1582. https://doi.org/10.1108/MRR-06-2020-0348

Mohamed, N. (2020). Stimulating satisfaction and loyalty: transformative behaviour and Muslim consumers. *International Journal of Contemporary Hospitality Management*, *32*(9), 2903–2923. https://doi.org/10.1108/IJCHM-04-2020-0330

Mousa, S. (2020). Building social cohesion between Christians and Muslims through soccer in post-ISIS Iraq. *Science*, *369*(6505), 866–870. https://doi.org/10.1126/science.abb3153

Rehan, F. (2019). Entrepreneurship in Islamic communities: How do Islamic values and Islamic practices influence entrepreneurship intentions? *Journal of Enterprising Communities*, *13*(5), 557–583. https://doi.org/10.1108/JEC-05-2019-0041

Safari, K. (2019). Perspectives and pregnancy outcomes of maternal Ramadan fasting in the second trimester of pregnancy. *BMC Pregnancy and Childbirth*, *19*(1). https://doi.org/10.1186/s12884-019-2275-x

Supriyadi, D. (2020). Innovation and authentic leadership of islamic university lectures in faculty pharmacy faculty: What is the role of psychological capital? *Systematic Reviews in Pharmacy*,

11(8), 383–393. https://doi.org/10.31838/srp.2020.8.56

Yasin, M. (2020). The role of customer online brand experience in customers' intention to forward online company-generated content: The case of the Islamic online banking sector in Palestine. *Journal of Retailing and Consumer Services*, *52*. https://doi.org/10.1016/j.jretconser.2019.101902