



Criterion-Referenced and Norm-Referenced Assessment: An Analysis of Score Processing Techniques in Islamic Religious Education

Ikhwan Al Hafiz^{1*}, Remiswal², Khadijah³

¹⁻³ Universitas Islam Negeri Imam Bonjol, Indonesia
email: ikhwan.al.hafiz@uinib.ac.id¹

Article Info :

Received:
21-04-2026
Revised:
02-05-2026
Accepted:
13-05-2026

Abstract

This study examines the conceptual foundations of Criterion-Referenced Assessment (CRA) and Norm-Referenced Assessment (NRA), with particular emphasis on their score-processing techniques and pedagogical relevance in Islamic Religious Education (IRE). A qualitative, non-empirical approach employing library research was adopted by systematically reviewing scholarly books, peer-reviewed journal articles, and relevant educational policy documents. The collected literature was analyzed through qualitative content analysis and comparative conceptual analysis to identify the characteristics, differences, pedagogical functions, and implications of both assessment approaches. The findings reveal that CRA evaluates students' achievement against predetermined competency standards, enabling objective interpretation of learning outcomes and supporting mastery learning through competency-based assessment. In contrast, NRA interprets achievement relative to group performance, making it more appropriate for ranking, classification, and selection purposes than for measuring competency attainment. The analysis further demonstrates that CRA is more compatible with the objectives of Islamic Religious Education because it facilitates comprehensive assessment across cognitive, affective, and psychomotor domains while supporting individualized feedback and competency development. Although NRA remains useful in specific educational contexts, its role in routine classroom assessment is complementary rather than primary. The study contributes to contemporary pedagogy by proposing an integrated conceptual perspective linking assessment paradigms, score-processing techniques, and competency-based evaluation in Islamic Religious Education.

Keywords: Criterion-Referenced Assessment, Competency-Based Assessment, Islamic Religious Education, Norm-Referenced Assessment, Score-Processing Techniques.



©2022 Authors.. This work is licensed under a Creative Commons Attribution-Non Commercial 4.0 International License.
(<https://creativecommons.org/licenses/by-nc/4.0/>)

INTRODUCTION

Assessment has become one of the most influential components of contemporary pedagogy because it not only determines the extent to which students achieve expected learning outcomes but also guides instructional improvement, curriculum refinement, and evidence-based educational policymaking. Contemporary educational systems have progressively shifted from content-oriented instruction toward competency-based learning, requiring assessment practices that accurately capture students' mastery of intended competencies rather than merely documenting academic performance (Sternberg et al., 2022). Recent developments in learning science further emphasize that assessment should function as an integral component of the learning process by providing meaningful evidence for instructional adjustment and continuous learner development (Littman et al., 2024). This transformation has encouraged educational researchers to reconsider traditional score interpretation models in light of validity, reliability, fairness, and instructional usefulness across increasingly diverse learning environments (Burr et al., 2026). Universities, schools, and curriculum authorities worldwide have consequently revisited assessment policies to ensure that educational decisions are based on meaningful interpretations of student achievement instead of isolated numerical outcomes (Geiser, 2022). These international developments indicate that assessment systems are expected to provide valid evidence capable of supporting both individual learning improvement and institutional accountability within competency-oriented education. In Islamic Religious Education (IRE), this expectation becomes considerably more demanding because learning outcomes encompass cognitive understanding, affective internalization of Islamic values, and psychomotor demonstrations of religious practices that

cannot be represented adequately through a single measurement perspective (Pranajaya et al., 2023). The complexity of these multidimensional competencies places assessment frameworks at the center of pedagogical discussions, making the selection of appropriate score interpretation approaches an essential issue for contemporary learning science.

Among the numerous assessment paradigms developed within educational measurement, criterion-referenced assessment and norm-referenced assessment remain the two most influential approaches for interpreting student achievement because each is grounded in fundamentally different assumptions regarding learning success and educational quality. Criterion-referenced assessment determines students' achievement by comparing performance with predetermined competency standards, whereas norm-referenced assessment interprets achievement relative to the performance of other students within the same reference group (Gaertner, 2022). These conceptual differences produce substantially different implications for educational decision-making, particularly regarding mastery learning, remediation, enrichment, student placement, and academic ranking. Recent empirical evidence indicates that criterion-referenced assessment has become increasingly compatible with competency-based curricula because it prioritizes individual mastery of learning objectives rather than comparative performance among learners (Hadi et al., 2024). Nevertheless, investigations conducted within higher education demonstrate that norm-referenced grading continues to influence students' academic expectations, motivation, and perceptions of competition, suggesting that comparative assessment still possesses practical relevance in specific educational contexts (Liu et al., 2024). Research focusing on Islamic education similarly reports that both assessment approaches contribute differently to evaluating learning outcomes, with criterion-referenced assessment supporting competency attainment while norm-referenced assessment remaining valuable for selection and classification purposes (Darmawan, 2024). Studies exploring assessment practices in Islamic Religious Education further reveal that teachers frequently combine both approaches depending on instructional objectives, classroom contexts, and institutional demands rather than adopting either paradigm exclusively (Irmayanti et al., 2024). The coexistence of these approaches demonstrates that effective assessment depends not merely on selecting a particular reference framework but on ensuring conceptual alignment between assessment purposes, learning objectives, and score interpretation procedures.

Despite the growing body of scholarship concerning criterion-referenced and norm-referenced assessment, important conceptual and methodological limitations remain insufficiently addressed, particularly regarding score-processing techniques that ultimately determine how educational evidence is interpreted and utilized. Existing studies predominantly discuss the implementation of criterion-referenced and norm-referenced assessment at the procedural level without critically examining how different score-processing strategies influence the validity and fairness of educational decisions (Kiswanto et al., 2024). Research investigating learning mastery frequently concentrates on achievement classification while paying comparatively limited attention to the methodological assumptions underlying score conversion, interpretation, and standard setting (Azijah et al., 2026). Scholarly discussions concerning score processing likewise tend to emphasize computational procedures instead of exploring how processed scores affect instructional interventions, learning evaluation, and educational accountability (Rohmah, 2024). Recent methodological scholarship has further questioned the practice of establishing assessment standards primarily on statistical difficulty because valid educational interpretation requires stronger alignment between assessment purposes, construct validity, and intended competency outcomes (Burr et al., 2026). Historical analyses also remind educational researchers that excessive dependence on norm-referenced testing has long been associated with concerns regarding educational inequality, social comparison, and biased interpretations of learners' abilities that may persist within contemporary assessment practices (Brumley et al., 2026). These unresolved issues become increasingly significant within Islamic Religious Education because multidimensional competencies involving knowledge, attitudes, and religious practices require score-processing techniques capable of producing educationally meaningful interpretations rather than merely statistically accurate classifications. Such limitations indicate that a comprehensive analysis integrating assessment paradigms with score-processing methodologies remains absent from the current literature despite its substantial theoretical and practical importance.

RESEARCH METHODS

This study employed a non-empirical qualitative research design using a library research approach to examine the conceptual foundations of criterion-referenced assessment (CRA) and norm-referenced assessment (NRA), with particular emphasis on score-processing techniques in Islamic Religious Education (IRE). Instead of collecting primary empirical data, the study synthesized theoretical and empirical evidence from scholarly books, peer-reviewed journal articles, and relevant educational policy documents addressing educational assessment, learning evaluation, and Islamic pedagogy. The literature was purposively selected based on four inclusion criteria: relevance to the research topic, methodological rigor, publication credibility, and contribution to the discussion of criterion-referenced and norm-referenced assessment in educational settings (Gaertner, 2022). Priority was given to recent studies published between 2022 and 2026 to ensure that the analysis reflected contemporary developments in competency-based assessment and learning evaluation, while seminal theoretical works were retained to strengthen the conceptual foundation of the study (Burr et al., 2026). This literature selection process enabled the construction of a comprehensive analytical framework that integrates assessment theory with pedagogical practices in Islamic Religious Education (Irmayanti et al., 2024).

The collected literature was analyzed using qualitative content analysis combined with a comparative conceptual analysis to identify similarities, differences, and pedagogical implications of criterion-referenced and norm-referenced assessment, particularly regarding score-processing techniques. The analytical procedure consisted of data reduction, open coding, thematic categorization, cross-literature comparison, and interpretative synthesis to examine how different score-processing approaches influence the interpretation of learning outcomes, competency achievement, and instructional decision-making in Islamic Religious Education (Rohmah, 2024). The analysis also explored the relationship between assessment references and competency-based evaluation by comparing findings from previous studies concerning the implementation of criterion-referenced and norm-referenced assessment in Islamic education (Darmawan, 2024). To enhance analytical rigor, conceptual triangulation was employed by comparing theoretical arguments and empirical findings from multiple scholarly sources, thereby strengthening the consistency and credibility of the resulting interpretations (Hadi et al., 2024). The findings were ultimately presented in a descriptive-analytical form to develop an integrated conceptual understanding of score-processing techniques and their relevance for improving assessment practices in contemporary Islamic Religious Education (Susilowati et al., 2024).

RESULTS AND DISCUSSION

Conceptual Foundations of Criterion-Referenced and Norm-Referenced Assessment

Assessment in competency-oriented education is fundamentally concerned with generating valid evidence regarding learners' achievement while simultaneously supporting instructional improvement and meaningful educational decision-making. The literature synthesis demonstrates that criterion-referenced assessment (CRA) and norm-referenced assessment (NRA) represent two complementary yet conceptually distinct paradigms that differ primarily in their interpretation of learning outcomes rather than in their measurement procedures. CRA evaluates learners' performance against predetermined standards or competency benchmarks, whereas NRA interprets achievement by comparing individual performance with that of a reference group, leading to substantially different educational consequences (Gaertner, 2022). These conceptual distinctions influence how teachers evaluate learning achievement, provide feedback, and determine subsequent instructional interventions. Within Islamic Religious Education (IRE), the assessment framework becomes increasingly significant because learning objectives encompass cognitive mastery, affective development, and psychomotor competence that cannot be adequately represented through comparative achievement alone (Pranajaya et al., 2023). The literature consistently indicates that selecting an appropriate assessment reference is closely associated with the instructional purpose rather than merely the availability of measurement instruments. This conceptual distinction forms the foundation for understanding how score-processing techniques should subsequently be designed and interpreted within competency-based learning.

The comparative characteristics identified from the reviewed literature further clarify the pedagogical orientation of both assessment paradigms. CRA emphasizes competency attainment by comparing student performance with explicitly defined learning criteria, while NRA focuses on

identifying relative achievement within a learning cohort through statistical comparison (Nurmanto & Ashari, 2023). Such differences extend beyond technical measurement because they influence teachers' instructional priorities, students' learning experiences, and the interpretation of educational success. Contemporary curriculum reforms increasingly advocate competency-oriented assessment, encouraging educators to evaluate whether students have mastered intended learning outcomes instead of emphasizing competitive ranking (Hadi et al., 2024). By contrast, NRA remains relevant for educational contexts requiring differentiation among learners, including admission, scholarship selection, and competitive academic programs, where relative standing constitutes an important consideration (Geiser, 2022). The reviewed studies also suggest that neither paradigm should be interpreted as universally superior because their effectiveness depends on the educational objectives, assessment context, and intended use of assessment results. Consequently, conceptual clarity regarding the characteristics of each assessment approach becomes essential before determining appropriate score-processing procedures.

The comparative analysis also reveals that the strengths and limitations of CRA and NRA emerge from their underlying philosophical assumptions regarding educational quality and learner achievement. CRA provides greater transparency in evaluating competency mastery because learners are assessed against explicit standards that remain independent of group performance, supporting equitable opportunities for demonstrating learning achievement (Sternberg et al., 2022). Several studies likewise indicate that criterion-referenced evaluation facilitates formative assessment, individualized feedback, and instructional remediation because achievement is interpreted according to predetermined competency indicators rather than statistical distribution (Susilowati et al., 2024). In contrast, NRA offers advantages when educational institutions require relative classification among learners, particularly for selection and placement purposes where ranking provides meaningful comparative information (Liu et al., 2024). Historical analyses nevertheless caution that excessive dependence on norm-referenced interpretation may unintentionally reinforce social comparison and reduce attention to individual learning progress, especially when relative ranking becomes the primary indicator of educational success (Brumley et al., 2026). These findings indicate that pedagogical effectiveness depends less on the assessment model itself than on the alignment between assessment purposes and score interpretation. Such alignment becomes particularly relevant in Islamic Religious Education because competency development extends beyond measurable cognitive performance.

Table 1. Comparative Characteristics of Criterion-Referenced Assessment and Norm-Referenced Assessment

Dimension	Criterion-Referenced Assessment (CRA)	Norm-Referenced Assessment (NRA)
Assessment reference	Predetermined competency standards	Performance of a comparison group
Primary objective	Measuring competency mastery	Determining relative ranking
Score interpretation	Absolute	Relative
Orientation	Individual achievement	Group comparison
Main educational function	Mastery learning, remediation, feedback	Selection, classification, placement
Suitability for IRE	Highly appropriate	Limited to specific contexts

Table 1 demonstrates that the fundamental distinction between CRA and NRA lies not only in the reference used to interpret scores but also in their educational orientation and pedagogical implications. CRA prioritizes evidence of competency attainment, allowing teachers to identify learning gaps and design appropriate remedial or enrichment activities based on clearly defined performance standards (Darmawan, 2024). Conversely, NRA emphasizes comparative performance, making it particularly useful when educational decisions require differentiation among students with relatively similar academic abilities (Irmayanti et al., 2024). Recent investigations into classroom assessment further indicate that combining both paradigms without a clear conceptual rationale may generate inconsistent interpretations of student achievement because identical raw scores may produce different

educational conclusions depending on the selected assessment reference (Kiswanto et al., 2024). This observation suggests that assessment validity depends not only on measurement accuracy but also on the appropriateness of the interpretive framework applied to learning outcomes. The reviewed literature therefore positions assessment references as pedagogical constructs rather than purely statistical mechanisms.

Within the context of Islamic Religious Education, the literature consistently supports the greater pedagogical compatibility of CRA because Islamic learning objectives emphasize individual competency development, character formation, and continuous improvement rather than competitive comparison among learners. Studies examining assessment implementation in Islamic education report that competency-based evaluation enables teachers to assess knowledge acquisition, religious attitudes, and practical worship skills using clearly established performance indicators that correspond to curriculum expectations (Nurhayani et al., 2024). Similar findings indicate that criterion-based evaluation provides stronger support for formative feedback and individualized instructional planning because achievement is interpreted according to competency attainment rather than class distribution (Nurhijatina et al., 2023). Educational standards governing Islamic Religious Education likewise prioritize learning achievement through competency indicators and expected learning outcomes, reinforcing the compatibility between CRA and competency-based curriculum implementation (Zubaidi, 2022). Even though NRA remains valuable for specific purposes such as academic competitions or student selection, its pedagogical contribution within routine classroom assessment appears comparatively limited because it provides relatively little information regarding students' actual mastery of expected competencies. These conceptual findings establish the theoretical basis for examining how different score-processing techniques influence educational interpretation, which constitutes the primary analytical focus of the subsequent section.

Comparative Analysis of Score Processing Techniques in Islamic Religious Education

The comparative analysis of score-processing techniques reveals that the distinction between criterion-referenced assessment (CRA) and norm-referenced assessment (NRA) extends beyond the interpretation of assessment outcomes to the methodological procedures through which raw scores are transformed into meaningful educational information. Although both approaches originate from identical measurement results, they generate different educational conclusions because they employ fundamentally different reference frameworks during score interpretation. Within CRA, raw scores are converted into indicators of competency mastery by comparing students' achievement with predetermined learning standards, whereas NRA interprets the same scores through statistical comparison with the performance of other learners in the reference group (Gaertner, 2022). This distinction illustrates that score processing is not merely a mathematical operation but also an interpretive process shaped by educational philosophy and assessment objectives. In competency-based education, the meaning attached to student scores determines the quality of instructional decisions, including mastery determination, feedback provision, remediation, and enrichment activities. The reviewed literature consistently demonstrates that educational assessment becomes more meaningful when score interpretation remains aligned with curriculum expectations rather than relying exclusively on statistical distributions (Burr et al., 2026). Such alignment is particularly important in Islamic Religious Education because learning outcomes encompass measurable knowledge alongside behavioral and spiritual competencies that require criterion-based interpretation.

The literature synthesis further indicates that score processing under CRA follows a relatively direct procedure because student achievement is evaluated against predetermined performance criteria established before instruction begins. The conversion of raw scores into percentages or competency levels enables teachers to determine whether students have successfully achieved the expected learning outcomes without considering the performance of their peers (Rohmah, 2024). Contemporary competency-based curricula similarly employ criterion-referenced score processing to identify mastery levels, diagnose learning difficulties, and determine the necessity of remedial instruction or enrichment activities after assessment has been completed (Hadi et al., 2024). This procedure strengthens instructional transparency because students understand the standards they are expected to achieve, while teachers can relate assessment outcomes directly to instructional objectives. Previous investigations within Islamic education further indicate that criterion-based score interpretation supports more consistent classroom evaluation because competency standards remain stable regardless of differences

in classroom composition or student ability distribution (Nurmanto & Ashari, 2023). The analytical evidence therefore suggests that CRA contributes to a more instructionally meaningful assessment system by connecting numerical scores with explicit competency expectations rather than comparative achievement.

A different analytical pattern emerges in NRA, where score processing depends primarily on the statistical characteristics of the reference group rather than predetermined competency standards. Raw scores are interpreted using measures such as the group mean, standard deviation, percentile ranks, or other comparative indicators to determine each student's relative position within the cohort (Geiser, 2022). This approach enables educators to distinguish high-performing, average, and low-performing students for purposes including selection, placement, scholarship allocation, and academic competition. The reviewed studies indicate that NRA remains particularly valuable when educational institutions require differentiation among students who have completed similar learning experiences but must be classified according to relative achievement (Liu et al., 2024). Nevertheless, comparative interpretation may produce substantially different conclusions despite identical raw scores because the educational meaning of achievement changes according to group performance. Historical and contemporary scholarship has repeatedly emphasized that such dependency on group distribution may reduce interpretive consistency when assessment is intended to determine competency mastery rather than competitive ranking (Brumley et al., 2026). These observations illustrate that statistical precision alone cannot guarantee educational validity if the chosen score-processing model is inconsistent with instructional objectives.

The conceptual differences between both score-processing approaches become more evident when identical raw scores are interpreted using different assessment references. Table 2 presents a simplified illustration derived from the comparative literature to demonstrate how the same assessment result may lead to different educational decisions depending on the selected interpretation framework.

Table 2. Illustrative Comparison of Score Processing under Criterion-Referenced and Norm-Referenced Assessment

Student	Raw Score	Criterion-Referenced Assessment	Norm-Referenced Assessment
A	36/40 (90%)	Mastered competency	High performer
B	30/40 (75%)	Mastered competency	Average performer
C	24/40 (60%)	Not yet mastered	Low performer

The illustration demonstrates that CRA interprets student performance according to competency attainment, whereas NRA classifies identical achievement according to comparative position within the learning group. Consequently, a learner categorized as "average" under NRA may nevertheless demonstrate complete competency mastery according to CRA, illustrating how educational conclusions depend on the assessment reference rather than the numerical score alone. Similar observations were reported by Kiswanto et al. (2024), who found that differences in score-processing procedures frequently produce contrasting interpretations despite identical assessment results. Azijah et al. (2026) likewise argue that learning mastery should not be inferred exclusively from relative score distributions because competency attainment requires explicit comparison with predetermined educational standards. These findings reinforce the view that score-processing techniques function as interpretive mechanisms rather than purely statistical procedures.

The implications of these analytical findings become increasingly significant within Islamic Religious Education because assessment encompasses cognitive achievement, affective development, and psychomotor competence simultaneously. Competencies such as Qur'anic recitation, prayer performance, moral behavior, and religious practice are evaluated according to explicit standards of performance that cannot be meaningfully interpreted through class ranking alone (Pranajaya et al., 2023). Empirical studies examining assessment implementation in Islamic education consistently report that criterion-referenced score processing provides teachers with clearer evidence regarding individual competency achievement, thereby supporting formative assessment, individualized feedback, and differentiated instructional planning (Irmayanti et al., 2024). Comparable findings also demonstrate that

criterion-based interpretation aligns more closely with competency-oriented curricula because educational success is determined by mastery rather than competitive superiority over peers (Nurhayani et al., 2024). Although NRA continues to contribute to specific educational purposes such as student selection and academic competitions, its routine application in classroom assessment appears less capable of representing holistic learning achievement within Islamic Religious Education. These observations position criterion-referenced score processing as the more pedagogically coherent approach for evaluating multidimensional competencies while recognizing the complementary role of norm-referenced interpretation in contexts requiring comparative educational decisions.

The comparative synthesis ultimately indicates that effective score processing should not be determined solely by statistical convenience but by the educational function that assessment is intended to serve. Criterion-referenced assessment offers greater instructional value because processed scores directly reflect competency achievement and facilitate evidence-based instructional improvement, whereas norm-referenced assessment contributes more effectively to classification and selection processes where relative performance constitutes the primary objective (Syafuddin & Herdah, 2025). Recent discussions on classroom assessment similarly emphasize that meaningful educational evaluation requires coherence among curriculum objectives, assessment design, score-processing procedures, and instructional follow-up rather than reliance on isolated measurement techniques (Susilowati et al., 2024). This conceptual integration strengthens the argument that score processing should be understood as an educational decision-making framework instead of a purely computational activity. From the perspective of contemporary pedagogy and learning science, assessment quality is determined not only by measurement accuracy but also by the extent to which processed scores generate valid, equitable, and pedagogically useful interpretations of student learning. This analytical perspective provides the basis for examining the broader pedagogical implications of integrating criterion-referenced and norm-referenced assessment within contemporary Islamic Religious Education, which is discussed in the following section.

Pedagogical Implications for Contemporary Islamic Religious Education

The comparative findings indicate that the effectiveness of educational assessment is determined not only by the quality of measurement instruments but also by the appropriateness of the assessment reference used to interpret learning outcomes. Within competency-oriented education, score interpretation should provide meaningful information regarding students' progress toward predefined learning objectives rather than merely describing their relative position among peers. This perspective reflects the growing emphasis within contemporary pedagogy on assessment as an integral component of learning rather than a mechanism for measuring achievement at the conclusion of instruction. The literature reviewed consistently suggests that assessment practices become more educationally valuable when they facilitate instructional improvement, individualized feedback, and continuous competency development instead of focusing exclusively on student classification (Sternberg et al., 2022). Such a paradigm shifts the role of assessment from a summative judgment toward a formative process capable of supporting sustainable learning. For Islamic Religious Education (IRE), where instructional objectives include knowledge acquisition, character formation, and religious practice, assessment should primarily function as evidence for educational development rather than competitive comparison.

The analytical synthesis therefore positions assessment references as pedagogical decisions that directly influence how learning quality is understood and improved. The analysis also demonstrates that criterion-referenced assessment possesses stronger theoretical compatibility with the principles of competency-based education implemented in contemporary Islamic Religious Education. Current educational reforms emphasize learning outcomes, competency attainment, differentiated instruction, and continuous feedback, all of which require assessment systems capable of identifying individual mastery instead of relative academic standing (Zubaidi, 2022). Studies examining assessment implementation in Islamic schools similarly report that criterion-based evaluation enables teachers to monitor students' progress more systematically because learning achievement is interpreted according to predetermined competency indicators rather than fluctuations in classroom performance (Darmawan, 2024). Such consistency contributes to greater fairness in educational evaluation because each learner is judged according to identical standards regardless of the academic profile of the class. The literature further indicates that criterion-referenced assessment encourages reflective teaching practices by allowing educators to evaluate the effectiveness of instructional strategies based on competency

achievement instead of comparative examination results (Hadi et al., 2024). These characteristics strengthen the pedagogical relevance of CRA for classroom assessment, particularly within educational environments emphasizing mastery learning and student-centered instruction. Consequently, the conceptual integration between curriculum objectives and assessment interpretation becomes a prerequisite for improving learning quality.

Although the present analysis demonstrates the pedagogical advantages of criterion-referenced assessment, the reviewed literature also indicates that norm-referenced assessment continues to perform important functions within particular educational contexts. Comparative assessment remains appropriate when educational institutions require objective differentiation among learners for purposes such as academic selection, scholarship allocation, placement, certification, or participation in competitive programs (Geiser, 2022). Norm-referenced interpretation provides efficient information concerning students' relative standing, allowing decision-makers to identify exceptional performance within large populations. Research investigating grading systems further suggests that comparative evaluation influences learners' expectations, academic motivation, and perceptions of competition, indicating that NRA continues to possess educational relevance when implemented within clearly defined institutional objectives (Liu et al., 2024). Historical critiques nevertheless remind educators that excessive dependence on comparative ranking may unintentionally reinforce inequity and reduce attention to individual learning growth if assessment becomes detached from competency attainment (Brumley et al., 2026). These observations indicate that NRA should be understood as a complementary assessment approach rather than a substitute for competency-oriented evaluation. Its application should therefore remain context-dependent and aligned with educational purposes that genuinely require comparative interpretation.

An equally significant implication emerging from this study concerns the relationship between score-processing techniques and instructional decision-making. The literature synthesis demonstrates that raw scores acquire educational meaning only after being interpreted through an appropriate assessment framework, implying that identical numerical results may generate different pedagogical conclusions depending on the selected score-processing procedure (Rohmah, 2024). Studies examining classroom assessment similarly emphasize that the quality of educational decisions—including remediation, enrichment, differentiated instruction, and competency certification—is directly influenced by the validity of score interpretation rather than the numerical score itself (Azijah et al., 2026). The reviewed evidence further indicates that systematic score processing contributes to greater transparency because teachers can justify assessment decisions using explicit competency criteria or clearly defined comparative standards (Kiswanto et al., 2024). Such transparency strengthens educational accountability while improving students' understanding of their own learning progress. From the perspective of learning science, assessment therefore functions as an interpretive process through which quantitative evidence is transformed into actionable pedagogical knowledge capable of supporting instructional improvement. This interpretation extends the discussion beyond measurement theory toward the broader role of assessment in facilitating meaningful learning experiences.

The multidimensional characteristics of Islamic Religious Education further reinforce the necessity of adopting assessment approaches that accommodate cognitive, affective, and psychomotor competencies simultaneously. Learning outcomes within IRE involve not only conceptual understanding of Islamic teachings but also observable religious practices, moral behavior, and the internalization of Islamic values, dimensions that require explicit competency indicators during evaluation (Pranajaya et al., 2023). Previous investigations into Islamic education consistently demonstrate that criterion-based assessment offers greater flexibility for evaluating these multidimensional competencies because performance standards can be developed for each learning domain while maintaining consistency in score interpretation (Irmayanti et al., 2024). Similar findings also indicate that competency-oriented assessment facilitates the implementation of formative evaluation by providing detailed information regarding learners' strengths, weaknesses, and areas requiring instructional intervention (Nurhayani et al., 2024). Educational assessment within Islamic Religious Education therefore benefits from integrating criterion-based evaluation as the primary assessment framework while reserving norm-referenced interpretation for specific circumstances involving comparative decision-making. Such integration supports a balanced assessment system that values both competency development and institutional accountability without compromising the educational objectives of Islamic learning.

The present study contributes to the literature on contemporary pedagogy and learning science by proposing an integrated conceptual perspective that connects assessment paradigms, score-processing techniques, and instructional decision-making within competency-based Islamic Religious Education. Previous discussions have generally examined criterion-referenced assessment, norm-referenced assessment, and score interpretation as separate themes, whereas the present analysis demonstrates that these dimensions are conceptually interdependent and collectively determine the educational meaning of assessment outcomes. This integrated perspective highlights that assessment quality cannot be evaluated solely through measurement accuracy but must also consider the coherence between learning objectives, assessment references, score-processing procedures, and pedagogical actions derived from assessment results. Such a conceptual framework expands existing scholarship by positioning score processing as an essential component of educational interpretation rather than a purely technical calculation. Future research may strengthen this framework through empirical investigations examining how different score-processing models influence teachers' assessment practices, students' learning achievement, and competency development across diverse educational settings. These directions are expected to enrich the ongoing development of evidence-based assessment while supporting more valid, equitable, and learner-centered evaluation practices within contemporary Islamic Religious Education.

CONCLUSION

This study concludes that criterion-referenced assessment (CRA) and norm-referenced assessment (NRA) represent two distinct assessment paradigms whose differences extend beyond score interpretation to encompass their pedagogical orientation, educational objectives, and implications for instructional decision-making. CRA evaluates students' achievement against predetermined competency standards, enabling educators to determine the extent to which individual learners have mastered the intended learning outcomes regardless of the performance of their peers. Consequently, the interpretation of assessment results under CRA is absolute, competency-oriented, and directly aligned with curriculum expectations. By contrast, NRA interprets learning outcomes through comparison with the performance of a reference group, making the resulting scores relative rather than competency-based. Such an approach is valuable when educational decisions require ranking, classification, or selection among students, although it provides limited information regarding the actual mastery of expected competencies. The comparative analysis also demonstrates that score-processing techniques are not merely statistical procedures but educational mechanisms that shape how learning evidence is interpreted and subsequently translated into instructional actions. The educational value of assessment therefore depends on the coherence between learning objectives, assessment references, score-processing procedures, and the pedagogical decisions derived from assessment outcomes.

The findings further indicate that CRA possesses stronger pedagogical relevance for Islamic Religious Education because its competency-oriented framework accommodates the multidimensional nature of learning outcomes encompassing cognitive, affective, and psychomotor domains. Through criterion-based score processing, teachers are able to provide more meaningful feedback, identify individual learning needs, and implement appropriate remedial or enrichment strategies based on explicit competency standards. Although NRA continues to contribute to educational contexts involving academic competition, student placement, or institutional selection, its application within routine classroom assessment should remain complementary rather than dominant. A balanced integration of both assessment approaches allows educators to address diverse educational purposes without compromising the principles of fairness, competency development, and meaningful learning evaluation. From the perspective of contemporary pedagogy and learning science, assessment quality should be evaluated not only by measurement accuracy but also by its capacity to generate valid, transparent, and instructionally useful interpretations of student learning. The conceptual framework developed in this study provides a theoretical foundation for strengthening competency-based assessment practices in Islamic Religious Education while expanding scholarly discussions concerning score-processing techniques in educational evaluation. Future studies are encouraged to empirically examine the implementation of these assessment models across different educational settings to further validate and refine the conceptual propositions advanced in this research.

REFERENCES

- Azizah, I. N., Nurkholifah, N. I., Rahma, S. N., Putri, M. M., & Udin, T. (2026). Analisis Butir Soal dan Ketuntasan Belajar Ditinjau dari Penilaian Acuan Norma dan Penilaian Acuan Patokan. *AL IBTIDA': Jurnal Program Studi Pendidikan Guru Madrasah Ibtidaiyah*, 14(1), 1-10. <https://ejournal.kopertais4.or.id/mataraman/index.php/alibtida/article/view/6496>
- Brumley, B., Brumley, L., & Morrison, D. (2026). Learning About the Past to Leave it Behind: Eugenics & Norm-Referenced Educational & Psychological Testing. *Critical Questions in Education*: 17, (40). https://www.researchgate.net/publication/400299722_Learning_About_the_Past_to_Leave_it_Behind_Eugenics_Norm-Referenced_Educational_Psychological_Testing
- Burr, S. A., Zahra, D., & Robinson, I. M. (2026). The Difficulty with Setting a Standard Based on Difficulty: The Role of Validity in Determining Assessment Standards. *Education Sciences*, 16(3), 1–11. <https://doi.org/10.3390/educsci16030488>
- Darmawan, H. (2024). The Comparative Analysis Of Learning Outcomes Using Norm-Referenced Test And Criterion-Referenced Test In Islamic Education. *Tarbiyah Islamiyah: Jurnal Ilmiah Pendidikan Agama Islam*, 14(2), 115–126. <https://doi.org/10.36668/jal.v8i1.105.3>
- Gaertner, M.N. (2022). Norm-Referenced Assessment. *Routledge*. <https://doi.org/10.4324/9781138609877-REE13-1>
- Geiser, S. (2022). Eligibility for Admission to the University of California After the SAT/ACT: Toward a Redefinition of Eligibility. *UC Berkeley Center for Studies in Higher Education: Research and Occasional Papers Series*, 2(3), 1–13. <https://escholarship.org/uc/item/87p9j2qf>
- Hadi, M. A. B., Saputra, P., Imamah, A., & Maisarah, I. (2024). The implementation of criterion referenced test approach in analysis of student learning outcomes at Madrasah Ibtidaiyah. *Al-Lahjah: Jurnal Pendidikan, Bahasa Arab, dan Kajian Linguistik Arab*, 7(2), 773-783. <https://ejournal.unwaha.ac.id/index.php/lahjah/article/view/4665/2287>
- Irmayanti, A. P., Remiswal, & Khadijah. (2024). Exploring the Practice of Norm-Referenced and Criterion-Referenced Assessment in Islamic Education. *At-Ta'lim: Media Informasi Pendidikan Islam*, 23(1), 90–104. <http://dx.doi.org/10.29300/attalim.v23i1.4327>
- Kiswanto, D., Arista, D., Fitrah, I. J., Annisa, M. N., & Qomari, N. (2024). Implementasi Penilaian Acuan Norma (PAN) dan Penilaian Acuan Patokan (PAP) dalam Pengolahan Hasil Belajar Siswa. *Risalah Jurnal Pendidikan dan Studi Islam*, 10(3), 1207-1219. https://doi.org/10.31943/jurnal_risalah.v10i3.1088
- Littman, R., Hochman, S., & Kalanthroff, E. (2024). Reliable affordances: A generative modeling approach for test-retest reliability of the affordances task. *Behavior Research Methods*, 56(3), 1984-1993. <https://link.springer.com/article/10.3758/s13428-023-02131-3>
- Liu, J., Wong, M., & Hard, B. (2024). The Impact of Norm- and Criterion-Referenced Grading Systems on Students' Course-Related Expectations. *Journal of the Scholarship of Teaching and Learning*, 24(4), 43–60. <https://doi.org/10.14434/josotl.v24i4.35434>
- Nurhayani, N., Zulhimmah, Z., Efendi, M. E., & Harahap, H. A. (2024). Penerapan Penilaian Beracuan Patokan Dan Norma Pada Pelajaran Sejarah Kebudayaan Islam Di Madrasah Aliyah Swasta Darul Hadits Huta Baringin. *Cognoscere: Jurnal Komunikasi dan Media Pendidikan*, 2(2), 55-63. <https://doi.org/10.61292/cognoscere.168>
- Nurhijatina, H., Larasati, N. J., & Bella, S. (2023). Analisis Classrooms Assessment: Remedial, Pengayaan, Pendekatan Acuan Patokan (PAP) Dan Pendekatan Acuan Normatif (PAN). *Didaktik: Jurnal Ilmiah PGSD STKIP Subang*, 9(5), 3216-3230. <https://doi.org/10.36989/didaktik.v9i5.2276>
- Nurmanto, Y., & Ashari, M. Y. (2023). Penilaian Acuan Patokan Dalam Evaluasi Pendidikan Islam. *Ats-Tsaqofi: Jurnal Pendidikan dan Manajemen Islam*, 5(1), 39-51. <https://doi.org/10.61181/ats-tsaqofi.v5i1.362>
- Pranajaya, S. A., Idris, J., & Abidin, Z. (2023). Integration of Cognitive, Affective, and Psychomotor Domain Scoring in Islamic Religious Education. *Sinergi International Journal of Education*, 1(2), 90-102. <https://doi.org/10.61194/education.v1i2.106>
- Rohmah, F. L. (2024). Pengolahan Hasil Belajar Peserta Didik. *Grata: Jurnal Inovasi Pendidikan*, 1(1), 37-43. <https://ejournal.mejailmiah.com/index.php/grata/article/view/30>

- Sternberg, R. J., Chowkase, A., Parra-Martinez, F. A., & Landy, J. (2022). Criterion-Referenced Assessment of Intelligence as Adaptation to the Environment: Is It Possible, Plausible, or Practical? *Journal of Intelligence*, 10(3), 2–18. <https://doi.org/10.3390/jintelligence10030057>
- Susilowati, E., Shaleh, S., & Zulfatulkhoiriyah, Z. (2024). Analisis Analisis Langkah-Langkah Pendekatan Acuan Patokan (PAP) Asesmen dalam Pendidikan Islam. *Moderasi : Journal of Islamic Studies*, 4(2), 194–206. <https://doi.org/10.54471/moderasi.v4i2.74>
- Syafruddin, L., & Herdah, H. (2025). Interpretasi Hasil Tes Penilaian Acuan Norma (PAN) dan Penilaian Acuan Kriteria (PAK) dalam Evaluasi Pembelajaran Bahasa Arab. *Ta'limi Journal of Arabic Education and Arabic Studies*, 4(2), 245-266. <https://doi.org/10.53038/tlmi.v4i2.281>
- Zubaidi, Z. (2022). Analisis Kebijakan Standar Penilaian Pendidikan Di Indonesia Pada Mata Pelajaran PAI. *Journal of Industrial Engineering & Management Research*, 3(1), 256-267. <https://doi.org/10.7777/jiemar.v3i1.418>