

The Urgency of Objective Score Processing: A Qualitative Study on The Understanding of Scalling System, Score Distribution, and Assessment Standards in Educational Evaluation

Anisa Meilawati^{1*}, Muh. Nur Rochim Maksu¹, Ari Anshori¹, Hana Ayu Safitri¹

¹ Faculty of Islamic Education, Universitas Muhammadiyah Surakarta, Surakarta, Indonesia

Abstract

Objective and transparent grading remains a critical challenge within educational evaluation system, particularly in contexts where inconsistency, subjectivity, and unclear assessment standards affect students' motivation and academic fairness. This study explores the urgency of objective score processing in education by examining teachers' understanding and practices related to scale selection, score distribution, and standardized assessment criteria. Using a qualitative research design, data were collected through in-depth interviews, observations, and documentation analysis in several senior high schools. Thematic analysis was employed to identify patterns in how educators interpret and implement objective grading principles. Findings indicate that while teachers recognize the importance of objective evaluation, variations in understanding of measurement scale (nominal, ordinal, interval, ratio), limited use of score distribution analysis, and inconsistent application of assessment standards contribute to discrepancies in grading outcomes. These inconsistencies often impact students' perceptions of fairness and may influence learning motivation. Furthermore, the study highlights structural challenges such as lack of professional training, minimal feedback mechanisms, and limited technological integration in score processing. This research underscores the need for systematic professional development and standardized assessment frameworks to promote fairness, validity, and accountability in educational evaluation. The study also suggests integrating statistical tools and digital assessment platforms as strategies to enhance objectivity and transparency.

Keyword : *Objective Scoring, Assessment Standards, Score Distribution, Qualitative Study, Educational Evaluation.*

Introduction Section

Educational evaluation plays a crucial role in determining the effectiveness and success of the learning process. Beyond measuring students' academic achievement, evaluation also serves as a basis for decision-making related to curriculum development, instructional improvement, and constructive feedback for both teachers and students. For that reason, objective score processing becomes an essential component to ensure that evaluation results accurately reflect students' abilities and performance in a fair, valid, and accountable manner.

According to the Indonesian National Education Standards Agency (BSNP, 2023), approximately 35% of educational institutions still encounter inconsistencies in grading due to the absence of clear assessment criteria and standardized procedures. This issue is further supported by findings from the OECD's PISA survey, which demonstrate that countries with transparent and objective assessment systems tend to obtain higher average PISA scores—above 500 in literacy and mathematics—while systems affected by subjectivity and inconsistency tend to score below 450. These data indicate that objectivity in grading is strongly associated with the quality of learning outcomes and students' academic performance (Lager & Lavonen, 2023).

The urgency of objective assessment becomes evident when examined from a socio-educational perspective. In 2023, Indonesia recorded a dropout rate of 6.5%. One of the contributing factors was low learning motivation, which students attributed to unfair or subjective evaluations.

A survey involving 1,000 students from various regions revealed that 42% of participants perceived their scores as unreflective of their actual effort and ability. This perception has led to distrust toward educational institutions and has the potential to create wider socio-educational implications if left unaddressed.

Objective score processing in education is far from a simple technical procedure; rather, it involves the systematic integration of psychological, pedagogical, and statistical considerations. Among its main challenges is the need to minimize bias, ensure consistency, and establish scientific accountability in the evaluation process. To achieve this objective, teachers and evaluators must possess adequate understanding of scaling systems, score distribution, and assessment standards. These three components are crucial in organizing, analyzing, and interpreting assessment results.

* Corresponding author: g000214092@student.ums.ac.id

Scaling systems are measurement tools used to categorize and interpret learning outcomes. In educational evaluation, scaling may take the form of nominal, ordinal, interval, or ratio scales—each offering different levels of precision and analytical utility. Research from Universitas Pendidikan Indonesia (2022) indicates that the use of interval and ratio scales in final examinations increases reliability by up to 20% compared to nominal or ordinal scales. Without an adequate understanding of scaling, educators risk misinterpreting scores, which may lead to inappropriate academic decisions (Levy-feldman, 2025).

Score distribution analysis is an equally essential aspect of evaluation. Distribution reflects how scores are spread among a group of students and can reveal whether a test meets fairness and difficulty criteria. A national-level study involving 500 senior high schools found that approximately 40% of final exam results exhibited non-normal distributions.

In 65% of such cases, parents expressed concerns regarding fairness and grading consistency. Abnormal score distributions often indicate instruments that are too easy, too difficult, or insufficiently discriminative, resulting in inaccurate representations of learning outcomes.

Assessment standards function as the normative framework for determining levels of achievement, passing criteria, and competency categories. A Ministry of Education study (Kemendikbud, 2022) reported that the implementation of standardized assessment criteria reduced complaint levels related to grading disputes by 50%. However, around 30% of institutions still lack standardized assessment practices, which contributes to disparities in student learning outcomes and academic recognition.

In practice, objective score processing requires the integration of scaling systems, score distribution analysis, and standardized criteria. This integration ensures that evaluation is not merely a data collection activity, but a systematic process enabling educators to interpret learning outcomes based on empirical evidence and scientific measurement. Technological advancements have significantly supported this need. The use of educational data management systems and statistical software has enabled faster and more accurate score processing.

A UNESCO (2024) report found that institutions employing digital scoring systems demonstrated a 30% increase in efficiency and a 40% reduction in recording errors. However, technology should be seen as a complement rather than a substitute for teachers' fundamental understanding of assessment principles.

In the context of global competitiveness, objective score processing becomes an urgent necessity for improving learning quality and producing competent graduates who can compete at national and international levels. Therefore, strengthening teachers' assessment literacy and promoting standardized, transparent, and accountable evaluation systems should remain a priority within the landscape of education reform.

Method

This study employed a qualitative research approach aimed at understanding teachers' perspectives and practices regarding objective score processing in educational evaluation. Data were collected through in-depth interviews with teachers and evaluators in several senior high schools, complemented by direct observations and document analysis. The use of multiple data collection techniques allowed researchers to obtain comprehensive insights into how scaling systems, score distribution, and assessment standards were interpreted and applied in practice.

The qualitative approach enabled the researcher to explore experiences, challenges, and strategies related to objective grading in a contextual and natural setting. Interview sessions were conducted to reveal teachers' understanding of measurement scales and their reasoning behind score allocation, while observations provided insights into the actual grading practices during evaluation periods. Document analysis was used to review score sheets, assessment rubrics, and internal school guidelines to identify how standards and criteria were implemented.

Data analysis was carried out using thematic analysis. The analysis process began with coding and categorizing relevant information, followed by identifying recurring themes related to scaling practices, score distribution patterns, and the use of assessment standards in evaluation. These themes were then interpreted to explain the urgency and implications of objective score processing within the education system. The findings derived from this analysis provide a contextual understanding of objective-based evaluation and highlight the need for improving assessment literacy among educators.

Result and Discussion

Understanding of Measurement Scales in Educational Evaluation

Measurement scales serve as fundamental tools for categorizing and interpreting learning outcomes. In the context of educational evaluation, scales may take the form of nominal, ordinal, interval, or ratio, each providing different levels of analytical precision. The findings of this study reveal that teachers possess varying levels of understanding regarding measurement scales, which contributes to differences in score interpretation. Teachers who apply interval or ratio scales demonstrate a more structured and accurate approach to grading, allowing them to distinguish performance gaps among students more effectively. Conversely, limited understanding of measurement scales often results in inconsistencies in score assignment, potentially leading to biased evaluation and misinterpretation of student performance (Lewis et al., n.d.).

In addition to technical concerns, measurement scales play a central role in representing learning outcomes based on the intended level of precision. From a theoretical perspective, interval and ratio scales allow educators to compare performance differences across individuals with greater accuracy, while ordinal and nominal scales merely categorize outcomes without providing meaningful distances between categories. In the context of educational evaluation, selecting an inappropriate scale may result in misinterpretation of student performance and limit the validity of pedagogical decisions. Scholars in educational measurement argue that scaling should always align with instructional objectives and the nature of assessment tasks, particularly in summative evaluations where grading carries high stakes (King & Patterson, 2008).

From a practical standpoint, the study found that teachers' preferences toward ordinal scales were frequently influenced by institutional habits and time efficiency. Many teachers viewed scaling as merely a technical requirement rather than a measurement framework that informs evaluation. Consequently, the emphasis placed on scaling remains procedural rather than analytical. This finding aligns with previous studies which suggest that teachers' assessment literacy significantly influences how measurement concepts are implemented in classrooms. Without adequate assessment literacy, teachers rely more on intuitive judgments rather than structured measurement strategies (Fernandes et al., 2023).

Interviews also revealed that teachers faced challenges in distinguishing between ordinal and interval interpretations when grading performance-based tasks. For instance, assigning a score of 80 versus 85 may imply equal performance differences across students, even though ordinal grading typically does not support that assumption. This mismatch illustrates a gap between the theoretical assumptions underlying scaling and the practical realities of school-based assessment. Such inconsistencies open opportunities for subjective interpretation, which contradicts the fundamental objective of grading fairness (Levy-feldman, 2025).

Another significant finding concerns the influence of institutional culture on scaling practices. Teachers working in schools with strong assessment guidelines tended to adopt more consistent scaling systems compared to those operating under flexible or teacher-dependent policies. In these contexts, scaling not only functions as a measurement tool but also as an accountability mechanism that reinforces transparency and comparability of student outcomes. This observation suggests that efforts to strengthen objective score processing should not rely solely on teacher competence, but also on collective policy frameworks at the school level (Lewis et al., n.d.).

Qualitative evidence from participants further indicates that teachers recognize the importance of scaling when dealing with heterogeneous student populations. Scaling enables educators to differentiate learning performance fairly, particularly in settings where academic variability is prominent. Teachers who adopted interval scaling reported increased confidence when communicating scores to students and parents, as the grading process appeared more systematic and defensible. This finding underscores the practical value of scaling in promoting fairness, reducing disputes, and facilitating constructive feedback (Chapman & Cook, 2005).

Finally, the integration of scaling concepts with technology emerged as an important theme in the study. Several participants noted that school-based information systems and digital platforms provided automated scaling features that simplified data processing. However, teachers also cautioned that reliance on digital tools without conceptual understanding could lead to superficial use of scaling without critical reflection. Thus, while technology offers efficiency benefits, conceptual mastery of scaling remains vital for achieving genuine objectivity in educational evaluation.

Score Distribution as an Indicator of Evaluation Quality

Score distribution illustrates how students' assessment results are spread across a group and serves as a critical indicator of the fairness, difficulty level, and discrimination power of evaluation instruments. The findings of this study reveal that non-normal score distributions frequently occur during summative examinations, suggesting that the assessment instruments used may not optimally differentiate students' performance levels. In many cases, the distribution pattern tends to cluster at the higher or lower ends, indicating that the tests may be either too easy or too difficult. Such patterns may weaken the credibility of evaluation and limit its ability to represent true student learning outcomes (Yvette, 2021).

In educational measurement, a normal distribution often reflects assessment instruments that are balanced in difficulty and possess adequate discrimination capacity. Conversely, skewed or bimodal distributions may imply weaknesses in item construction, misalignment between instructional objectives and assessment tasks, or inconsistencies in scoring procedures. These theoretical considerations are consistent with qualitative data in this study, where teachers acknowledged that examinations often produce score clusters without adequate diagnostic value.

Interview results further indicate that teachers rarely conduct distribution analysis due to limited training in statistical interpretation and a lack of institutional expectations concerning data-based evaluation. Teachers expressed that their primary concern lies in assigning final grades rather than analyzing distribution patterns for pedagogical refinement. This practical orientation suggests that distribution analysis is perceived as an additional workload rather than an integral component of score processing, even though it directly relates to evaluation validity (Yassin et al., 2025).

From a qualitative perspective, teachers' limited engagement with score distribution analysis stems from both competence-related and systemic factors. Competence-related challenges include lack of statistical literacy and insufficient exposure to educational measurement theory during teacher training. Systemic factors involve institutional cultures that emphasize administrative compliance rather than analytical reflection. In schools where assessment is treated as a routine rather than an analytical process, distribution analysis receives minimal attention.

Despite these challenges, several participants recognized the potential value of distribution analysis in improving assessment quality. Teachers who explored distribution trends reported gaining insights into item difficulty, instructional gaps, and classroom heterogeneity. Such insights could help refine assessment design, improve fairness, and inform instructional decisions. For instance, an extremely skewed distribution may indicate that instruction did not adequately prepare students for the assessed content, whereas a perfectly flat distribution may suggest that the test lacks discrimination ability (Priyanto et al., 2024).

Moreover, score distribution analysis facilitates communication and accountability among stakeholders. Parents and students often question score disparities, especially when evaluations are perceived as subjective. In such contexts, visualizing score distribution through histograms or statistical summaries can enhance transparency and mitigate conflicts. Teachers noted that distribution analysis helps them justify scores more objectively and reduces misunderstandings regarding evaluation outcomes.

Technological tools also play a role in supporting distribution analysis. Several teachers reported that digital platforms, school information systems, and spreadsheet software provide automatic distribution summaries that can be used for reflection. However, similar to the case of measurement scales, reliance on technology does not guarantee conceptual understanding. Without sufficient assessment literacy, distribution interpretation risks becoming superficial and divorced from instructional and evaluative objectives (Buhari, 2023).

Overall, the findings highlight that score distribution analysis constitutes an important component of objective score processing. Strengthening teachers' competence in interpreting distribution patterns, along with institutional support for data-based evaluation, can enhance the credibility, fairness, and diagnostic relevance of educational assessments.

Assessment Standards as a Foundation for Grading Objectivity

Assessment standards serve as normative benchmarks that define the criteria for academic achievement, passing thresholds, and performance categories. The findings of this study reveal that teachers acknowledge the importance of assessment standards for maintaining fairness and accountability in educational evaluation. However, inconsistencies in standard implementation remain a persistent challenge, resulting in variations among teachers in interpreting and applying the same assessment criteria. These inconsistencies have implications for the comparability of student performance and the accuracy of grades awarded (Prasetyo & Komariah, 2025).

From a theoretical perspective, assessment standards provide a structured framework that guides educators in judging learning outcomes and aligning assessments with instructional goals. Standards-based assessment is expected to ensure that students are evaluated based on predetermined criteria rather than subjective judgments. In practice, assessment standards contribute to reducing personal bias, enhancing uniformity in scoring procedures, and strengthening students' and parents' trust in the evaluation system. Without clear standards, grading becomes vulnerable to both conscious and unconscious bias, potentially undermining the purpose of objective score processing (Literacy et al., 2025).

Interview data also suggest that teachers' familiarity with assessment standards varies significantly. Some teachers demonstrated strong awareness of competency-based rubrics and passing criteria, whereas others relied heavily on intuitive judgments when determining final grades. This disparity often arises from differences in professional development experiences, institutional expectations, and exposure to assessment training. Teachers who received structured training in assessment standards reported more confidence and consistency in grading, highlighting the role of capacity-building initiatives in enhancing evaluation quality.

Institutional support plays an important role in facilitating the implementation of assessment standards. Schools with well-defined assessment policies, rubric guidelines, and monitoring procedures tend to exhibit more consistent grading practices. Conversely, institutions that leave assessment decisions entirely to individual teachers risk creating fragmented evaluation systems with limited transparency. This finding aligns with prior research emphasizing that standardized assessment frameworks contribute to equitable educational evaluation and improved learning accountability.

Teachers also noted challenges in balancing standards-based assessment with contextual and individual differences among students. While standards help ensure uniform expectations, they may create tension when students' abilities, learning backgrounds, or circumstances vary widely. For instance, strict adherence to performance criteria may disadvantage students who require differentiated support or alternative assessment modalities. Thus, teachers often negotiate between maintaining fairness through standards and accommodating diversity through flexibility. This negotiation further illustrates the complexity of score processing in educational settings (History, 2025).

Qualitative evidence also highlights that the absence of clear standards can trigger disputes among stakeholders. Students and parents frequently question score discrepancies, especially in high-stakes evaluation contexts. Teachers reported that standards-based assessment offers a convincing rationale when explaining grades to stakeholders, thereby reducing conflict and reinforcing transparency. Standardization thus serves both pedagogical and relational functions within the schooling process.

Lastly, participants emphasized the importance of aligning standards with curriculum goals and instructional content. Misalignment between standards and teaching practices may diminish the validity of assessment results and lead to evaluation that does not accurately represent learning outcomes. This finding underscores the necessity of continuous review and calibration of assessment standards, ensuring that they remain relevant and responsive to curricular reforms and student needs.

Integration of Measurement Scales, Score Distribution, and Assessment Standards in Objective Score Processing

The findings of this study show that the integration of measurement scales, score distribution, and assessment standards constitutes the core foundation of objective score processing. These three components are interrelated and cannot function independently if the goal of evaluation is to achieve fairness, accuracy, and accountability. Measurement scales determine how learners' performance is quantified, distribution analysis provides insight into the quality of assessment instruments, and standards serve as normative references that guide judgment. Without integration, objective grading becomes difficult to realize in practice (Rumia et al., 2025).

From a theoretical perspective, integration reflects the alignment of assessment design with measurement principles and evaluative criteria. Scholars in educational assessment argue that scaling alone cannot ensure fairness in grading unless it is accompanied by distributional evidence and standards-based interpretation. For instance, interval scaling may provide high measurement precision, yet if it is paired with non-normal distributions or unclear standards, the resulting scores may still become questionable. Integration therefore functions as a quality assurance mechanism that validates the scoring process (Hapsari et al., 2025).

Qualitative evidence also indicates that teachers experience challenges in integrating these components due to limited assessment literacy. Some participants reported that even though they understood scaling and standards, they seldom analyzed score distribution, which reduced the analytical depth of evaluation. Conversely, teachers who engaged in distribution analysis reported greater confidence in interpreting scores and aligning them with competency criteria. This indicates that integration not only strengthens technical accuracy but also enhances pedagogical interpretation.

The integration of these components further influences communication and accountability in educational contexts. When scores are generated through consistent scaling, validated through distribution analysis, and justified with clear standards, teachers are able to communicate grading decisions more transparently to students, parents, and administrators. This transparency reduces ambiguity and minimizes conflicts related to perceived unfairness in grading. In educational environments where accountability plays an increasingly important role, such integration becomes essential.

Institutional support emerged as a crucial factor enabling integration. Schools that provided standardized rubrics, assessment guidelines, and digital scoring tools allowed teachers to apply scaling, distribution, and standards in a more systematic manner. Meanwhile, schools lacking such support tended to rely on individual teacher judgment, resulting in fragmented evaluation practices. These findings support the notion that objective score processing is not merely an individual competence, but a systemic practice requiring organizational frameworks and policies.

The role of technology also intersects with integration. Several teachers used digital platforms to calculate scores, generate distribution graphs, and align assessments with competency indicators. Technology simplifies the integration process by reducing manual calculation errors and enabling reflective evaluation. However, participants emphasized that technology alone cannot guarantee objectivity without conceptual understanding. Digital tools support integration, but theoretical comprehension sustains it.

Overall, integration not only promotes objectivity but also reinforces comprehensive evaluation practices in contemporary education. The alignment of scaling, distribution, and standards ensures that assessment outcomes are credible, meaningful, and pedagogically interpretable. This study suggests that strengthening integration through teacher training, institutional policies, and technological support can significantly enhance the quality of educational evaluation.

Conclusion

This study concludes that objective score processing in educational evaluation requires the integration of measurement scales, score distribution analysis, and standards-based assessment criteria. These three components function interdependently in strengthening the fairness, transparency, and validity of grading practices.

The findings reveal that while teachers acknowledge the importance of objective evaluation, inconsistencies remain in the application of scaling systems, distribution analysis, and assessment standards. Limited assessment literacy, institutional culture, and variations in training contribute to differences in scoring practices among educators. Objective score processing is therefore not merely a technical activity, but a comprehensive evaluative procedure that must be conceptually and methodologically grounded.

Furthermore, the study highlights that objective evaluation cannot rely solely on one component. Measurement scales provide the basic structure of quantification, distribution analysis validates the interpretability of assessment results, and standards act as normative references for decision-making. Without integration, grading becomes vulnerable to subjectivity and bias, undermining the credibility of educational evaluation. Thus, strengthening the integration of these components represents a critical need for improving assessment quality and enhancing learning accountability in contemporary educational settings.

Acknowledgement

The author would like to express sincere appreciation to the participating teachers and school institutions for their cooperation and contributions during the data collection process. Their willingness to share insights and experiences greatly supported the completion of this study. The author also extends gratitude to academic colleagues who provided constructive feedback and encouragement throughout the preparation of this paper. Their support has been valuable in refining the research and its presentation. In addition, the author acknowledges the assistance and academic environment that facilitated this research, particularly in providing access to relevant literature and educational resources. Such support played an important role in strengthening the conceptual and methodological foundation of this study. Finally, appreciation is expressed to family and peers who continuously offered moral support throughout the writing process. Their presence and encouragement have been a meaningful source of motivation in completing this work.

References

- Buhari, M. R. (2023). *Digital Technology and Local Policy : An Evidence-Based Collaborative Model for Sports Talent Identification*. 4(2), 161–175. <https://doi.org/10.58524/jcss.v4i2.903>
- Chapman, G., & Cook, L. C. (2005). *Moving From Assessment To Treatment Of School Refusal Behavior In Youth*. 1(1), 46–51.
- Fernandes, C., Santiago, D. B., & Barbosa, A. (2023). *Mapping Measurement Scales for the Assessment of Learning Environments*. 16(2), 164–179. <https://doi.org/10.5539/ies.v16n2p164>
- Hapsari, T., Rif, M., & Kusumastuti, N. (2025). *Enhancing numeracy literacy through algebra learning materials : A vocational high school context*. 16(01), 409–425. <https://doi.org/10.24042/ajpm.v16i1.27935>
- History, A. (2025). *ISLAMIZATION OF SCIENCE AND ISLAMIC EDUCATIONAL*. 2(2).
- King, C. M., & Patterson, N. G. (2008). *Portfolio assessment : Making connections , guiding change*. 7(3), 4–9.
- Lager, A., & Lavonen, J. (2023). *education sciences Engaging Students in Scientific Practices in a Remote Setting*. 1–12.
- Levy-feldman, I. (2025). *The Role of Assessment in Improving Education and Promoting Educational Equity*.
- Lewis, R. K., Sembiante, S. F., Lewis, R. K., Ed, M., Sembiante, S. F., & Ph, D. (n.d.). *Research and Practice in Transition : Improving Support and Advocacy of Transgender Middle School Students*. 5(1).
- Literacy, D., In, S., & Schools, E. (2025). *The Effectiveness of the Animated Video-Based Tarl Model on*. 9(5), 4568–4582.
- Prasetyo, J., & Komariah, K. (2025). *Assessing the Effectiveness of the Desa Prima Program in Empowering Women Economically*. 12(2), 210–225.
- Priyanto, S., Sahara, S., & Nugraha, H. D. (2024). *TITLE : The Certification Program Evaluation for Students in Vocational Schools Using the CSE UCLA Model Jurnal Kependidikan :*
- Rumia, M., Simorangkir, R., & Deliviana, E. (2025). *Jurnal Pendidikan Progresif Effectiveness of the School Literacy Movement Program : An Impact Study Using the Kirkpatrick Model in Kuningan Regency*. 15(02). <https://doi.org/10.23960/jpp.v15i2.pp8>
- Yassin, A., Bashir, A., Surjono, H. D., Afdal, Z., & Novianto, V. (2025). *Design and assessment of effective multimedia-based courseware for student quantitative data analysis*. 14(4), 3103–3115. <https://doi.org/10.11591/ijere.v14i4.32055>
- Yvette, U. (2021). *REID (Research and Evaluation in Education) Evaluation of the implementation of educational assessment standards at Madrasah Tsanawiyah Modern Islamic Boarding Evaluation of the implementation of educational assessment standards at Madrasah Tsanawiyah Modern Islamic Boarding School*. 7(2).