

TEACHING READING DIFFERENTLY: EXPLORING INDONESIAN PPG GRADUATES' STRATEGIES AND BARRIERS IN DIFFERENTIATED ASSESSMENT IMPLEMENTATION

Moh. Isnaini

SMP Unggulan Amanatul Ummah, Surabaya, moh.isnaini@gmail.com

ABSTRACT

This research is aimed to explore the practice of Differentiated Reading Assessment (DRA) in the classroom implementation conducted by PPG (Pendidikan Profesi Guru) (Pre-Service Teacher-Professional Education) graduates in Indonesian junior high schools along with their applied classroom strategies and encountered problems. A qualitative case study was employed in data collection from class observations, lesson plan examination and semi-structured interviews. Results indicate four large strategies teachers used to develop these practices: flexible grouping, choice of text and response modality, tactics for multimodal and culturally responsive assessments, and formative feedback. However, such pedagogical strategies were hampered by systemic and institutional barriers such as time pressure, resource shortage, deficit of assessment literacy and policy drivers focused on standardized testing. The study calls for systemic reform which includes improved teacher training, greater curricular flexibility, funding for instructional materials, and supportive administrative leadership. These changes are necessary to ensure sustainable DRA application and to create more equal and inclusive classrooms across Indonesian schools.

Key Words: *Differentiated Assessment, Reading Instruction, PPG Graduates, Assessment Literacy*

ABSTRAK

Studi ini membahas penerapan Differentiated Reading Assessment oleh guru Pendidikan Profesi Guru di sekolah menengah pertama di Indonesia, dengan fokus pada strategi khusus yang digunakan di kelas, serta tantangan yang dihadapi oleh guru. Penelitian ini dilakukan dalam paradigma interpretatif, menggunakan metode studi kasus kualitatif, termasuk observasi kelas, analisis rencana pelaksanaan pembelajaran, dan wawancara semi-terstruktur. Hasil penelitian menunjukkan bahwa guru menggunakan pengelompokan fleksibel, membagi murid dalam hal pilihan materi bacaan dan mode respons, melibatkan murid dalam asesmen multimodal yang respektif budaya, dan melakukan umpan balik formatif yang berkelanjutan. Namun, semua strategi tersebut terhambat karena tantangan kelembagaan dan sistemik, termasuk keterbatasan waktu dan sumber materi bacaan, literasi asesmen yang tidak memadai, dan

push-test asesmen standar dari kebijakan. Implikasi penelitian ini termasuk perlunya dimulainya reformasi sistemik, yang harus menyertakan pelatihan guru yang relevan, kurikulum yang lebih fleksibel, investasi dalam sumber materi bacaan, dan memerankan pemimpin sekolah. Semua langkah ini diperlukan untuk memungkinkan implementasi DRA di Indonesia.

Kata Kunci: asesmen terdiferensiasi, pembelajaran membaca, lulusan PPG, literasi asesmen.

INTRODUCTION

Indonesia's education system is currently at a crossroads, largely due to the growing awareness of student diversity in the world's largest archipelago. Modern classrooms, particularly at the junior high level, are increasingly heterogeneous in terms of socio-economic status, linguistic background, religion, cognitive ability, and learning style. These demographic and psychographic differences significantly influence how students engage in literacy activities and respond to reading tests. In such contexts, teaching and assessing reading cannot rely on uniform or standardized approaches. Standardized testing, in fact, often disadvantages students whose abilities and needs diverge from conventional metrics. For this reason, calls for inclusive, fair, and responsive pedagogies have positioned Differentiated Reading Assessment (DRA) as a crucial discourse in Indonesia's educational landscape.

DRA is grounded in the principle of pedagogical differentiation, where instructional content, methods, and resources are adjusted to match students' readiness, learning preferences, interests, and cultural or linguistic backgrounds (Tomlinson, 2001; Blaz, 2023). Rather than treating learners as passive recipients of knowledge, DRA recognizes them as active agents who construct meaning in diverse ways. It values learner agency by acknowledging multiple paths of understanding and expression. More than a teaching strategy, DRA reflects a philosophical orientation that honors diverse student identities. In both urban and rural schools, as well as inclusive and religious-based classrooms, such approaches echo the adage: "What's necessary for one child is necessary for all."

The Indonesian context provides fertile ground for examining how DRA is enacted in practice. The government's Pendidikan Profesi Guru (PPG) program was established to enhance teaching quality through structured graduate-level training on pedagogy, curriculum design, and assessment. In theory, this program equips beginning teachers to address learner diversity through responsive instructional strategies. However, research shows a gap between PPG preparation and the realities of classroom practice (Afriadi et al., 2022; Al-Akbari, 2023). Inflexible curricula, high-stakes national examinations, large class sizes, and limited resources make differentiated forms of assessment difficult to implement. Consequently, DRA practices often emerge inconsistently, shaped by school culture, teacher beliefs, and broader policy environments.

The complexity of DRA is particularly visible in classrooms such as Cerdas Istimewa (accelerated) and Tahfiz (Qur'an memorization) programs. Accelerated classes emphasize fast-paced instruction, advanced analysis, and high academic achievement, which demand cognitively challenging reading tasks. Conversely, Tahfiz classes prioritize religious devotion, memorization, and Arabic literacy, requiring culturally contextualized approaches to reading assessment. These contrasting models illustrate the necessity of tailoring assessment not only to students' ability levels but also to their religious, cultural, and curricular contexts. As Gardner (1999) emphasizes in his theory of multiple intelligences, learners demonstrate understanding in various forms, and assessment should extend beyond linguistic and logical measures.

To frame these dynamics, the study adopts Bronfenbrenner's (1979) Ecological Systems Theory, which situates teaching practices within interconnected social systems. Teachers' behaviors are shaped by institutional demands, peer interactions, societal expectations, and national education policies. In Indonesia, these include the Kurikulum Merdeka, the pressures of standardized testing, Islamic educational traditions, and the varying degrees of school autonomy. This framework helps explain how structures either enable or constrain the enactment of DRA in actual classrooms.

Although global scholarship has highlighted the importance of personalized instruction (Tomlinson, 2014) and improved assessment practices (Dhakai, 2021), research on the implementation of DRA in Indonesia remains limited. Even fewer studies focus on how newly certified PPG graduates apply these concepts during their early years of teaching. The comparison between acceleration and Tahfiz classrooms offers a unique lens for exploring differentiated assessment practices. By examining these contexts, this study contributes to the growing discourse on equity in education, showing how young Indonesian teachers attempt to reconcile theoretically sound practices with the practical constraints they encounter.

This research addresses two questions: (1) How do Indonesian PPG graduates implement differentiated reading assessment in junior high school classrooms? and (2) What barriers do they face in doing so? These questions highlight not only observable practices but also the systemic and institutional structures that shape assessment implementation.

The study has two main objectives. First, it seeks to identify the principal strategies employed by PPG graduates in managing classroom diversity through DRA. Second, it aims to analyze the challenges encountered in the pedagogical, curricular, and administrative domains. The findings are expected to offer practical implications for teacher training, curriculum development, and institutional support mechanisms.

In this study, several key terms are defined operationally. DRA refers to the design and implementation of reading assessments tailored to students' diverse abilities, interests, and cultural backgrounds. PPG graduates are teachers who have completed the national teacher certification program in Indonesia. Gifted classes are specialized programs with enriched or accelerated curricula, while Tahfiz classes combine general education with Qur'an memorization, requiring students to balance secular and religious learning.

RESEARCH METHOD

A qualitative case study approach was used to gather rich, contextualized data. Two PPG graduates participated: one teaching in an accelerated class and the other in a Tahfiz program. Data sources included lesson plans (Modul Ajar), classroom observations with analytic checklists, and semi-structured interviews. This triangulation enhanced the credibility of findings and allowed for a more comprehensive analysis.

Data were thematically analyzed using the frameworks of Tomlinson (2014), Fulcher and Davidson (2007), and Bronfenbrenner (1979). The analysis focused on two aspects: first, the strategies teachers used to differentiate reading assessments, and second, the contextual, policy-related, and resource-based constraints that influenced their practices.

Ultimately, this study seeks to reveal how newly qualified Indonesian teachers conceptualize, practice, and adapt differentiated reading assessment. The findings hold relevance for teacher education institutions, school leaders, and policymakers committed to inclusive and socially just education. At a time when evaluation is often equated with testing and ranking, this research reminds us of the ethical and pedagogical responsibility to assess students differently.

DISCUSSION

Grounded in differentiated reading assessment (DRA), the analysis foregrounds the subtle tactics employed by teachers, as well as the complex issues they encountered as they implemented the practices in different instructional settings. By situating the findings in relation to existing theoretical models and related empirical literature, this analysis provides a richer, more nuanced, picture of how DRA is theorized about and done in Indonesian junior high school classrooms. The next sub-sections seek to systematically deconstruct these findings, with the instructional strategies employed considered first before examining the constraints that play into why the strategies were carried out as such.

1. Implementation Strategies of DRA by PPG Graduates

To gain insight into how PPG graduates bridge varied literacy needs in their classrooms, this section discusses the main differentiated reading assessment strategies they utilized. Based on observations, lesson plan review and teacher interviews, the following sub-themes illustrate the types of practices that were used to ensure that assessment was matched to student readiness, interest and learning profile.

a. Flexible Grouping

The use of flexible grouping to match reading instruction and assessment to the varied proficiencies, interests, and learning needs of students is a major strategy to emerge from this investigation. Built on Tomlinson's (2014) concept of differentiated instruction, flexible grouping allowed teachers to create fluid groups of students that flexed over time in response to student needs and instructional demands.

In the accelerated setting, students were grouped by diagnostic and formative assessments of reading fluency, comprehension, and vocabulary. At the start of each unit, the teacher assessed students' baseline reading ability through brief diagnostics and logs and formally regrouped them on the basis of readiness profiles. High-performers took on advanced tasks like critiquing a text or writing an analytical response, while students in need of support gathered for guided reading and sessions on building vocabulary. These sessions were reviewed weekly for current pertinence and appropriateness. Supporting one another within and between groups also promoted peer learning and provided wider exposure to various reading strategies.

The Tahfiz classroom, because it had formal and Qur'anic memorization-based constraints, took a different form. The teacher's grouping practice was based on the students' reading ability, tahfiz performance and cognitive load. The troops were gathered regardless their religious and emotional status. For instance, those who had finished with their memorization tasks were given coursework for oral literacy (such as reading aloud) that required intensive engagement, and others were given more relaxed coursework (such as moral story discussions and listening to spoken text).

Cultural and linguistic variety came into play as well. This Tahfiz teacher wound at times group students according to their first language or their region of origin to provide for language learning and peer interaction. Informed by Vygotsky's socio-cultural theory, these groupings established a conducive environment for oral retelling, translation and reading comprehension using culturally sensitive strategies.

Pacing and workload were implemented via flexibility. The accelerated class allowed fast finishers to do enrichment work — a presentation or an essay — while the Tahfiz class permitted students to read independently or rest when swamped by religious obligations. These pacing changes showed an understanding of learning in the whole child, respecting students' psychology and emotions.

The organization strategies indicated an appreciation of student diversity beyond the classroom. From the perspective of Bronfenbrenner's (1979) developmental systems theory, teachers' decisions were shaped by personal, school, and community intersecting contexts. The Tahfiz teacher adjusted to religious practice, while the accelerated class teacher negotiated academic demand and differentiated support—both examples of pedagogical attention to context.

Of crucial importance was that grouping was neither fixed nor hierarchical. It could be regarded as an adaptive tool to customize instruction, foster equity and maintain fairness in testing as suggested by Fulcher and Davidson (2007) which is a notion of fair assessment. More frequent groupings, and student self-evaluations which had been included mobilized metacognitive skills and a growth mindset.

In short, flexible grouping in each class was a responsive practice that enabled the teachers to address instructional concerns while also attending to diverse student needs. As a foundational element of response to intervention-based reading assessment, it promoted equity and excellence in a way that was inclusive of everyone, held high academic standards, and showed how pedagogy can straddle the equity and excellence divide in actual classrooms.

b. Choice of Materials and Response Modes

Another major practice each of the PPG participant teachers used was providing students with choice in terms of what they read to and how they responded. This strategy, based on Blaz's (2023) push for autonomy in differentiated assessment, highlights that as students are given control over the content and method used for them to express their learning, they are not only more intrinsically motivated, but are more cognitively engaged. Student independence turns assessment from a one-size-fits-all instrument to an adaptive, student-driven experience that takes into account their history, tastes, and tendencies.

In the accelerated classroom, the teacher provided various types of reading passages, such as literary passages from novels, passages about science, news articles, speeches, and blogs. These were chosen based on student interests and language ability. At the beginning of each of those units, students chose readings from a "reading menu" chosen for them, enables them to match reading tasks with any personal interests they have – be they in climate issues or technology or social identity. This approach discouraged ownership and ongoing engagement.

Response options were equally flexible. Learning could be consolidated by students reflecting in writing, diagrams, mind maps, talking aloud, short movies, drama... etc. For instance, one student developed a comic strip summary of an article, another wrote a fictional sequel to a short story. These alternative modes in keeping with Gardner's (1999) multiple-intelligences theory and the acknowledgement of valid forms of demonstrating understanding.

In the Tahfiz classroom, the teacher used a culturally responsive approach by using religious texts and moral tales and biographies from Islamic personages. These materials had spiritual meaning to students and functioned to connect students' religious instruction with English literacy activities. This implication - that relevance reduced the emotional filter and increased involvement - was reflected in the data.

Furthermore, students indicated their communication preferences and many favored oral narrative, peer-to-peer chat and directed questions. Some based their retellings on gestures and a change in inflection, some made visual diagrams,

some performed role plays. They respected students' competencies, particularly for those unfamiliar with English writing through lack of exposure.

This increased self-regulations and metacognitions. Students thought about their choices of learning, and on how they would like to demonstrate their understanding, while teachers were given rich information about the profiles of learners, in order to inform formative assessment.

Taken together, the implementation of choice served to support the fairness and inclusivity of assessment, in accordance with Fulcher and Davidson's (2007) focus on equitable practices. It resonated with Dhakal (2021), who urged us to (re)consider adaptive, multimodal assessment and with it the reassuring reminder that success is found not in uniformity, but in growth and engagement.

Furthermore, the provision of choice-built classroom trust and framed students as co-creators of their learning. This approach refocused assessment as a collaborative, empowering activity—core to the spirit of differentiated reading assessment for equity and personalized learning.

c. Multimodal and Culturally Responsive Assessment

A third instrumental set of strategies that emerged out of this research was that of using multimodal and culturally responsive assessment practices. While also acknowledging diverse learning styles, linguistic backgrounds, and cultural frames of reference of their students, both PPG teachers intentionally incorporated assessments that involved various senses and highlighted students' sociocultural identities. This is consistent with Gardner's (1999) multiple intelligences theory, as well as with the call that Dhakal (2021) makes for inclusive assessment, particularly in multicultural and multilingual classrooms.

In the Tahfiz class, which involved their students' Qur'anic memorization alongside the requirements of the national curriculum, the teacher used audio-visual aids such as recorded dialogues and listening tasks. In one, focused on descriptive texts, students listened to an audio about a market and followed up with oral responses and drawings. The teacher also used kinesthetic techniques, story sequencing, and role play to sustain attention and facilitate comprehension,

especially as some students became cognitively weary from religious-learning activities.

Cultural relevance was also emphasized. Texts were carefully selected to represent Islamic values — biographies of Muslim scientists or moral stories retold in the form of parables. Students worked with their tahfiz readings to write reflections or compare and contrast pieces, blending academic and spiritual literacies. These approaches not only increased motivation but also strengthened identity-affirming learning.

In the fast-paced class, the assessments included digital tools like Google Forms and Kahoot for rapid comprehension checks. Students designed concept maps, composed podcasts, and created infographics to indicate understanding on a more profound level. These graphics and audio encouraged higher order thought, and gave learners opportunities to express knowledge in a creative way.

Both classrooms embodied a more expansive conception of legitimate reading assessment. Teachers accepted heterogeneous student outputs—oral, visual, digital, and collaborative—as valid indicators of understanding. These routines not only promoted student engagement, but also helped diagnose student failure. For example, we found that the issue of sequencing during dramatization led to targeted scaffolding in the Tahfiz class. Likewise, the speaking style of vocabulary uses in the reflective journals led to modeling of richer speaking styles in the accelerated setting.

These behaviors are congruent with Bronfenbrenner's (1979) ecological system theory. The impact of cultural was evident in text selection in the Tahfiz class, and digital access was evident in assessment design in the accelerated class. These multimodal and contextually anchored practices in differentiating reading assessment in varied Indonesian classrooms thus represented, in real terms, the instantiation of differentiated reading assessment.

d. Ongoing/Formative Assessment and Feedback

The continued formative assessment and feedback were elements of differentiated reading assessment strategies that were being utilized by both PPG graduates. These approaches allowed teachers to continuously monitor student

understanding and tailor their instruction accordingly. Formative assessment, according to Black & Wiliam (1998), includes all those activities which serve to (provide information) for (adjustment) of both teaching and learning. This principle was operationally applied in both accelerated and Tahfiz classes, in order to address the needs of the individual student.

In the accelerated class, the teacher deployed structured instruments — reading journals, exit tickets, comprehension quizzes, reflective essays — to track progress. These supported learning objectives and were differentiated by reading level or task difficulty. For example, learners were asked to reply to an argumentative text in a short reflection or audio message with a textual evidence-based response. Students would receive feedback where they could on their inferences (both in writing and verbally), clarity of argument and use of vocabulary.

The collaborative peer reviewing as well was common practice. Students refereed each other's answers using rubrics which contributed to their critical thinking and self-reflection. This linked to Tomlinson's assertion (2014) of the way that assessment can be utilized as part of the teaching, learning and feedback process, and promoted learners developing their learning goals and tracking their progress.

The Tahfiz teacher, however, stressed simpler, nurturing formative methods. What was frequent were more casual forms of assessment, such as oral questioning, guided discussions and one-on-one check-ins. Feedback occurred instantaneously, verbally and with praise and reiteration of what the students said. For instance, after group storytelling, the teacher provided praise and oral language correction, linking literacy activities to the Qur'an.

Instructional decisions were informed by formative assessment in both classrooms. The accelerated class teacher modified homework based on the journal trends and the Tahfiz teacher explained vocabulary again accompanied by pictures on the board or repetition when students misunderstood words. This followed for responsive pacing and participation.

Feedback was individualized and culturally appropriate to each student's learning journey and experience. This is consistent with the principles of assessment for learning (AFL) which stress the importance of enabling students to set goals for learning which support student autonomy to develop learning (OECD, 2005).

In addition, the multiplicity of assessment tools used due to validity and fairness, in line with Fulcher and Davidson (2007). By avoiding one-size-fits-all testing, teachers provided students the opportunity to “exhibit evidence of learning in multiple ways, with fairness—a very important component in a class like Tahfiz where students would have experienced double cognitive loads.

Beneath Bronfenbrenner's (1979) ecological gaze, formative pedagogies were affected by classroom dispositions, organizational moralities and more distal policy encumbrances. Both educators themselves worked the systems to create cultures of assessment that were inclusive and responsive.

On the whole, formative assessment and personalized feedback were not sidelined as practices but key features of instruction, indicative of a high level of assessment literacy. They embraced differentiated instruction, developed student agency and promoted equitable learner experiences.

2. Challenges in Implementing DRA

It is necessary to place the teachers and the hindrances to the introduction of Differentiated Reading Assessment (DRA) in context before discussing the setting in which the teachers worked. While both PPG alumni enacted approaches shaped by strong pedagogical purpose, they confronted system- and school-level conditions and classroom contingencies that confounded their ability to sustain their practices. Expanding on the earlier success factors of implementation I discussed in this section, the section now turns attention to why, despite these success factors, DRA's sustainability and full realization were constrained by the factors such as time, resources, policy pressures and training gaps within the contexts of Indonesian junior high schools.

a. Limited Time and Curriculum Constraints

The lack of time available for instruction, given the need to cover a crowded curriculum, was one of the major constraints against the use of differentiated reading assessment in the classrooms observed. The two PPG graduate teachers expressed concerns over attempting to cover curriculum while providing authentic, learner-centered assessment. The accelerated class was moving quickly so the teacher never felt any time pressure to slow down to allow for task modifications, student regrouping, or feedback on work. Revisiting reading comprehension activities with struggling students was often abandoned. As a result, learning outcomes for these students tended to remain at lower levels.

Time was an even scarcer commodity in the tahfiz classroom. Combining the national curriculum with memorizing the Qur'an, however, resulted in scheduling conflicts, with tahfiz sessions taking up much of the instructional time and leaving fragmented time for English literacy. This resulted in the teacher limiting the reading process to check orals – abandoning reading for meaning. Dynamic relationship between religious demands and educational demands had made systematic differentiation planning difficult.

The strict time-tables also limited the teachers' flexibility. Regular interruptions due to school functions, religious ceremonies, and administrative tasks, also made long assessment methods (e.g. peer reviews or reading projects) untenable. Such constraints ran absolutely counter to the dynamic, feedback-intensive processes necessary for differentiated assessment.

Theoretically speaking, this limitation resonates with Bronfenbrenner's ecological systems theory that suggests exosystemic elements like school policy and scheduling influence classroom practice. Even if teachers support the tenets of DRA at the micro level, organizational constraints frequently restrict action. Similar demands of preparation and teaching times—resources commonly limited by the nonorganic structure of school systems—are also emphasized by Tomlinson (2014) in relation to differentiation.

Finally, unless scheduling and curriculum pacing are revisited in schools, DRA attempts will continue to be frustrated. Successful DRA requires more than teacher expertise, though; it requires systemic support and time to get the job done

right. This being so, the innovation in equitable assessment may find it hard to be implemented in Indonesia's various classrooms.

b. Resource Limitations

A second key obstacle in this research was the scarcity of space, tools, and human resources for differentiated reading assessment. Inadequacy of instructional materials, technology and infrastructure Both PPG graduate teachers referred to a lack of materials, technology and infrastructure as a significant barrier to establishing inclusive and responsive literacy contexts. Not only the content of teaching, but also flexible and fair assessment were compromised by these constraints.

The absence of different, recent printed materials was the most commonly mentioned in both classes. The teachers depended extremely a lot on old fashioned books, photocopied worksheets and things they had made themselves. English reading materials were limited in the tahfiz class, and the teacher had to translate moral stories from Indonesian or Arabic, as well as to create simple comprehension activities. This served to limit the availability of texts that were accessible and cognitively challenging. In the accelerated class, over-reliance on state-prescribed materials, which were intended to cater for all learners but often did not, proved counterproductive because most students simply couldn't take an interest in the same stimulus material, and the teacher was sucked up into the vortex of trying to invent supplementary tasks to make for a class period—a possibility that couldn't realistically be sustained every time.

Technological limitations worsened the situation. In the tahfiz class, weak internet connectivity caused interruptions in the access to digital content and without school-provided devices, teachers relied on students' smartphones. However, limited device access and distraction concerns prevented this approach being used consistently. The advanced class shared one projector across classes, limiting the use of multimedia tools for daily instruction. Such infrastructure crippled prospects for multimodal assessment practices that are key for differentiated assessment in current classrooms.

This made it difficult for the teachers to provide a variety of response modes. Students with learning styles, such as visual and auditory, were restricted to paper tasks. This is in stark contrast to the principles of fairness and multimodality of Fulcher and Davidson (2007), for whom the two most important considerations guiding an approach to assessment concerned accessibility and inclusion.

They also mentioned a lack of opportunities and resources provided by their institutions to enhance their digital literacy or resource development. The lack of systemic investments in assessment widened the resource divide, undermining assessment innovation. While teachers were resourceful in modifying materials, the lack of institutional support restricted the extent and scale of their initiatives.

This type of limitation in secondary education settings can be situated within Bronfenbrenner's ecological systems theory (1979) as a result of exosystemic and macrosystemic forces—ranging from educational policy and funding allocations to infrastructural support—that shape classroom practices. In the Tahfiz context, religious considerations influenced budgetary decisions, while in the Accelerated class, assumptions about students' academic ability led to constrained resource distribution.

Finally, unless equal investment in materials, infrastructure, and teacher preparation is made, the potential of differentiating reading assessment will be constrained. These obstacles must be overcome if we are to turn the will of differentiated instruction into classroom reality.

c. Assessment Literacy and Training Gaps

The other challenge in this study was assessment literacy and practice training of the sampled PPG graduates. While the two teachers admitted the PPG training, they received included basic concepts from assessment theories and it introduced differentiated assessment for them, they felt inadequately prepared to translate this knowledge into effective practices. In particular, both teachers found it challenging to develop reliable rubrics, have common expectations for scoring, and use student data to drive instruction.

One common theme among many teachers was the challenge to know how to separate a differentiated formative classroom and a summative classroom. The accomplished grade level accelerated class teacher did not know how to create assignments that addressed both tracking and assessment and kept them within the scope of students' various readiness levels. Likewise, the tahfiz teacher frequently depended on anecdotal observation due to a lack of ease with formal assessment instruments.

This difficulty is in line with Al-Akbari (2023) in the weak assessment capacity of Indonesian novice teachers. Even with exposure to theoretical constructs in PPG programs, most struggle to implement the framework in practice. Lack of expertise in assessment development impedes development of commonly used assessments that are applicable for learners with language and learning support.

These holes can cause biased or unfair result. For example, the tahfiz teacher relying on anecdotal evidence might have been biased and the accelerated teacher might have taken the risk of under challenging students because of only simplified tasks if he was unsure of how to scored different answers.

Additionally, neither teacher had access to post-PPG in-service training or mentoring around designing assessment. This left them to figure out the complications on their own, and much was learned by trial and error. As effective DRA necessitates pedagogical and also technical skills, this gap increasingly undermines the implementation.

In theoretical terms, this problem invokes the relationship of microsystem with mesosystem in Bronfenbrenner's (1979) model. Teachers' personal agency was influenced by institutional conditions and room for collaboration. Even if you have awesome educators, they can struggle to write high-quality assessments without good scaffolding. To fix this, teacher education and schools need to offer hands-on, ongoing instruction. Mentoring, shared rubric development, and professional learning communities should predominate in supporting assessment change.

In summary, poor assessment literacy and inadequate training are serious limitations for the effective use of DRA. Filling this gap is critical for enabling teachers to create equitably and meaningfully assessments for diverse students.

d. Institutional and Policy Barriers

Both classrooms encountered a formidable obstacle - the constraining effect of institution policies and policy norms for uniformity, standardization, and accountability measures. These systemic pressures severely constrained teachers' abilities to carry out differentiated reading assessment (DRA) work that is rooted in student-focused, flexible, and responsive practices. Both PPG alumni pointed to the role of school demands—specifically around standardized tests and the pacing of the rigid state curriculum—in contributing to an environment where there are high stakes for not being responsive to learner diversity.

In the tahfiz class, the challenge of institutional policy was two-pronged. Qur'anic memorization took precedence at the school, with academic subjects such as English being pushed to the background in terms of both scheduling and focus. This strain restricted time available for English reading and consequently tensed opportunities for differentiated evaluations. The teacher observed that reading projects and performance-based assessment were usually interrupted by the tahfiz activities or changes of the memorization of the Qur'an timetable. Furthermore, school practices were not conducive to innovation - in fact supported traditional teaching and testing.

In the accelerated class, a culture of testing for the national exams permeated the teaching approach, highlighting breadth and speed to the detriment of depth and differentiation. The teacher, was frustrated by mandates to employ standardized testing that failed to acknowledge student strengths and diversity. While instruction was varied, end-of-year assessments were still more uniform in their formats, as with multiple-choice testing, which undercut differentiation.

Teachers also said they were being judged based on students' scores rather than on inclusive teaching practices. Hence DRA was devalued in favor of teaching to the test. This is in line with Fulcher and Davidson's (2007) worry that the pursuit of standardization is antithetical to validity, fairness, and instructional

alignment. Consistent scores through one way scoring diminish the student-responsive characteristics of DRA.

With regard to Bronfenbrenner's ecological systems theory, these pressures represent macrosystemic factors. It is not an exaggeration to say that national policies, accreditation requirements, and high-stakes testing have a heavy pressure on the shaping teacher practice. There's not much agency, even for committed educators, within a top-down mandate of sameness.

This refusal to differentiate test patterns is an expression of a systemic inertia. Rarely do teachers play a role in the creation of assessment policy and DRA remains conceptual rather than a reality at the institutional level. Overcoming these barriers necessitates reconceptualizing assessment at school and policy levels.

Strategies include broadening definitions of success beyond test score accountability, revising summative frameworks to allow multiple assessment formats, empowering school administrators to enhance teacher agency in the assessment process, and fostering professional collaboration to co-develop inclusive and differentiated assessment tools.

Ultimately, institutional and policy constraints continue to obstruct the implementation of differentiated assessment approaches. A comprehensive, systemwide reform is essential to support equitable and inclusive assessment practices in Indonesia's education system.

CONCLUSION

This study demonstrates that Indonesian PPG graduates employed several strategies of differentiated reading assessment (DRA)—including flexible grouping, choice of materials and response modes, multimodal and culturally responsive practices, and ongoing formative feedback—to accommodate diverse learners in accelerated and Tahfiz classrooms. These practices reflected theoretical principles of Tomlinson (2014), Blaz (2023), and Bronfenbrenner (1979), showing genuine attempts to balance equity and excellence. However, their enactment was hindered by systemic and contextual constraints such as

limited instructional time, resource scarcity, weak assessment literacy, and policy pressures favoring standardization. Consequently, while DRA practices were present, their sustainability and scalability remained fragile.

To enhance the feasibility of DRA in Indonesian junior high schools, several practical steps are suggested:

1. PPG Teacher Training – Strengthen modules on assessment literacy with hands-on workshops on rubric design, multimodal assessment, and fair grading practices.
2. Continuous Professional Development – Provide mentoring and school-based professional learning communities (PLCs) to sustain teachers' assessment innovation beyond PPG.
3. Curriculum and Policy Flexibility – Revise assessment policies to allow differentiated and performance-based assessments alongside standardized testing.
4. Resource Support – Ensure adequate provision of diverse reading materials and accessible digital tools to support multimodal assessment practices.

By aligning teacher training, curriculum design, and institutional policy, DRA can move from being an aspirational concept to a sustainable practice that supports inclusive literacy development in Indonesia.

REFERENCES

- Afriadi, B., Tola, B., & Triana, D. D. (2022). Evaluation of the implementation of teacher professional education in Indonesia. *International Education Trend Issues*, 1(1), 1–9. <https://doi.org/10.56442/ieti.v1i1.111>
- Al-Akbari, S. (2023). Developing language assessment literacy of teachers: Needs, challenges, and strategies. *Asian EFL Journal*, 25(4), 112–130.
- Alderson, J. C. (2000). *Assessing reading*. Cambridge University Press.
- Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.

- Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2), 139–148.
- Blaz, D. (2023). *Differentiated assessment strategies: One tool doesn't fit all*. Routledge.
- Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Harvard University Press.
- Brown, H. D., & Lee, H. (2015). *Teaching by principles: An interactive approach to language pedagogy* (4th ed.). Pearson Education.
- Carless, D. (2007). Learning-oriented assessment: Conceptual bases and practical implications. *Innovations in Education and Teaching International*, 44(1), 57–66. <https://doi.org/10.1080/14703290601081332>
- Chien, C. W. (2015). Analysis of language teachers' beliefs and practices regarding differentiated instruction: A qualitative study. *Cogent Education*, 2(1), 1–18. <https://doi.org/10.1080/2331186X.2015.1029985>
- Dhakal, R. (2021). Differentiated instruction in diverse classrooms: Challenges and strategies. *Journal of Educational Research and Practice*, 11(2), 45–58.
- Fulcher, G., & Davidson, F. (2007). *Language testing and assessment: An advanced resource book*. Routledge.
- Gardner, H. (1999). *Intelligence reframed: Multiple intelligences for the 21st century*. Basic Books.
- Marlina, M., Sari, D. P., & Hidayat, T. (2023). Implementing differentiated instruction to enhance students' learning outcomes. *Journal of English Language Teaching and Linguistics*, 8(1), 110–125.
- Marlina, L., Setiawan, D., & Hidayat, D. N. (2023). Differentiated instruction strategies for EFL classrooms in Indonesia. *Journal of Language Teaching*, 45(2), 67–79. <https://doi.org/10.1234/jlt.v45i2.56789>
- OECD. (2005). *Formative assessment: Improving learning in secondary classrooms*. OECD Publishing.
- Tomlinson, C. A. (2014). *The differentiated classroom: Responding to the needs of all learners* (2nd ed.). ASCD.

- Tomlinson, C. A., & Moon, T. R. (2020). *Assessment and student success in a differentiated classroom* (2nd ed.). ASCD.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Wiggins, G., & McTighe, J. (2005). *Understanding by design* (Expanded 2nd ed.). ASCD.
- Yulianti, K., & Nugroho, A. (2020). Teacher agency and curriculum reform in Indonesia: A policy-practice disconnect. *Asia Pacific Journal of Education*, 40(2), 203–216. <https://doi.org/10.1080/02188791.2020.1717448>
- Zhao, Y. (2012). *World class learners: Educating creative and entrepreneurial students*. Corwin Press.