INDONESIAN JOURNAL OF CURRICULUM
AND EDUCATIONAL TECHNOLOGY STUDIES
2026, Vol. 14, No.1 , 1-21, DOI: https://doi.org/10.15294/ijcets.v16i1.30475

Rethinking Curriculum Planning through Backward Design: A Qualitative Study of Understanding by Design (UbD) in an Inclusive School Context

Alisa Faith Kurniawan[1], Niam Wahzudik[2]

[1,2]Faculty of Education and Psychology, Universitas Negeri Semarang, Semarang, Indonesia.

Abstract

This study examines the implementation of Understanding by Design (UbD) within an inclusive school context in Indonesia, focusing on tensions between centralized curriculum standards and localized pedagogical values. The research examines how inclusive schools negotiate backward-design principles amid dual accountability demands. Using a qualitative case study approach, data were collected through autoethnography, semi-structured interviews, classroom observations, and curriculum document analysis across two hubs of a progressive inclusive school network in different regional contexts. Findings reveal three interconnected tensions across UbD stages: (1) goal-setting challenges arising from navigating national curriculum standards alongside the school's internal values framework (Citta); (2) an assessment implementation gap in which higher-order cognitive objectives documented in curriculum plans were consistently enacted at lower cognitive levels due to structural constraints; and (3) learning plan localization challenges reflected in the uneven implementation of the school's 5M pedagogical framework. These tensions collectively illustrate what this study conceptualizes as implementation pragmatism. The study contributes theoretically by extending UbD scholarship into inclusive education contexts within the Global South, highlighting the limits of backward design under conditions of heterogeneity and regulatory pressure. Practically, the findings inform organizational and pedagogical practices, as well as school leaders, about the need for systemic alignment among inclusive curriculum aspirations, assessment structures, and classroom realities.

ARTICLE HISTORY

Received 6 January 2026
Accepted 3 March 2026

KEY WORDS

Understanding by Design (UbD); Inclusive Education; Curriculum Implementation; Assessment Implementation; Indonesian Education

INTRODUCTION

Inclusive education has gained global recognition as essential for ensuring equitable learning opportunities for all students, including those with diverse learning needs (Azuka et al., 2024). Educational systems must evolve to accommodate natural variation in human brain function and cognition, recognizing that every individual has unique patterns of thinking, learning, and processing information (Hamilton & Petty, 2023; Mpolomoka et al., 2025). Around 15–20% of the world's population is neurodiverse, making it imperative for educational institutions to recognize these learners' specific needs and truly include them to enrich their learning experience (Le Cunff et al., 2025)

However, conventional curriculum approaches often fall short in addressing the wide spe-

✉Corresponding author: alisafaith.k@gmail.com

ctrum of learners in inclusive classrooms, particularly neurodiverse learners who struggle in traditional educational settings designed primarily for neurotypical students (Videla et al., 2025). Traditional education systems fail to recognize that neurodiversity represents natural variations rather than deficits, overlooking unique cognitive strengths such as creativity, problem-solving skills, and attention to detail (Cole et al., 2024; Nweje, 2025). Inclusive instructional design must recognize and support diverse learning styles and adapt teaching methods, materials, and assessments to provide equal learning opportunities for all students.

The Understanding by Design (UbD) model offers an alternative through its backward design principle: (1) beginning with learning outcomes, (2) followed by assessment evidence, and (3) concluding with instructional planning (Wiggins S McTighe, 2005). While UbD has been applied widely in structured curriculum settings, its implementation within inclusive contexts, especially in Indonesia, remains underexplored (Nur Safitri & Putri Fatmawati, 2024). Recent research shows promise: found that Indonesian inclusive schoolteachers with strong UbD understanding demonstrated higher independence and creativity in designing mathematics evaluation (Karim, 2024; Pramesti & Dewi, 2023a). Wulandari (2024) emphasized UbD's alignment with 21st-century education, though practical challenges persist in shifting from test-focused to understanding-driven planning.

A critical gap remains in understanding how inclusive schools interpret and enact Understanding by Design (UbD) under conditions of learner heterogeneity and contextual constraint. While UbD's structured emphasis on alignment supports curricular coherence, inclusive classrooms require ongoing pedagogical responsiveness, creating tensions between design intentions and enacted practice (Florian & Black-Hawkins, 2011; Wiggins & Mctighe, 2005). Existing studies emphasize the importance of teacher beliefs and inclusive commitment (Azuka et al., 2024; Fawns et al., 2025), yet they offer limited insight into how such commitments are mediated by assessment demands, institutional accountability, and locally embedded pedagogical values. Although recent research has examined UbD implementation within Indonesian educational contexts (Dari et al., 2024; Nur Safitri & Putri Fatmawati, 2024), these studies largely foreground curriculum planning and instructional structure, leaving unexplored how inclusive values are negotiated at the level of assessment enactment and classroom decision-making. Consequently, little is known about how centralized UbD frameworks are pragmatically adapted within inclusive Indonesian schools, particularly when global curriculum models intersect with national mandates and locally situated pedagogical traditions.

LITERATURE REVIEW

A. Understanding by Design

The Understanding by Design framework, developed by Grant Wiggins and Jay McTighe in the late 1990s, emerged in response to persistent challenges in curriculum design, where educators struggled with content overload and superficial learning (Wiggins & McTighe, 2005). UbD's theoretical foundation draws on constructivist learning theory and Bloom's revised taxonomy, emphasizing that curriculum design must create conditions for transferable understanding that goes beyond content recall. This represents a fundamental shift in positioning understanding as the primary educational goal (Wiggins & McTighe, 2005).

The backward design process consists of three sequential stages that systematically align learning goals, assessment, and instruction: (1) Identify Desired Results, it requires educators to clarify learning priorities by distinguishing between content worth being familiar with, content important to know and do, and enduring understandings that warrant deep, lasting comprehension; (2) Determine Acceptable Evidence, it establishes how students will demonstrate their understanding before instruction begins; (3) Plan Learning Experiences and Instruction, it involves designing instructional activities and resources that enable students to develop targeted understandings and successfully complete assessments (Wiggins & McTighe, 2012).

Research on UbD implementation reveals both significant benefits and persistent challenges across diverse educational contexts (Aslam et al., 2024; Cole et al., 2024). Studies demonstrate that backward design improves instructional clarity, strengthens alignment between objectives and assessments, and enhances teachers' ability to articulate learning goals (Roth, 2007). Teachers implementing UbD report increased confidence in promoting student understanding and greater intentionality in instructional decision-making (Pramesti & Dewi, 2023b; Roth, 2007). Students in UbD-designed courses show improved performance on transfer tasks and deeper engagement with essential questions compared to traditional curriculum approaches (Mctighe & Wiggins, 2012.)

However, successful implementation requires substantial support structures. Time constraints represent the most frequently cited barrier, as backward design demands significant upfront investment in curriculum mapping and assessment development (ÇeliKman Hanratty & EveyiK Aydin, 2023). Teachers transitioning to UbD often struggle with identifying truly essential understandings versus covering all content, particularly when facing external pressure from standardized testing requirements (Joshi, 2021; Tumlos-Castillo, 2015). Effective implementation necessitates sustained professional development, administrative support for collaborative planning time, and ongoing coaching to refine curriculum designs (Joshi, 2021; Tumlos-Castillo, 2015). Research indicates that schools achieving sustained UbD implementation establish professional learning communities focused on examining student work and revising units based on assessment evidence (Joshi, 2021).

The UbD framework aligns strongly with inclusive education principles and Universal Design for Learning (UDL) by providing multiple pathways to demonstrating understanding (Flood et al., 2025). The backward design process naturally accommodates diverse learners by focusing on essential understandings while allowing differentiated instruction in how students access content, process information, and demonstrate learning (Tomlinson & McTighe., 2006). Research indicates that "an enthusiastically inclusive approach makes a difference for students in the quality of learning and their interest in learning" (Florian & Black-Hawkins, 2011). The framework's emphasis on authentic assessment and multiple demonstration methods particularly benefits learners with diverse needs, providing alternatives to traditional written tests (Ralabate, 2016)

Recent qualitative research has established "a holistic framework for understanding the challenges in inclusive education, the core mission of inclusive schools, and the essential qualities of teachers" (Ainscow, 2020), highlighting the importance of flexible curriculum structures. UbD's six facets create space for varied learning profiles, as students can demonstrate understanding through different combinations of facets based on their strengths (McTighe & Wiggins, 2012). Nevertheless, challenges remain in balancing common learning goals with individualized supports, particularly when designing performance tasks that are both rigorous and accessible (Ralabate, 2016). Contemporary research explores the integration of technology with UbD principles to enhance accessibility and global applicability across diverse educational contexts, though systematic research on UbD implementation in inclusive classrooms remains limited.

B. Understanding by Design in Indonesian Educational Contexts

The Understanding by Design framework entered Indonesia through adoption, gaining momentum following the 2013 Curriculum's shift toward competency-based learning, and was reinforced by the Merdeka Curriculum's emphasis on curriculum flexibility and essential competencies. Indonesian educators adopted UbD to address persistent challenges: rote memorization, teaching to the test, and disconnection between learning goals and classroom practices (Defianty, 2021; Smith, 2025; Teacher Standards and Accreditation, 2025).

Core principles of Merdeka Curriculum include teaching at the right level, such as matching instruction to students' actual developmental readiness rather than assuming grade-level uniformity, and prioritizing essential competencies over exhaustive content coverage (Keputusan

Menteri Pendidikan, Kebudayaan, Riset, Dan Teknologi Republik Indonesia Tentang Pedoman Penerapan Kurikulum Dalam Rangka Pemulihan Pembelajaran, 2022) This philosophy aligns conceptually with Understanding by Design's emphasis on identifying enduring understandings rather than covering all possible content, creating natural synergies with backward design approaches (Wiggins & Mctighe, 2005). Schools implementing the Merdeka Curriculum receive significant curricular autonomy, with the flexibility to adapt portions of the learning content to personalized contexts, community needs, and student interests, while maintaining alignment with core competency standards. The Merdeka Inclusive School (pseudonym) addresses Citta (a character and aspirations framework) and the 5M (a five-stage learning method) framework. However, this autonomy requires substantial teacher capacity for curriculum design and assessment development, and capabilities that backward design frameworks like UbD can support (OECD, 2024; Smith, 2025).

METHOD

This qualitative case study employed autoethnography, interviews, observations, and document analysis to investigate UbD implementation at The Merdeka Inclusive School, a progressive Indonesian inclusive school network integrating UbD with localized values (Citta, 5M).

The researcher's positionality as a learning specialist with nine months of engagement at the school (six months internship, three months observation) was acknowledged through reflexive journaling. Autoethnography provided insider insights into UbD implementation challenges, triangulated with interviews, observations, and documents to enhance credibility.

A. Data Collection

Data collection employed four complementary methods over nine months by: (1) autoethnographic reflections, (2) curriculum document analysis, (3) semi-structured interviews, and (4) participant observations. Each method is detailed below.

The UbD unit "Bumi dan Alam Semesta" (10 weeks) was analyzed across all three UbD stages: Stage 1 (Enduring Understandings, Essential Questions, Learning Outcomes), Stage 2 (performance tasks, authentic assessments), and Stage 3 (learning activities incorporating the 5M framework).

Four participants were purposively sampled: one inclusive program coordinator, two primary teachers in an inclusive classroom, and one shadow teacher, all with at least 2 years of experience implementing UbD. Interviews (60-120 minutes, Bahasa Indonesia) explored teachers' interpretations of UbD principles, the framework's flexibility in inclusive settings, and the integration of personalized value. Participant observations were conducted weekly over 9 months, documenting classroom dynamics and UbD implementation in real time through field notes. Data collection continued until theoretical saturation.

B. Data Analysis

Data analysis followed Miles and Huberman's (1994) framework, involving data reduction, data display, and conclusion drawing, and was conducted iteratively alongside data collection to achieve theoretical saturation.

Analysis employed both deductive coding aligned with UbD's three-stage framework (Stage 1: Enduring Understandings, Essential Questions, Learning Outcomes; Stage 2: performance tasks, authentic assessments; Stage 3: learning activities incorporating 5M and Citta) and inductive coding to capture emergent themes beyond the predetermined structure. The autoethnographic journal entries were analyzed first to identify preliminary themes from lived experience, which

then informed refinement of interview protocols and observation focus areas. Interview transcripts were coded line-by-line in Bahasa Indonesia to preserve cultural nuances, with themes translated to English for reporting.

Manual coding processes involved multiple readings, margin coding with reflexive memos documenting analytical decisions, color-coding to distinguish data sources and theme categories, and thematic mapping using data matrices to identify patterns across participants. Triangulation occurred through constant comparison: autoethnographic insights were validated or challenged against teacher interview data, observation field notes verified reported practices, and curriculum document excerpts provided concrete evidence of implementation approaches. All analytical decisions were documented through reflexive memos, establishing clear audit trails linking raw data to interpretive claims.

C. Credibility and Limitations

Credibility was enhanced through methodological triangulation across autoethnography, interviews, observations, and document analysis; peer debriefing with colleagues outside the research context; prolonged engagement (9 months); and reflexive journaling documenting positionality and analytical decisions. The manual coding process provided transparent audit trails, documented coding decisions, and iterative theme refinement.

This study was conducted at The Merdeka Inclusive School's Grogol (West Jakarta) and Semarang (Central Java) hubs. The Merdeka Inclusive School operates as a centralized network of inclusive schools across Indonesia, managed by a central organization that provides standardized UbD implementation guidelines while allowing individual hubs to adapt practices to their local contexts. This dual structure, which centralizes principles while allowing local adaptations, is a key dynamic investigated in this study.

Several limitations affect generalizability, though these are methodologically appropriate for case study research. First, the modest sample size (n=5, including the researcher in the autoethnography) limits statistical generalization to broader populations but is justified to achieve theoretical saturation, in which data collection continued until no new themes emerged and the phenomenon was comprehensively understood (Yin, 2018). The purposive sampling of experienced UbD practitioners ensured the selection of information-rich cases for deep exploration.

Second, this study examined two hubs within The Merdeka Inclusive School network. While this limits breadth, the dual-hub design strengthens theoretical contribution by demonstrating how centralized UbD guidelines manifest differently across distinct regional contexts (Grogol/West Jakarta and Semarang/Central Java), revealing patterns in standardization-adaptation dynamics that single-site studies would miss. Findings offer analytical generalization—transferable theoretical insights applicable to similar progressive Indonesian schools navigating dual-framework tensions (Yin, 2016).

Third, the researcher's insider role as a learning specialist introduces potential bias in interpretation, which is mitigated through reflexive journaling, triangulation with external perspectives, and peer debriefing, while providing irreplaceable access to tacit implementation knowledge that external researchers could not obtain.

These limitations represent inherent trade-offs in case study methodology, prioritizing contextual depth and theoretical development over statistical generalization (Yin, 2016). The study's contribution lies not in universal claims but in generating transferable theoretical insights about centralized-decentralized curriculum implementation dynamics that practitioners in similar contexts can evaluate for relevance to their settings.

Ethical approval was obtained from The Merdeka Inclusive School Center and its organization; informed consent was obtained from all participants with a clear explanation of voluntary

participation and withdrawal rights; and data were anonymized using pseudonyms to protect The Merdeka Inclusive School's network identity and individual participants.

RESULT AND DISCUSSION

This section presents findings from the implementation of Understanding by Design (UbD) at The Merdeka Inclusive School. The findings reveal three primary tensions emerging from The Merdeka Inclusive School's adaptation of UbD within Indonesia's dual accountability system (national Kurikulum Merdeka standards and the school's internal Citta competencies):

1. Stage 1 (Goal Setting) Navigating Dual Framework Demands: Teachers balance national standards with the school's localized values (Citta and 5M), creating ongoing negotiation in formulating learning goals, enduring understandings, essential questions, and knowledge-skills mapping across mixed-age, inclusive classrooms.
2. Stage 2 (Assessment) The Implementation Gap: Curriculum documents specify higher-order thinking objectives (Bloom's C4-C6), yet classroom enactment consistently reverts to lower cognitive levels (C1-C2). This gap reflects contradictions between authentic, personalized assessment and national regulation.
3. Stage 3 (Learning Plan) Localization Challenges: the school's 5M pedagogical framework (humanizing relationships, understanding concepts, empowering context, building sustainability, and choosing challenges) encounters implementation constraints such as time, resources, and student readiness, while flexibility for diverse learners creates coherence challenges in maintaining systematic progression.

These tensions are not isolated implementation difficulties but rather interconnected dynamics that reveal what this study conceptualizes as implementation pragmatism. The section proceeds as follows: First, Stage 1 examines goal-setting tensions, including dual-framework navigation (Section 1.1), depth-versus-coverage pressures (Section 1.2), inquiry-versus-assessment-preparation dilemmas (Section 1.3), and knowledge-skills mapping challenges in mixed-age contexts (Section 1.4). Second, Stage 2 analyzes the assessment implementation gap through the lens of Bloom's Taxonomy paradox (Section 2.1). Third, Stage 3 explores challenges in learning plan localization through the realities of implementing the 5M framework (Section 3.1). Finally, the Synthesis section in the Conclusion integrates findings across stages, situates them within Indonesian education policy, articulates the concept of implementation pragmatism, and discusses implications for practice and theory.

A. Stage 1: Goal Setting, Tensions, and Their Implications

1. Navigating Dual Framework Demands

Teachers at The Merdeka Inclusive School navigate a complex dual-layered curriculum system that integrates Indonesia's national Kurikulum Merdeka standards with the school's internal competency framework, Citta. This dual structure creates ongoing tension in goal formulation, requiring teachers to simultaneously address mandated national learning outcomes while honoring the values-based educational philosophy. Document analysis of the "Bumi dan Alam Semesta" (Earth and Universe) thematic unit reveals this tension embedded in the Established Goals component of Stage 1, where teachers must answer two distinct sets of expectations with different philosophical foundations.

The Kurikulum Merdeka emphasizes specific, measurable competencies aligned with national standards across subject disciplines, for example, requiring Grade 1-3 students to "identify characteristics of living and non-living things" (Kemendikbud, 2022) and "understand basic concepts of ecosystems." These competencies prioritize knowledge acquisition and cognitive skill

development, which are measured through formative and summative assessments rather than high-stakes standardized testing.

In contrast, SMM's Citta framework centers on holistic character development and philosophical inquiry. The same "Bumi dan Alam Semesta" (Earth and Universe) unit incorporates Citta-driven goals such as: "Bagaimana aku memahami diri sendiri, alam, dan juga yang ada di sekitarku?" (How do I understand myself, nature, and what exists around me?) and "Bagaimana bumi dan alam semesta yang ada di dalamnya saling berkaitan satu sama lain?" (How are Earth and the universe within it interconnected?). These questions emphasize self-awareness, relational understanding, and systems thinking, like competencies less amenable to conventional measurement (Corcoran, 2023; Plotinsky, 2024).

The curriculum document shows attempted integration: national science standards about ecosystems appear alongside philosophical questions about human-nature relationships. However, interviews reveal this integration as aspirational rather than seamless, with teachers noting challenges in balancing scientific rigor and philosophical depth in inclusive classrooms.

The UbD lesson plan for "Bumi dan Alam Semesta" explicitly reveals how this dual-framework tension becomes materially embedded under Established Goals. The document lists national competencies in bullet points (e.g., "Students will understand differences between living and non-living things," "Students will identify types and parts of plants") and is followed by Citta-driven essential questions. This juxtaposition on paper suggests parallel rather than integrated frameworks.

The Inclusive Program Coordinator articulated this tension directly:

> *Kadang-kadang susah juga, ya. Kita harus mikir dua kali, ini sudah sesuai dengan standar nasional belum? Tetapi tetap juga harus sesuai dengan nilai-nilai sekolah*" (Sometimes it's difficult. We must think twice, does this align with national standards yet? But it also must align with The Merdeka Inclusive School's values) (Inclusive Program Coordinator, 2025a).

This pressure intensified due to parental concerns about academic development, as the teacher reported:

> *Orang tua juga bertanya: apakah anak saya siap? Mereka appreciates nilai-nilai sekolah, tapi tetap khawatir soal academic achievement"* (Parents also ask: is my child ready? They appreciate the school's values but still worry about academic achievement) (Inclusive Class Teacher A, 2025a).

This stems from observed differences between the Merdeka Inclusive School and mainstream students. The school's mastery-based approach resulted in slower pacing and simplified content. Week 1 planning meetings exemplified this tension. Teachers debated whether to prioritize ecosystem vocabulary (Kurikulum Merdeka) or experiential connection to nature (*Citta value*), revealing philosophical splits: "meeting standards first, then adding values" versus "starting with values and finding standards within them." The compromise satisfied neither approach, with one teacher noting: "*Sepertinya kita akan adjust sambil jalan*" (It seems we'll adjust as we go along).

The tension between the dual frameworks produced differential outcomes. For neurodiverse learners, slower pacing builds confidence and foundational skills. For example, Student B (ADHD, Grade 2) progressed from "I don't know what to write" to "I saw a caterpillar. I think it will become a butterfly" by presenting a lesson in Week 2. However, they rarely encountered challenges beyond comfortable levels. For neurotypical learners, outcomes were complex. They developed strong conceptual foundations and engaged with Citta's holistic values, yet academic content was simplified to accommodate all. Student A (neurotypical, Grade 3) wrote basic observations just like "I saw ants carrying leaves," but informal conversation revealed deeper capacity: "Maybe for food? Or building homes?" This analytical thinking rarely appeared in formal work because the curriculum didn't demand it.

This finding challenges assumptions about framework integration. When frameworks embody different epistemologies, such as cognitive versus holistic (Wiggins & Mctighe, 2005), integration becomes a philosophical negotiation rather than a technical alignment (Videla et al., 2025). The pattern exemplifies *implementation pragmatism:* teachers strategically prioritize implementable elements over philosophically preferred but difficult-to-assess dimensions (Defianty, 2021; Teacher Standards and Accreditation, 2025).

Table 1. Lesson Planning 3rd Trimester "Bumi dan Alam Semesta" (Earth and Universe)

Week	Main Theme	Key Learning Focus	Core Activities	Evidence/Assessment
1	Earth and the Universe (Tuning In)	Introduction to theme; living vs non-living; concept of time; emotions; 2D shapes	Journaling, environmental observation, Y-chart, emotion discussion, shape exploration	Classification table, mind map, emotion worksheet, shape creations
	What Is Around Us?	Ecosystems; parts of plants; characteristics of living things; day–night activities; animal classification	Plant observation; videos; discussions; drawing; group classification; mini presentations	Plant worksheet; living things poster; day–night activity board; animal classification poster
2	2D Shapes & Life Cycles	Attributes of shapes; symmetry; creative shape construction; life cycles; formative performances	Origami; playdough; shape construction; performance practice (song/poem/life cycle)	Symmetry board; shape worksheets; formative performance
3	Introduction to Data	Data concepts; tally marks; pictograms; bar charts	Journaling; data concept discussion; creating tables and tallies; reading graphs	Data worksheets; tables & tallies; pictograms
4	Data & Probability	Reading and comparing data; advanced 2D shape exploration	Data review, graph-reading exercises, and extended shape creation	Data worksheets; advanced shape products
5	Beginning of Summative Project	Summative overview; product planning; rubric familiarization	Idea brainstorming; initial project planning; project journaling	Project draft; summative journal
6	Summative Product Development	Simple engineering; plant growth experiment; problem–solution identification; story-based project	Planting experiment; creating solutions; developing storyboard; building project components	Plant journal; engineering draft; project storyboard
7	Finalization & Presentation	Product finalization; presentation rehearsal; exhibition	Product refinement; presentation practice; LMS submission; term exhibition	Final product: presentation, LMS upload, exhibition display

2. Enduring Understanding: Depth vs Coverage in Practice

The Merdeka Inclusive School's curriculum structure emphasizes "deep learning," in which material progresses vertically across grade levels rather than horizontally within single grades. The Inclusive Program Coordinator explained:

The Inclusive Program Coordinator explained:

> *Struktur makro kurikulum SMM disusun vertikal antar jenjang. Jadi pendekatannya seperti deep learning yang materinya berkelanjutan"* (The macro curriculum structure is arranged vertically across levels, so the approach is like deep learning, the material is continuous) (Inclusive Program Coordinator, 2025b).

This aligns with Wiggins and McTighe's emphasis on enduring understandings, which are transferable concepts that transcend single units. For *"Bumi dan Alam Semesta"* (Earth and Universes), there are enduring understandings included: *"Alam semesta terdiri dari makhluk hidup dan benda mati yang saling berinteraksi"* (The universe consists of living and non-living things that interact with each other).

In the Week 1 document analysis, the lesson plan listed three learning objectives: understanding time concepts in daily activities, differentiating between animals and plants, and identifying issues in natural resource conservation. Within a 90-minute session (actual learning time ~60 minutes after routines), teachers allocated 30 minutes for theme explanation and small-group discussion, followed by 15 minutes for nature-walk observation and table creation. This structure allowed only a surface-level introduction to each concept, preventing the sustained inquiry depth and the enduring understanding intended (Dolmans et al., 2016; G. Wibowo et al., 2025; H. A. Wibowo & Waluya, 2025.).

Observations across Weeks 1-5 showed consistent breadth prioritization driven by: (1) front-loaded content pacing (Weeks 1-5) before project work (Weeks 6-10), (2) assessment accountability pressures ensuring coverage of required knowledge, and (3) mixed-age scaffolding demands across Grade 1-3 developmental spans. Observations in Week 3 (*method ilmiah*/scientific method lesson) showed that teachers spent significant instructional time ensuring that Grade 1 students understood basic observation vocabulary before Grade 2-3 students could proceed to hypothesis formation. This pattern reflects what Kumaya et al. (2025) describe as implementation gaps between reform aspirations and classroom realities.

The coverage-depth tension produced differentiated impacts reflecting students' incoming academic capital. Teachers and observations revealed that the school's approach was a rapid topic rotation with project-based synthesis. This benefited students entering with strong foundational skills and self-directed learning capacities, while creating challenges for those requiring more structured academic development.

During trimester observations, ecosystem lessons illustrated this pattern. Students engaged in a brief concept introduction (e.g., an animal habitats video, a simple classification activity) before quickly transitioning to hands-on projects (planting mung beans, creating dioramas).

The Teacher B reflected:

> *Anak-anak yang dari awal sudah punya strong foundation, mereka bisa thrive di sistem kita. Mereka bisa explore sendiri, connect konsep-konsep, and create sophisticated projects. Tapi untuk anak-anak yang masih butuh explicit instruction, yang butuh lebih banyak academic pressure untuk develop skills, mereka bisa hilang fokus. Mereka happy, mereka creative, tapi academically kurang ter-push"* (Children who from the start had strong foundations, they can thrive in our system. They can explore independently, connect concepts, and create sophisticated projects. But children who still need explicit instruction, who need more academic pressure to develop skills, get lost. They're happy and creative, but not academically pushed enough (Inclusive Class Teacher B, 2025b).

This reveals a troubling equity paradox: The Merdeka's inclusive, "merdeka" (liberate) philosophy, which is designed to liberate learners from rigid traditional structures, inadvertently reproduces advantage for already-privileged students while failing to provide the scaffolding that struggling learners need (Peltier, 2024). Students lacking strong prior academic preparation received concept exposure without sufficient structured instruction to build genuine understanding (Antonio & Prudente, 2023; OECD, 2024; Peltier, 2024). As one teacher reflected, *"mereka tahu konsepannya doang tapi academically kurang"* (they only know the concepts superficially but academically insufficient).

For neurodiverse learners specifically, rapid topic rotation compounded these challenges, as Shadow Teacher said:

> *Anak-anak yang butuh extra time untuk truly understand itu tantangan tersendiri karena kita cepat move on dari suatu topik ke yang lain, pengkondisiannya itu yang butuh extra. Mereka baru mulai paham ekosistem, eh sudah harus belajar data"* (Children who need extra time to truly understand things pose a challenge because we quickly move on from one topic to another, and they need extra conditioning. They are just beginning to understand ecosystems, but already they have to learn about data.) (Shadow Teacher, 2025a).

This created tension between the school's stated mastery philosophy and enacted spiral curriculum practice. The pattern challenges assumptions about mastery-based learning in mixed-age, inclusive classrooms. True mastery requires differentiated pacing, but some students need more time, others less (Dabell, 2025; Hull, 2024). This exemplifies implementation pragmatism: teachers compromise pedagogical ideals to maintain classroom coherence and coverage accountability, even when this undermines stated commitments to depth and mastery (Bergmann, 2022).

3. Essential Questions Under Pressure

The Merdeka Inclusive School employed open-ended Essential Questions per thematic unit, designed to provoke exploration and sustained inquiry. For the *"Bumi dan Alam Semesta"* ("Earth and Universe") theme, the questions included: *"Bagaimana semua yang ada di alam saling berkaitan?"* (How are all things in nature interconnected?) and *"Mengapa pelestarian sumber daya alam penting untuk kelangsungan makhluk hidup?"* (Why is natural resource conservation important for living things' survival?). These questions aligned with McTighe and Wiggins's (2005) emphasis on provocative inquiries that "point toward important, transferable ideas within (and sometimes across) disciplines" (p. 110) and require students to "think, dig, search, and research" (p. 106).

Document analysis revealed sophisticated question design. Unlike traditional textbook questions that seek single correct answers, The Merdeka Inclusive School's Essential Questions deliberately invite multiple perspectives and require synthesis across content domains such as ecosystems, weather, conservation, and human impact. Week 1's question *"Ada apa saja di alam sekitar kita?"* (What exists in nature around us?) progressed by Week 5 to *"Mengapa pelestarian sumber daya alam penting untuk kelangsungan makhluk hidup?"* (Why is natural resource conservation important for living things' survival?) demonstrating intentional scaffolding toward higher-order thinking.

Classroom observations revealed how this tension manifested in practice. During Week 4's lessons exploring connections between weather and ecosystems within the Essential Question *"Bagaimana bumi dan alam semesta saling berkaitan?"* (How are Earth and the universe interconnected?), students initially engaged in rich discussion about how weather affects plant growth, animal behavior, and human activities. However, when transitioning to worksheet activities, teachers reformulated the open inquiry into structured formats: *"Sebutkan tiga cara cuaca mempengaruhi tumbuhan"* (Name three ways weather affects plants). This shift from explorato-

ry dialogue to closed-ended responses occurred repeatedly across observed lessons, suggesting teachers felt compelled to "translate" Essential Questions into assessment-compatible formats that provided tangible evidence of learning (Chen, 2024; Tóth & Csapó, 2022).

The tension intensified in mixed-age classrooms where developmental readiness varied dramatically. Grade 1 students, who are still developing basic literacy skills, struggled to engage meaningfully with abstract Essential Questions such as *"Mengapa semua yang ada di alam sekitar saling berkaitan?"* (Why are things in nature all connected?) requires conceptual thinking beyond their immediate experience. Meanwhile, Grade 3 students capable of sophisticated responses received simplified prompts to maintain classroom coherence. Teachers reported additional challenges specific to inclusive contexts—Essential Questions assume a metacognitive capacity to recognize when thinking remains superficial rather than achieving deeper understanding, a skill many neurodiverse learners had not yet developed. One teacher noted:

> *Mereka excited untuk jawab pertanyaan besar, tapi kadang jawabannya masih di permukaan. Susah untuk tau apakah mereka truly understand atau hanya enthusiastic.* (They're excited to answer big questions, but sometimes their answers remain surface-level. It's difficult to know whether they truly understand or are just enthusiastic) (Inclusive Class Teacher A, 2025c).

Those findings reveal fundamental tensions in adapting UbD to assessment-constrained contexts. Wiggins & McTighe (2005) designed Essential Questions to resist reductionist teaching, arguing they should "cause genuine and relevant inquiry into big ideas and core content" (p. 110). However, this presumes assessment systems reward such inquiry (Fawns et al., 2025). When external accountability structures prioritize discrete knowledge demonstration over sustained inquiry, Essential Questions risk becoming decorative rather than functional curriculum elements (Fawns et al., 2025; Tóth & Csapó, 2022). For Indonesian inclusive education, this reveals that Essential Questions cannot function under assessment pressure without systemic reform which implementing UbD's inquiry-based pedagogy demands not merely teacher training but fundamental restructuring of accountability systems that currently reward procedural recall over sustained inquiry.

4. Knowledge-Skills Mapping Challenges

The fourth component of Stage 1 involves identifying specific knowledge and skills students must acquire. While Section 1.2 identified mixed-age classroom complexity as one factor driving coverage prioritization, the challenge extends beyond time allocation to fundamental difficulties in mapping competencies across developmental spans. At The Merdeka Inclusive School, this posed unique challenges as teachers mapped competencies across disciplinary boundaries and in Grade 1-3 mixed-age environments. The knowledge component for *"Bumi dan Alam Semesta"* (Earth and Universe) encompassed factual information (e.g., ecosystem vocabulary, plant parts), conceptual understanding (e.g., interdependence, life cycles), and procedural knowledge (e.g., scientific observation methods). Skills included both academic competencies (classification, data recording) and socio-emotional capabilities (collaboration, environmental stewardship), reflecting Anderson and Krathwohl's (2009) revised taxonomy distinguishing knowledge types and cognitive processes.

Document analysis revealed an ambitious scope. The 10-week unit specified knowledge spanning multiple disciplines: science (ecosystems, weather, conservation), mathematics (geometric shapes, data analysis, measurement), language arts (narrative writing, poetry), and arts (dioramas, creative projects). Simultaneously, teachers needed to address Citta competencies, for example, philosophical inquiry, relational awareness, contextual empowerment, creating what one coordinator termed "layered complexity":

> K*ita harus pastikan anak-anak dapat pengetahuan dasar yang dibutuhkan, tapi juga harus*

> *integrate dengan pendekatan holistik The Merdeka Inclusive School.* (We must ensure children get the basic knowledge needed, but also must integrate with The Merdeka Inclusive School's holistic approach) (Inclusive Program Coordinator, 2025c).

This challenge intensified in mixed-age settings where knowledge and skill baselines varied dramatically. Week 2's plant observation exemplified this: Grade 1 needed vocabulary (leaf, stem, root), Grade 2 added plant functions and systematic recording, Grade 3 required ecosystem connections and comparative analysis—three distinct levels within one lesson. Teachers attempted differentiation through tiered worksheets, with Grade 1 drawing and labeling parts, Grade 2 completing structured observation tables, and Grade 3 writing analytical paragraphs. However, whole-class instruction time focused on Grade 1-2 level content to ensure foundational understanding, leaving Grade 3 students without sufficient cognitive challenge in shared learning segments.

The skills component proved particularly complex. Teachers simultaneously developed academic skills, such as critical thinking, scientific inquiry, and mathematical reasoning, alongside character development aligned with Citta values. One teacher articulated this tension:

> *Setiap anak di kelas campuran ini levelnya berbeda-beda.* (Every child in this mixed class has different levels.) (Inclusive Class Teacher B, 2025c).

Observations confirmed this dilemma. During Week 3's scientific method lesson, teachers needed Grade 1 students to master basic observation (describing what they see) while Grade 3 students should practice hypothesis formation and evidence evaluation. The 30-minute shared instruction segment prioritized Grade 1-2 needs—defining observation, modeling descriptive language, and guided practice with teacher support. Grade 3 students completed these activities quickly, then received independent extension tasks (more complex observation scenarios) while teachers continued scaffolding younger students. This pattern created what observers termed "wait time inequity"—advanced learners spent significant instructional time waiting or working independently rather than receiving direct instruction at their level of challenge.

The knowledge-skills mapping challenge produced three primary consequences. First, curriculum documents specified comprehensive competencies across all grades, but the enacted curriculum consistently prioritized foundational levels to maintain coherence in mixed-age classrooms. Analysis of student work samples from Weeks 1-5 showed that Grade 3 students produced work qualitatively similar to that of Grade 2 peers despite differentiated task instructions. When asked why her ecosystem diagram resembled simpler versions, one Grade 3 student explained: *"Itu yang diajarkan"* (That's what was taught), revealing that whole-class instruction anchored expectations regardless of individual capacity.

Second, skills requiring sustained practice and progressive complexity—particularly higher-order thinking and academic discourse—received inconsistent attention. Teachers reported that explaining skills at multiple levels simultaneously proved cognitively demanding, leading to simplified instruction benefiting struggling learners but insufficiently challenging advanced students (Fawns et al., 2025; Teacher Standards and Accreditation, 2025). The shadow teacher observed:

> *Untuk keterampilan yang kompleks seperti analisis atau evaluasi, guru cenderung fokus ke level yang paling banyak anak bisa ikuti. Yang sudah advanced, mereka kurang dapat scaffolding untuk naik ke level berikutnya.* (For complex skills like analysis or evaluation, teachers tend to focus on the level most children can follow. Those already advanced, they don't get scaffolding to move to the next level) (Shadow Teacher, 2025c).

Third, integrating knowledge-skills with the Citta framework required constant negotiation between explicit competency requirements and implicit value formation. Teachers spent planning meetings debating whether specific lessons should prioritize knowledge acquisition (identifying animal characteristics) or values development (appreciating biodiversity, developing

environmental care). This finding challenges assumptions about differentiation in mixed-age inclusive classrooms. While educational literature advocates heterogeneous grouping for social benefits and peer learning (Vygotsky, 1978), The Merdeka Inclusive School's experience reveals that significant age/ability spans create knowledge-skill mapping complexity that exceeds the capacity of differentiation strategies. When developmental distances span three grade levels, teachers face impossible choices: teach to the middle (leaving both ends underserved), prioritize foundations (limiting advanced learners), or attempt individualization (creating management and coherence challenges). The pattern suggests that the UbD framework is insufficiently addressed; backward design assumes relatively stable learner profiles when planning "acceptable evidence," an assumption violated by wide-ranging mixed-age contexts where acceptable evidence differs fundamentally across the span (Wiggins & McTighe).

For Indonesian inclusive education, this tension exposes limitations in progressive pedagogical assumptions. The Merdeka Inclusive School's commitment to inclusion and mixed-age learning reflects admirable philosophical values, yet implementation reveals structural constraints on achieving both depth and differentiation simultaneously. Successful knowledge-skills mapping in such contexts may require acknowledging trade-offs—accepting that prioritizing inclusion and social learning necessarily involves compromises in academic challenge for some learners—rather than maintaining that all pedagogical goals can be achieved concurrently through sufficient teacher skill (UC Berkeley, 2025). This honest reckoning with inclusive education's complexities better serves students than idealized narratives obscuring real implementation dilemmas (UC Berkeley, 2025).

B. Stage 2: Assessment Implementation Gap and Its Causes

1. The Bloom's Taxonomy Paradox

Document analysis of the "Bumi dan Alam Semesta" (Earth and Universe) unit revealed assessment tasks explicitly aligned with Bloom's Taxonomy spanning C1 (remembering) to C6 (creating), suggesting systematic progression in cognitive demand. Curriculum documents specified higher-order objectives: "Students will evaluate the effectiveness of environmental conservation strategies" (C6-Evaluating) and "Students will create innovative solutions for local environmental problems" (C6-Creating). Stage 2 assessment designs appeared sophisticated, incorporating authentic performance tasks, student self-reflection, and portfolio documentation aligned with UbD's emphasis on varied evidence of understanding (Anderson, 2009; Wiggins & Mctighe, 2005).

However, systematic classroom observations across Weeks 1-8 revealed persistent gaps between documented cognitive levels and the instruction enacted. While curriculum documents specified C4-C6 objectives, actual classroom activities consistently operated at C1-C2 levels. Systematic classroom observations revealed consistent cognitive downgrading across all content domains (Table 2). While curriculum documents specified C4-C6 objectives emphasizing analysis, evaluation, and creation, enacted instruction consistently operated at C1-C2 levels, requiring recall and basic comprehension. This pattern appeared in science (ecosystem analysis -> sorting cards), mathematics (creating patterns -> identifying shapes), and language arts (composing narratives -> writing three sentences) (Anderson, 2009).

The systematic nature of this downgrading indicated structural rather than incidental causes. Teacher interviews revealed that cognitive downgrading stemmed from pragmatic responses to structural constraints rather than pedagogical misconceptions. Faced with the structural constraints identified in Sections 1.2 and 1.4—compressed timelines, mixed-age heterogeneity, and coverage pressures—teachers made pragmatic decisions about cognitively achievable activities within existing conditions.

Di dokumen tertulis C4-C5, tapi saat mengajar kita sesuaikan dengan kondisi kelas. (In do-

cuments, it's written as C4-C5, but when teaching, we adjust to classroom conditions) (Inclusive Class Teacher A, 2025d)

However, this theoretical awareness did not lead to adequate cognitive progress. While teachers differentiated tasks by grade level, the pace of cognitive development remained remarkably slow. Observations revealed Grade 1 students at C1 (recall), Grade 2 at C2 (comprehension), and Grade 3 still at C2-C3; three years progressing only 1-2 cognitive levels. The ecosystem unit exemplified this: Grade 1 identified living/non-living things; Grade 2 listed examples; Grade 3 wrote characteristics, but none analyzed ecological relationships or evaluated conservation strategies, as required by the documented C4-C6 objectives. This conservative progression made higher-order outcomes structurally impossible; students moving from C1 to C3 over three years left no time to reach documented C5-C6 goals.

This created "performative rigor"—documents maintained C4-C6 goals while instruction remained trapped in the C1-C3 range. Assessment rubrics revealed this mismatch: the ecosystem diorama assessed identification and description (achievable given slow progression) rather than analysis and evaluation (unachievable given conservative pace). The resulting system normalized a gap between the pace of cognitive development and curriculum expectations. This gap between documented rigor and enacted instruction produced significant consequences for learning outcomes (Pritchett, 2025).

The Bloom's Taxonomy paradox produced significant consequences for learning outcomes. Student work samples from summative projects revealed a pattern: sophisticated creative products (detailed dioramas, colorful posters, elaborate models) demonstrating C1-C2 cognitive engagement—recall and comprehension—rather than the analysis, evaluation, and creation documented in the specified objectives. A Grade 3 student's ecosystem diorama illustrated this disconnect. The project showed meticulous craft—carefully arranged plants, animals, water features, and decorative elements. When asked to explain ecosystem interactions (the C4 objective), the student responded: *"Ada kelinci, ada wortel, kelinci makan wortel"* [There's a rabbit, there's a carrot, the rabbit eats the carrot]. This demonstrated awareness that animals eat plants, but not analysis of interdependencies, energy flow, or systemic relationships, the enduring understanding emphasized.

For neurodiverse learners, consequences proved particularly detrimental. Students requiring explicit instruction in higher-order thinking strategies, metacognitive scaffolding, structured analytical frameworks, and modeled evaluation criteria rarely encountered such teaching (Azuka et al., 2024). The assumption that students would naturally progress from remembering to analyzing through exposure and project work disadvantaged learners needing systematic cognitive skill development (Azuka et al., 2024).

The pattern reveals fundamental tensions in UbD implementation within resource-constrained inclusive contexts. Wiggins and McTighe (2005) designed Stage 2 to ensure assessments elicit genuine understanding rather than superficial performance. However, this presumes instructional capacity to develop higher-order thinking—adequate time, manageable class sizes, homogeneous-enough groups for efficient scaffolding, and teacher expertise in cognitive skill instruction (Bergmann, 2022; Pak et al., 2020). When these conditions don't exist, the gap between documented objectives and enacted assessment becomes inevitable (Bergmann, 2022). Teachers face impossible choices: maintain aspirational C4-C6 objectives knowing implementation will fall short, or document achievable C1-C2 objectives, acknowledging this contradicts UbD's emphasis on understanding over recall (Adams, 2015; Bergmann, 2022).

This finding extends research on implementation gaps in Bloom's Taxonomy. While existing literature documents teacher difficulty in designing higher-order assessments (Anderson, 2009), The Merdeka Inclusive School case reveals that even when assessments are well-designed on paper, structural constraints—time, heterogeneity, resource limitations—can render implementation impossible. For Indonesian inclusive education, this suggests that adopting sophis-

ticated curriculum frameworks without addressing systemic constraints produces what we term implementation fiction—curriculum documents specifying practices unsupported by contextual realities, creating professional dissonance as teachers navigate gaps between documented intentions and feasible enactment (Talley & Huddleston, 2024).

Table 2 Bloom's Taxonomy Implementation Gap Analysis

Week	Documented Learning Objective	Documented Bloom's Level	Enacted Classroom Activity	Actual Bloom's Level	Implementation Gap
2	"Students will analyze relationships between living and non-living components in ecosystems"	C4 (Analyzing)	Watched rainforest video; sorted picture cards into "living things" and "non-living things" categories	C1-C2 (Remembering/ Understanding)	-2 to -3 levels
5	"Students will evaluate different conservation methods and justify which approaches would be most effective in local contexts"	C5-C6 (Evaluating/ Creating)	Listed 3 ways to protect the environment: created posters showing recycling, planting trees, and anti-littering messages	C1-C2 (Remembering/ Understanding)	-3 to -4 levels
3	"Students will apply the scientific method to formulate and test hypotheses"	C3-C4 (Applying/ Analyzing)	Defined "observation"; practiced describing what they see with teacher guidance	C1-C2 (Remembering/ Understanding)	-1 to -2 levels
Math (unspecified)	"Students will create original geometric patterns demonstrating understanding of shape properties"	C6 (Creating)	Identified shapes in the classroom and drew them	C1-C2 (Remembering/ Understanding)	-4 to -5 levels
Language Arts (unspecified)	"Students will compose narrative texts integrating scientific concepts with creative storytelling"	C5-C6 (Evaluating/ Creating)	Wrote 3 sentences about what they observed during nature walk	C2-C3 (Understanding/ Applying)	-2 to -3 levels

C. Stage 3: Learning Plan Localization and Implementation Realities

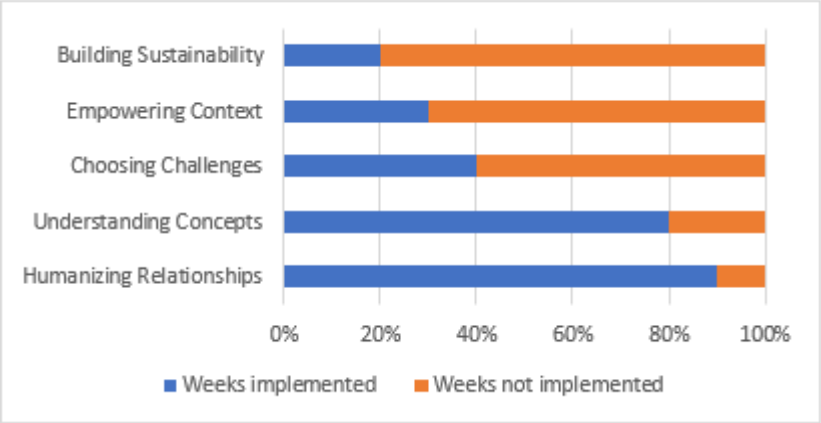
1. The 5M Framework: Imbalance and Structural Constraints

The Merdeka Inclusive School's Stage 3 learning plans center on the 5M methodological framework: Humanizing Relationships, Understanding Concepts, Empowering Context, Building Sustainability, and Choosing Challenges. Curriculum documents clarify that the 5M framework functions as a trimester-level structure rather than a per-lesson requirement—individual lessons

may emphasize 2-3 principles, provided that all five receive adequate attention throughout the 10-week unit. Document analysis of the *"Bumi dan Alam Semesta"* (Earth and Universe) unit revealed that lesson plans were systematically coded which "The Ms" each activity addressed: Week 1's nature walk specified Humanizing Relationship by peer collaboration during observation, Understanding Concepts by identifying living vs. non-living things, and Empowering Context by connecting learning to the school environment. Week 3's scientific method lesson emphasized Understanding Concepts through hypothesis formation and Choosing Challenges in selecting observation subjects. This documentation suggested teachers understood the method's flexibility, selecting appropriate principles for each lesson's learning goals while maintaining the trimester's overall 5M integrity (OECD, 2024).

However, systematic analysis of lesson plans and classroom observations across all 10 weeks revealed a significant imbalance in the 5M distribution. Frequency analysis showed by the Picture 1 below:

Picture 1. 5M Framework Implementation Frequency



Furthermore, when underrepresented, "The Ms" appeared; their implementation often remained superficial. Week 6's lesson documented Empowering Context by showing a video on local conservation issues, but this passive media consumption fundamentally differed from the principle's intended meaning—actively empowering students to investigate and act in their real contexts. Week 7 coded Choosing Challenges by offering a choice among creating dioramas, writing stories, or designing action plans, yet observations revealed that this predetermined menu limited genuine differentiation of challenge compared to the principal's vision of students identifying their own learning edges. The pattern indicated not merely lesson-level flexibility but a systematic underimplementation of specific "The Ms" across the entire trimester. It shows a structural imbalance contradicting the framework's requirement that all five principles receive adequate attention over the unit's duration.

This imbalanced distribution reflected teachers' strategic responses to differential implementation barriers across the five principles (Dacus-Hare, 2023; Pak et al., 2020). Humanizing relationships and understanding concepts proved most feasible because they aligned with conventional classroom practices and required minimal additional resources.

Teachers reported comfort with these practices:

> *Memanusiakan hubungan itu natural buat kita. Setiap hari kita buka dengan circle time, anak-anak kerja kelompok, kita facilitate interaksi. Itu bukan sesuatu yang extra harus direncanakan"* (Humanizing relationships is natural for us. Every day we start with circle time, children work in groups, we facilitate interaction. That's not something extra we must plan) (Inclusive Class Teacher A, 2025e)

Similarly, Understanding Concepts represented core instructional work, for example, introducing content, explaining ideas, guiding practice, a central to teaching regardless of pedago-

gical framework. Teachers seamlessly integrated this into lesson structures without additional preparation or resources. In contrast, Empowering Context, Building Sustainability, and Choosing Challenges encountered substantial structural barriers, making repeated implementation prohibitively difficult. Empowering Context required resources and coordination extending beyond classroom boundaries (Dacus-Hare, 2023; Pak et al., 2020).

The Inclusive Program Coordinator explained:

> *Untuk truly empowering context, kita butuh bawa anak keluar, connect dengan real-world settings. Itu butuh planning ekstensif dan kita nggak bisa lakukan itu setiap minggu"* (To truly empower context, we need to take children outside, connect with real-world settings. That requires extensive planning, and We can't do that every week) (Inclusive Program Coordinator, 2025d).

Week 6's documented plan to visit a conservation area exemplified these barriers. The backup plan, showing a conservation documentary, satisfied curriculum documentation requirements but failed to provide the embodied contextual engagement the principal intended. After this experience, subsequent weeks reverted to classroom-bound activities requiring less logistical complexity.

Building Sustainability faced different but equally constraining barriers rooted in curriculum structure rather than resources. This principle emphasizes revisiting concepts over time with increasing complexity, the spiral curriculum model supporting memory consolidation and deepening understanding (Ireland & Mouthaan, 2020). However, the Merdeka Inclusive School's trimester system created temporal fragmentation, working against sustainability. Each 10-week trimester addressed distinct themes: Trimester 1 focused on *"Aku dan Sekitarku"* (Me and My Surroundings), Trimester 2 on *"Aku Penjelajah Dunia"* (I Am a World Explorer), Trimester 3 on *"Bumi dan Alam Semesta"* (Earth and Universe), culminating in Trimester 4 with *"Aku Pemimpin Masa Depan"* (I Am a Future Leader). While themes were built progressively from self-awareness toward global citizenship, substantial content gaps existed between trimesters; environmental concepts introduced in Trimester 3 (ecosystems, conservation, interdependence) had little foundation in the previous trimesters' focus on personal identity and cultural exploration. The curriculum's conceptual progression emphasized expanding social spheres (self -> world -> planet -> leadership) rather than deepening disciplinary understanding within domains. This structure made the implementation of Building Sustainability paradoxical: teachers could build sustainability within the 10-week unit (as the plant growth observations in Weeks 2-7 attempted) but could not construct the longer developmental arcs across trimesters, as the principal had philosophically intended. Scientific concepts introduced in Trimester 3 appeared isolated rather than building on sustained inquiry begun earlier (Pak et al., 2020).

Choosing Challenges encountered implementation barriers rooted in students' developmental readiness and teachers' pedagogical capacity. This principle reflects constructivist commitments to learner agency, which allows students to identify their learning edges and select appropriately challenging tasks matching their zones of proximal development (Vygotsky, 1978). However, this requires metacognitive capacities that many students, particularly younger children and neurodiverse learners, have not yet developed (Teacher Standards and Accreditation, 2025). Week 7's differentiated ecosystem project offered a choice among dioramas, stories, or action plans to provide varied levels of challenge. Observations revealed that without explicit scaffolding in self-assessment, understanding one's current capabilities, recognizing growth areas, and gauging task difficulty are difficult. The Shadow Teacher noted:

> *Anak-anak pilih based on apa yang mereka suka atau apa yang temen mereka pilih, bukan based on apa yang akan stretch learning mereka. Student B always pilih hands-on crafts karena dia enjoy, padahal dia perlu di-push untuk develop writing skills yang masih weak* (Children choose based on what they like or what their friends choose, not based on what will stretch their learning. Student B always picks hands-on crafts because he enjoys them,

even though he needs to be pushed to develop writing skills that are still weak (Shadow Teacher, 2025).

The uneven implementation of the 5M framework components reveals a critical tension between *design flexibility* and *implementation equity*, challenging assumptions in the contemporary curriculum design literature. While the framework's trimester-level structure intentionally provides temporal flexibility by acknowledging classroom realities and respecting teacher agency, the research site's experience demonstrates that flexibility operates as a necessary *but insufficient* condition for balanced implementation (Kossek et al., 2015)

For neurodiverse learners specifically, under implementing the Empowering Context, limited access to embodied learning is crucial for abstract concept comprehension (Moore, 2007). Student B's (ADHD, Grade 2) case illustrates this dependency vividly: intense engagement and spontaneous conceptual connection during garden-based observation contrasted sharply with fragmented understanding when instruction shifted to classroom-bound, abstract formats. This pattern suggests that for many neurodiverse learners, embodied context is not merely one beneficial learning mode among several equivalent options; it may constitute the *primary pathway* through which abstract concepts become meaningful (Schilhab, 2021). When curriculum implementation systematically underemphasizes embodied learning due to logistical barriers, it effectively creates exclusionary conditions for learners who require concrete, sensory-rich experiences to access abstract content (Barnes, 2024; Schilhab, 2021).

CONCLUSION

This study demonstrates that implementing Understanding by Design (UbD) in inclusive school contexts is not a straightforward application of backward-design principles but rather a negotiated process shaped by structural constraints, accountability pressures, and learner heterogeneity. Through a qualitative case study across two regional hubs, the findings reveal persistent tensions in goal setting, assessment enactment, and learning plan implementation as centralized UbD guidelines intersect with national curriculum mandates and localized pedagogical values. These dynamics are conceptualized as implementation pragmatism, whereby teachers strategically adapt aspirational curriculum designs to what is institutionally and pedagogically feasible in mixed-age inclusive classrooms. While such pragmatism enables curriculum continuity under real-world conditions, it simultaneously constrains higher-order cognitive development, assessment authenticity, and equitable learning challenge, particularly for neurodiverse learners who require explicit cognitive scaffolding. Theoretically, this study extends UbD scholarship by situating backward design within inclusive education contexts in the Global South and foregrounding the structural gap between curriculum design intentions and enacted classroom realities. In practice, the findings call for stronger alignment among inclusive curriculum frameworks, assessment accountability systems, and sustained teacher capacity development at the school, organizational & pedagogical levels. As a qualitative case study, this research prioritizes contextual depth over statistical generalization and offers transferable theoretical insights rather than universal claims. Future research should examine UbD implementation across broader inclusive school networks and investigate the longitudinal effects of implementation pragmatism on learner trajectories under varying assessment and inclusion regimes.

REFERENCES

Adams, N. E. (2015). Bloom's taxonomy of cognitive learning objectives. J*ournal of the Medical Library Association : JMLA*, 103(3), 152–153. https://doi.org/10.3163/1536-5050.103.3.010

Ainscow, M. (2020). Promoting inclusion and equity in education: Lessons from international experiences. *Nordic Journal of Studies in Educational Policy*, 6(1), 7–16. https://doi.org/10.1080/20020317.2020.1729587

Anderson, L. W. (Ed.). (2009). A *taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives* (Complete ed., [Nachdr.]). Longman.

Antonio, R. P., & Prudente, M. S. (2023). Effects of inquiry-based approaches on students' higher-order thinking skills in science: A Meta-Analysis. *International Journal of Education in Mathematics, Science and Technology*, 12(1), 251–281. https://doi.org/10.46328/ijemst.3216

Aslam, A., Ahamd, S., Siller, H. S., & Nasreen, A. (2024). Impact of the understanding by design model on the science academic achievement of fifth grade students in Pakistan. *Asia-Pacific Science Education,* 10(1), 113–153. https://doi.org/10.1163/23641177-bja10078

Asmayawati, Yufiarti, & Yetti, E. (2024). Pedagogical innovation and curricular adaptation in enhancing digital literacy: A local wisdom approach for sustainable development in Indonesia context. J*ournal of Open Innovation: Technology, Market, and Complexity*, 10(1), 100233. https://doi.org/10.1016/j.joitmc.2024.100233

Azuka, C. V., Wei, C. R., Ikechukwu, U. L., & Nwachukwu, L. (2024). *Inclusive instructional design for neurodiverse learners*. 1(2).

Barnes, E. (2024). *Creating outdoor learning opportunities for neurodiverse students: an outdoor learning curriculum for neurodiverse students.*

Bergmann, J. (2022). The biggest problem with mastery-based learning and how to solve it. *ASCD*. https://www.ascd.org/blogs/the-biggest-problem-with-mastery-based-learning-and-how-to-solve-it

ÇeliKman Hanratty, G., & EveyiK Aydin, E. (2023). Designing units with the UBD framework to teach english as a foreign language: benefits and challenges. *Sakarya University Journal of Education*, 13(3), 435–455. https://doi.org/10.19126/suje.1277604

Chen, S. (2024). During the reform: how teachers' lived experiences influence their beliefs and practices of project- based learning. *International Journal of Educational Reform*, 10567879241265096. https://doi.org/10.1177/10567879241265096

Cole, J. Y., Graham, J., Norvell, S., & Schaber, P. (2024). Universal design for learning principles impact on students with neurodiverse learning styles. *Journal of Occupational Therapy Education*, 8(2). https://doi.org/10.26681/jote.2024.080204

Corcoran, E. (2023). *Backward design*. In University of Illinois Chicago.

Dabell, J. (2025). *Mixed ability teaching for mastery: making it work in your class*. https://thirdspacelearning.com/blog/every-school-leader-know-mastery-mixed-ability-maths-class/

Dacus-Hare, D. (2023). *The barriers teachers face when implementing the universal design for learning framework* [Master of Science, Dominican University of California]. https://doi.org/10.33015/dominican.edu/2023.EDU.15

Dari, W., Hidayat, S., & Wulandari, E. (2024). The understanding by design strategy in 21st-century education. *Biosfer: Jurnal Tadris Biologi*, 14(2), 169. https://doi.org/10.24042/biosfer.v14i2.15818

Defianty, M. (2021). Asesmen pengganti UN akan tetap jadi momok jika guru dan sekolah masih pegang budaya tes. *The Conversation*. https://theconversation.com/asesmen-pengganti-un-akan-tetap-jadi-momok-jika-guru-dan-sekolah-masih-pegang-budaya-tes-155135

Dolmans, D. H. J. M., Loyens, S. M. M., Marcq, H., & Gijbels, D. (2016). Deep and surface learning in problem-based learning: A review of the literature. *Advances in Health Sciences Education*, 21(5), 1087–1112. https://doi.org/10.1007/s10459-015-9645-6

Fawns, T., Bearman, M., Dawson, P., Nieminen, J. H., Ashford-Rowe, K., Willey, K., Jensen, L. X., Damşa, C., & Press, N. (2025). Authentic assessment: From panacea to criticality. *Assessment & Evaluation in Higher Education*, 50(3), 396–408. https://doi.org/10.1080/02602938.2024.2404634

Flood, M., Frizzarin, A., & Carlos, M. (2025). *Universal design for learning. all means all! opentextbook for diversity in education*. https://book.all-means-all.education/ama-2025-en/chapter/universal-design-for-learning/

Florian, L., & Black-Hawkins, K. (2011). Exploring inclusive pedagogy. *British Educational Research Journal*, 37(5), 813–828. https://doi.org/10.1080/01411926.2010.501096

Hamilton, L. G., & Petty, S. (2023). Compassionate pedagogy for neurodiversity in higher education: A conceptual analysis. *Frontiers in Psychology*, 14, 1093290. https://doi.org/10.3389/fpsyg.2023.1093290

Hull, C. (2024). Mastery Learning: The educator's guide. *Otus*. https://otus.com/resources/guides/mastery-learning

Ireland, J., & Mouthaan, M. (n.d.). *Perspectives on curriculum design: Comparing the spiral and the network models*.

Joshi, S. P. (2021). *Evaluation of the Implementation Of Understanding By Design Processes In Select Minnesota Public Schools*.

Karim, H. (2024). Meningkatkan profesionalisme guru dan prestasi siswa dalam pembelajaran ekonomi melalui model understanding by design (UBD) dan differentiated instruction (DI) di SMA Yapidh Kota Bekasi. *FOCUS*, 5(1), 37–42. https://doi.org/10.37010/fcs.v5i1.1129

Keputusan Menteri Pendidikan, Kebudayaan, Riset, Dan Teknologi Republik Indonesia Tentang Pedoman Penerapan Kurikulum Dalam Rangka Pemulihan Pembelajaran, Pub. L. No. NOMOR 56/M/2022.

Kossek, E. E., Thompson, R. J., & Lautsch, B. A. (2015). Balanced workplace flexibility: avoiding the traps. *California Management Review*, 57(4), 5–25. https://doi.org/10.1525/cmr.2015.57.4.5

Kumayas, T. A., Luntungan, G. S., Arsai, A. N., & Tatipang, D. P. (n.d.). *Curriculum reform meets classroom realities: selected indonesian teachers' perspectives on the merdeka curriculum*.

Le Cunff, A.-L., Martis, B.-L., Glover, C., Ahmed, E., Ford, R., Giampietro, V., & Dommett, E. J. (2025). Cognitive load and neurodiversity in online education: A preliminary framework for educational research and policy. *Frontiers in Education*, 9, 1437673. https://doi.org/10.3389/feduc.2024.1437673

Mctighe, J., & Wiggins, G. (2012). Understanding by design® framework.

Moore, S. L. (2007). David H. Rose, Anne Meyer, Teaching every student in the digital age: universal design for learning: association for supervision and curriculum development. *Educational Technology Research and Development*, 55(5), 521–525. https://doi.org/10.1007/s11423-007-9056-3

Mpolomoka, D. L., Muvombo, M., Noyoo, N., Mwale, J., & Chirwa, E. (2025). Teaching neurodiverse early childhood learners: a bio-ethnographic study. *European Journal of Public Health Studies*, 8(3). https://doi.org/10.46827/ejphs.v8i3.223

Nur Safitri, E., & Putri Fatmawati, D. (2024). Analysis of inclusive education curriculum in physical education, sports, and health subjects at brajan elementary school, *JURNAL PENDIDIKAN*, 25(1), 74–88. https://doi.org/10.52850/jpn.v25i1.15155

Nweje, U. (2025). STEM for all abilities: unlocking the potential of special-needs and neurodiverse students in america's innovation pipeline. *International Journal of Science and Research Archive*, 17(1), 604–617. https://doi.org/10.30574/ijsra.2025.17.1.2837

OECD. (2024). *Curriculum flexibility and autonomy: promoting a thriving learning environment*. OECD Publishing. https://doi.org/10.1787/eccbbac2-en

Pak, K., Polikoff, M. S., Desimone, L. M., & Saldívar García, E. (2020). The adaptive challenges of curriculum implementation: insights for educational leaders driving standards-based reform. *AERA Open*, 6(2), 2332858420932828. https://doi.org/10.1177/2332858420932828

Peltier, C. (2024). Explicit instruction – what students with disabilities need most. *Oklahoma Education Journal*, 2(3).

Plotinsky, M. (2024). *Making space for your autonomy within a set curriculum* [Online post]. Edutopia. https://www.edutopia.org/article/teacher-autonomy-set-curriculum/

Pramesti, N., & Dewi, L. (2023a). The Implementation of understanding by design approach in mathematics learning on elementary school. *(JIML) JOURNAL OF INNOVATIVE MATHEMATICS LEARNING*, 6(2), 124–131. https://doi.org/10.22460/jiml.v6i2.16304

Pramesti, N., & Dewi, L. (2023b). The implementation of understanding by design approach in mathematics learning on elementary school. *(JIML) JOURNAL OF INNOVATIVE MATHEMATICS LEARNING*, 6(2), 124–131. https://doi.org/10.22460/jiml.v6i2.16304

Pritchett, L. (2025). Addressing the learning crisis: An emergent consensus. In T. Besley, I. Bucelli, & A. Velasco (Eds.), *The London Consensus: Economic Principles for the 21st Century* (pp. 391–432). LSE Press. https://doi.org/10.31389/lsepress.tlc.l

Ralabate, P. K. (2016). *The step-by-step guide for teaching all learners*.

Rizky, A. A., & Amal, F. (2025). Revisiting Indonesia's English curriculum: policy gaps and classroom realities. *Jurnal Tahuri*, 22(2), 102–119. https://doi.org/10.30598/tahurivol22issue2page102-119

Roth, D. (2007). Understanding by design: a framework for effecting curricular development and assessment. *CBE—Life Sciences Education*, 6(2), 95–97. https://doi.org/10.1187/cbe.07-03-0012

Schilhab, T. (2021). Nature experiences in science education in school: review featuring learning gains,

investments, and costs in view of embodied cognition. *Frontiers in Education*, 6, 739408. https://doi.org/10.3389/feduc.2021.739408

Smith, P. (2025). *The open guide to teaching and learning in higher education* [Online post].

Talley, S., & Huddleston, A. P. (2024). Principled resistance: how teachers can survive and thrive in a time of standardization. *In The Heart of Literacy Connects Us All!* (Vol. 11). Texas Association for Literacy Education.

Teacher Standards and Accreditation. (2025). *Backward design model* [Online post]. NSW Government.

Tomlinson, C. A., & McTighe., J. (2006). *Integrating differentiated instruction and understanding by design: Connecting content and kids.* Association for Supervision and Curriculum Development.

Tóth, E., & Csapó, B. (2022). Teachers' beliefs about assessment and accountability. *Educational Assessment, Evaluation and Accountability*, 34(4), 459–481. https://doi.org/10.1007/s11092-022-09396-w

Tumlos-Castillo, L. M. (2015). Initial evaluation of the understanding by design (UBD) framework in writing learning modules. 22.

UC Berkeley. (2025). *Teaching a mixed level disparate class* [Online post]. Center for Teaching & Learning.

Videla, R., Aros, M. B., Parada, F., Kausel, L., Sandoval-Obando, E., Jorquera, D., Ibacache, D., Maluenda, S., Rodríguez-Herrero, P., Cerpa, C., González, M. J., Chávez, M., & Ramírez, P. (2025). Neurodiversity: Post-cognitivist foundations of the 3E approach for educational inclusion of autistic students with technology. *Frontiers in Human Neuroscience*, 18, 1493863. https://doi.org/10.3389/fnhum.2024.1493863

Vygotsky, L. (1978). *Mind in society: the development of higher psychological processes*. Harvard University Press.

Wibowo, G., Gunawan, D., & Mardiana, D. (2025). *Implementasi pendekatan pembelajaran mendalam (deep learning) dalam meningkatkan pemahaman konsep siswa di sekolah dasar*. 10.

Wibowo, H. A., & Waluya, B. (n.d.). *Menggugah kesadaran pedagogis: peran guru dalam menghindari surface learning dan mendorong deep learning.*

Wiggins, G., & Mctighe, J. (2005). *Understanding by design*. Association for Supervision and Curriculum Development.

Yin, R. K. (2016). *Qualitative research from start to finish* (Second edition). The Guilford Press.