

CHAPTER I

INTRODUCTION

A. Background of the Problem

Assessment is a crucial component of English language education, as it plays an essential role in identifying students' needs, measuring learning progress, and informing instructional strategies. It involves a continuous process of collecting, analyzing, and using information to improve learning outcomes¹. Traditionally, English language assessment in Indonesia has emphasized standardized testing, often focusing on discrete linguistic elements such as grammar and vocabulary. However, recent pedagogical approaches advocate for more authentic assessments that support language use in real contexts and emphasize learning improvement rather than judgment.

In response to the evolving demands of education, the Indonesian Ministry of Education, Culture, Research, and Technology introduced the *Merdeka* Curriculum, which provides schools and teachers with greater autonomy to design learning that is flexible, contextualized, and aligned with students' needs. It prioritizes

¹ H. Douglas Brown, *Language Assessment: Principles and Classroom Practices* (New York: Pearson Education, 2004). 4

foundational learning, empowers teachers to enhance instructional practices, and aims to create a joyful school experience². According to Regulation of the Minister of Education, Culture, Research, and Technology Number 12 of 2024³, *Merdeka* Curriculum is a curriculum that provides flexibility and focuses on essential materials to develop students' competencies as lifelong learners with Pancasila values. This curriculum is based on the principles of character development, flexibility, and essential learning content.

Assessment within the *Merdeka* Curriculum framework functions not solely as a means of measuring learning outcomes, but as an essential and continuous component of the instructional process. In accordance with the official regulation, "Assessment is utilized at the beginning, during, and end of the learning process to understand students' learning needs and the development of the learning process they have undergone" (Regulation of the Minister of Education, Culture, Research, and Technology No. 12 of 2024). This perspective underscores the role of assessment that enable teachers to adjust

² OECD, "Education in Indonesia a Decade on Transforming Education in Indonesia: Examining the Landscape of Current Reforms," *OECD Education Policy Perspectives* (Paris, 2024), <https://doi.org/https://doi.org/10.1787/9ff8d407-en>. 6

³ The Ministry of Education, Culture, Research, and Technology "Regulation of the Minister of Education, Culture, Research, and Technology of the Republic of Indonesia Number 12 of 2024," 2024, jdih.kemdikbud.go.id. 1-6

instruction to support continuous student development. According to Anggreana et al.⁴, the *Merdeka* Curriculum emphasizes four key principles of assessment: (1) assessment should be used to improve learning, not just to measure it, (2) assessment should be fair to all students and should produce consistent results, (3) assessment results should be clear and easy to understand, and should provide useful information to students, parents, and teachers, and (4) assessment results should be used to refine the quality of teaching and learning. These principles provide the foundation for how English language assessment should be carried out within the curriculum.

Nevertheless, the implementation of English language assessment in the *Merdeka* Curriculum faces various challenges. Although teachers are encouraged to develop authentic assessments, in reality, many still rely on traditional formats due to limited training, lack of resources, or uncertainty about how to design and interpret meaningful assessment tasks⁵. The challenges also include the fact that teachers now have more responsibility to craft creative teaching

⁴ Yogi Anggreana et al., *Panduan Pembelajaran Dan Ascsmen Pendidikan Anak Usia Dini, Pendidikan Dasar, Dan Menengah* (Ministry of Education, Culture, Research, and Technology, 2024), <https://kurikulum.kemdikbud.go.id/tentang>. 4-9

⁵ Robert Randall et al., *THE LEARNING GAP SERIES-TWO Reforming Indonesia's Curriculum: How Kurikulum Merdeka Aims to Address Learning Loss and Learning Outcomes in Literacy and Numeracy* (Jakarta: INOVASI, 2022), <https://www.inovasi.or.id>. 5

materials, learning goals, and the sequence of those goals⁶. Furthermore, Wuwur⁷ showed that the *Merdeka* Curriculum presents significant hurdles for teachers throughout planning, teaching, and assessment due to limitations in creativity, resources, and school environments. This reflects a gap between theoretical expectations and classroom practices, especially in secondary education, where teachers often struggle to balance curriculum autonomy with institutional demands.

Accordingly, the evaluation of English assessment implementation within Indonesia's *Merdeka* Curriculum is crucial to assess its effectiveness. To evaluate the impact of English assessment in the *Merdeka* Curriculum, the Kirkpatrick Model⁸, which outlines four levels of evaluation (Reaction, Learning, Behavior, and Results), provides a practical framework. Beginning with assessing teacher's and students' reactions, the evaluation would then measure the perception of teachers and students on the *Merdeka* Curriculum's assessment

⁶ M M Jannah and H Rasyid, "Kurikulum Merdeka: Persepsi Guru Pendidikan Anak Usia Dini," *Jurnal Obsesi: Jurnal Pendidikan Anak Usia Dini* 7, no. 7 (January 11, 2023): 197–210

⁷ Erwin Simon Paulus Olak Wuwur, "Problematika Implementasi Kurikulum Merdeka Di Sekolah Dasar," *Jurnal Ilmu Pendidikan* 3, no. 1 (April 2023): 1–9, <https://ejurnal.politeknikpratama.ac.id/index.php>.

⁸ Donad L. Kirkpatrick and James D. Kirkpatrick, *Implementing The Four Levels: A Practical Guide for Effective Evaluation of Training Programs*, First (San Fransico: Berrett-Koehler Publishers, Inc., 2007). 23-107

practices. Subsequently, the evaluation would examine changes in and students' behaviors in the classroom, specifically focusing on how assessment is integrated into teaching and learning. Finally, the evaluation would analyze the results of the new assessment practices, including improvements in student English proficiency and the overall effectiveness of the curriculum.

While educational research on assessment in the *Merdeka* Curriculum continues to grow, most studies focus on general implementation or conceptual frameworks. Research that specifically examines English language assessment practices under this curriculum remains limited. As of 2024, the curriculum has not been implemented nationally. According to Anindito, Head of the Education Standards, Curriculum, and Assessment Agency, the *Merdeka* Curriculum, developed since 2020 and implemented and evaluated in stages since 2021, has been adopted by approximately 80% of Indonesian schools, including 6,000 in disadvantaged areas, with the remaining schools having until the 2027/2028 academic year to transition⁹. Consequently, further research is required to understand comprehensively how

⁹ Ester Lince Napitupulu, "Kurikulum Merdeka Resmi Jadi Kurikulum Nasional," *Kompas*, March 27, 2024, <https://www.kompas.id/baca/humaniora/2024/03/27/kurikulum-merdeka-resmi-jadi-kurikulum-nasional>. 1

English language assessment is implemented, and how it can be evaluated.

These circumstances highlight the necessity of conducting a focused investigation on the implementation of English language assessment in the context of the Kurikulum Merdeka. Accordingly, the present study is titled “Evaluating the Implementation of English Language Assessment in the *Merdeka* Curriculum (An Explanatory Sequential Mixed-Methods Study at SMA Islam Al-Azhar 6 Cilegon Academic Year 2024/2025).”

B. Identification of the Problem

Based on the background above, the problems identified in this study are:

1. Limited research on how students and teachers perceive the implementation of English language assessments in the *Merdeka* Curriculum.
2. Lack of clarity regarding how English language assessment contribute to students’ perceived gains in English language skills.
3. The need to explore the extent to which students exhibit changes in English language usage as a result of the assessment.

4. Concerns about the overall impact of English language assessments in the *Merdeka* Curriculum on students' proficiency.

C. Formulation of Problem

This study aims to answer the following questions:

1. How do students and teacher perceive the implementation of English language assessment in the *Merdeka* Curriculum?
2. In what ways English language assessment within the *Merdeka* Curriculum contribute to students' gains in English language skills?
3. How do students exhibit changes in their English language usage, following the implementation of English language assessments in the *Merdeka* Curriculum?
4. What is the overall impact of the English language assessments practices on students' English proficiency within the *Merdeka* Curriculum?

D. Limitation of the Problem

To maintain focus, this study limits itself to the following areas:

1. Students' and teacher's perceptions of the implementation of English language assessment in the *Merdeka* Curriculum.

2. The contribution of English language assessment within the *Merdeka* Curriculum to students' gains in English language skills.
3. Observed changes in students' language usage following the implementation of English language assessment in the *Merdeka* Curriculum.
4. The overall impact of the English language assessment on students' English proficiency within the *Merdeka* Curriculum.

E. Objectives of Research

This study seeks to:

1. Explore how students and teacher perceive the implementation of English language assessment under the *Merdeka* Curriculum framework.
2. Investigate the contribution of English language assessment within the *Merdeka* Curriculum to students' gains in English language skills.
3. Examine how students exhibit changes in their English language usage, following the implementation of English language assessments in the *Merdeka* Curriculum.

4. Explore overall impact of the English language assessment on students' English proficiency within the *Merdeka* Curriculum.

F. Significance of Research

Firstly, this study contributes academically by offering empirical evidence on the implementation of English language assessment within the framework of the Merdeka Curriculum. While previous studies have broadly explored the curriculum's design and challenges, few have focused specifically on how assessment practices are experienced and interpreted by both teachers and students. The insights derived from this research are expected to enrich the academic discourse on curriculum evaluation and language assessment in the Indonesian secondary school context.

Secondly, this research serves as a valuable resource for English teachers and school administrators by providing practical implications related to assessment design and implementation. The findings can inform teaching strategies, especially in selecting and applying assessment methods that align with the curriculum's emphasis on flexibility, student-centered learning, and holistic development. Therefore, this study may support teachers in refining their assessment techniques to better serve diverse learners.

Thirdly, the study offers benefits for curriculum developers and policymakers. The evaluation of students' and teacher's responses to current assessment practices may highlight areas that require further support or policy refinement. By identifying both strengths and gaps in the assessment process, the research can guide the development of curriculum enhancements.

Finally, this study provides a foundation for future research in English language assessment and curriculum evaluation. Its methodological approach and findings can be used as a reference point for scholars seeking to investigate the effectiveness of educational reforms in similar contexts. As the Merdeka Curriculum continues to evolve, such research is crucial in ensuring that its implementation remains aligned with educational goals and responds to the real needs of schools.

G. Previous Research

Examining the existing literature is crucial for developing a deeper understanding of English assessment practices in the context of the *Merdeka* Curriculum. Therefore, the researcher presents several studies offering an overview of the *Merdeka* Curriculum.

Arifa¹⁰ conducted a policy-based study under the Research Center of the Indonesian House of Representatives, aiming to analyze the implementation and challenges of the *Merdeka* Curriculum. Her descriptive-analytical report utilized secondary data from government documents, national statistics, and expert commentary. The study highlighted major obstacles in the implementation process, particularly concerning the readiness of human resources. Teachers were found to lack alignment in terms of competence, pedagogical skills, and mindset with the demands of the new curriculum. The paper also emphasized infrastructural and budgetary constraints, suggesting that the *Merdeka* Curriculum's success depends heavily on systematic support from the government. However, Arifa's study did not involve primary data collection nor focus specifically on assessment practices, particularly English language assessment. In contrast, this research, *Evaluating The Implementation of English Language Assessment in the Merdeka Curriculum*, adopts an evaluative design using Kirkpatrick's Four-Level Model to evaluate the implementation of English assessment under the *Merdeka* Curriculum.

¹⁰ Ficka Nurul Arifa, "Implementation of Independent Curriculum and Its Challenges," *Info Singkat* XIV, no. 9 (2022): 22–30, <http://puslit.dpr.go.id>.

Rahina and Syamsi¹¹ conducted a descriptive study to examine the perceptions of Indonesian language teachers at the junior secondary level (SMP/MTs) in Yogyakarta regarding the implementation of the *Merdeka* Curriculum. The study employed a combination of questionnaires, interviews, and classroom observations involving 15 teachers as respondents. Their findings indicated that teachers' overall perception of the *Merdeka* Curriculum was at a “medium” level, with an aggregated score of 66.7%. While the respondents generally supported the curriculum's vision and were moderately prepared for its implementation, the study also revealed lingering uncertainty and the need for further readiness, especially as the curriculum was still in its early stages of national rollout. Unlike this study, which focuses specifically on English language assessment within the *Merdeka* Curriculum using Kirkpatrick's Four-Level Evaluation Model, Rahina and Syamsi's research offered a broader view of teachers' perceptions without addressing specific assessment practices or evaluating outcomes in a structured framework.

¹¹ Annisa Cahya Rahina and Kastam Syamsi, “Independent Curriculum in the Perception of Indonesian Language Teachers of Smp/Mts in Yogyakarta,” *International Journal of Linguistics, Literature and Translation* 6, no. 2 (February 4, 2023): 54–57, <https://doi.org/10.32996/ijllt>.

Additionally, Astuti et al¹² conducted a qualitative study to explore challenges in implementing assessment under the *Merdeka* Curriculum. The study involved teachers from SDN 1 Melinggih Kelod, Gianyar Regency, Bali, which had adopted the new curriculum. Data were collected through observation, interviews, and expert discussions, and analyzed using an interactive model that included data reduction, data display, and conclusion drawing. The study identified two main areas of concern: intracurricular assessment and the Project for Strengthening the Pancasila Student Profile (P5). In intracurricular assessment, teachers struggled to distinguish between formative and summative assessments, with many relying solely on written tests for formative purposes and misunderstanding the development process of summative assessments. Furthermore, teachers still viewed report cards as the sole indicators of student achievement. As for the P5 project, teachers mistakenly perceived its assessment as a final-stage activity conducted only during showcase events, which led to additional challenges, including financial constraints in organizing those events. While this study provides insight into assessment-related

¹² Ni Putu Eni Astuti et al., "Permasalahan Asesmen Pada Kurikulum Merdeka," *Cetta: Jurnal Ilmu Pendidikan* 7, no. 1 (2024): 22–32, <https://jayapanguspress.pencrbit.org/index.php/cetta/article/view/2954>.

misconceptions in early implementation contexts, it does not specifically investigate English language assessment. In contrast, this study focuses on English assessment within the *Merdeka* Curriculum.

In line with that finding, Rosidah, et al¹³ employed a descriptive approach through a literature review, supplemented with a Focus Group Discussion (FGD) and brief interviews, to investigate teachers' readiness in implementing authentic assessment within the *Merdeka* Curriculum. The study engaged 50 teachers from Gondang, Mojokerto, East Java, and explored aspects of assessment preparation, implementation, and reporting. Findings revealed that although teachers demonstrated a theoretical understanding of authentic assessment, they encountered substantial difficulties in practice. Challenges included the perceived complexity of constructing and converting rubrics, confusion when rubrics spanned multiple subjects, and the absence of formal training on authentic assessment planning. As a result, the actual implementation often did not align with the intended standards, and teachers continued to rely on rubrics from the teacher's handbook without further adaptation. While this study addresses real-world

¹³ Cholifah Tur Rosidah, Pana Pramulia, and Wahyu Susiloningsih, "Analisis Kesiapan Guru Mengimplementasikan Asesmen Autentik Dalam Kurikulum Merdeka Belajar," *Jurnal Pendidikan Dasar*, 2021, 87–103, <https://doi.org/doi.org/10.21009/JPD.012.08>.

barriers to authentic assessment, it does not evaluate specific subject-area assessments or analyze implementation outcomes using an established evaluation framework. By contrast, the study, *Evaluating The Implementation of English Language Assessment in the Merdeka Curriculum*, focuses on English language assessment and employs Kirkpatrick's Four-Level Model to systematically evaluate how the *Merdeka Curriculum* is experienced and enacted at various outcome levels.

In summary, the reviewed studies have explored various aspects of the *Merdeka Curriculum*, such as teacher readiness, perceptions, and challenges in implementing assessment. While these studies highlight important issues, most of them focus on general curriculum implementation, different subject areas, or early-stage teacher understanding, without evaluating assessment practices in a structured manner. Moreover, none of the studies specifically investigate English language assessment or examine its effectiveness using a comprehensive evaluation framework. To address this gap, the present study applies Kirkpatrick's Four-Level Evaluation Model to systematically evaluate the implementation of English assessment in

the Merdeka Curriculum, drawing on data from questionnaires, interviews, observations, and student scores.

H. Novelty

Research on the *Merdeka* Curriculum has primarily centered on its general implementation, teacher readiness, and pedagogical challenges. However, studies that specifically examine the implementation of English language assessment within this curriculum remain limited. This study seeks to fill that gap by focusing explicitly on how English assessment is perceived by both students and teacher under the *Merdeka* Curriculum framework.

Furthermore, this research offers a novel contribution by employing the Kirkpatrick Evaluation Model to analyze the assessment's implementation and impact across four levels: reaction, learning, behavior, and result. By adopting this framework, the study not only evaluates assessment outcomes but also captures perceptions, behavioral changes, and proficiency development holistically—an approach that has not been widely applied in the context of Indonesian curriculum evaluation.