

Exploring Teacher Assessment Literacy in Practice within the Implementation of *Kurikulum Merdeka*: A Case Study of EFL Teachers in Temanggung Regency

Wahyu Puji Lestari
Universitas Negeri Semarang
wahyulesta123@students.unnes.ac.id

Prof. Puji Astuti, S.Pd., M.Pd., Ph.D.
Universitas Negeri Semarang
puji.astuti.ssu@mail.unnes.ac.id

Abstract

Many teachers still find difficulties in practicing the paradigm of *Kurikulum Merdeka*, specifically regarding assessment. Several studies have identified the relationships between teachers' assessment literacy and their practices. However, investigation into those two fields according to the *Kurikulum Merdeka* are still scarce. Therefore, this study attempts to explore the EFL teachers' assessment literacy in practice regarding the implementation of the *Kurikulum Merdeka*, as well as the challenges and institutional support for teachers. The study employs qualitative design in form of exploratory case study basing on Xu and Brown's (2016) model framework of teacher assessment literacy in practice (TALiP). Three teachers and principals were involved as the main participants. This study found mixed findings for each aspect of TALiP and there was connection between teachers' conceptions, knowledge, and practices of assessment. Teachers were found to have implemented the assessment as instructed in the *Kurikulum Merdeka*, although not yet optimal. Furthermore, contextual factors affecting the practices were mainly in form of external factors, such as time constraint, excessive workload, and diverse students. Therefore, schools have attempted to provide supports by promoting professional development program, optimizing school community of practice, developing communication and collaboration, building reflection culture, and increasing teacher as well as student motivation. The findings of this study are expected to provide inputs for educational stakeholders in order to find solutions for better implementation of the *Kurikulum Merdeka*, especially in the language assessment field in a larger context.

Keywords: assessment literacy, assessment practice, conception of assessment, *Kurikulum Merdeka*

INTRODUCTION

Teachers can conduct assessments almost at anytime, anywhere, and in any way. The new educational policy in Indonesia reveals a new paradigm about assessment that should be practiced in classrooms, meaning that the controversial and high-stakes national examination in the Indonesian education system is terminated (Astuti et al., 2023). Although some concepts regarding the approach, methodology, and paradigm in the *Kurikulum Merdeka* are nothing new, some teachers expressed their

inconsistency in implementing them in classrooms. Many teachers find difficulties in comprehending and implementing the curriculum instructions, especially on assessment.

In 2021, the Ministry of Education, Culture, Research, and Technology has officially adopted a new paradigm for the curriculum labeled as the *Kurikulum Merdeka Belajar* or the Emancipated Curriculum (hereinafter referred to as the *Kurikulum Merdeka*). The shift from the 2013 Curriculum to the *Kurikulum Merdeka* was based on several previous studies that

evaluated the implementation of the 2013 Curriculum. Further in 2025, the Ministry has issued new regulations regarding national educational policy with an emphasis on the Deep Learning approach. However, this study regarding the *Kurikulum Merdeka* remains relevant because the Ministry states in Deep Learning Academic Paper (2025), that Deep Learning was proposed as a refinement of the *Kurikulum Merdeka* (pp. 2-3) in which the emphasis of assessment recommendation still lies on the formative and summative assessment. Nevertheless, the current curriculum is preferred to be called as 'National Curriculum'.

Regarding the previous curriculum, Widodo (2016) explained several drawbacks of the 2013 Curriculum, including a) the nature of the curriculum is highly prescriptive in that it dictates what and how to teach and learn English; b) the assessment still prioritizes cognitively demanding assessment even though it emphasizes both process- and product-based assessment; c) the curriculum puts emphasis on idealized guidelines, which do not recognize crucial elements of what the curriculum means; and d) curriculum does not detail curriculum materials, pedagogy, and assessment, informed by relevant theories of language, language learning, and language teaching. Along with this, Marijan in Anugrahwati and Agustien (2015) stated that the problems faced by teachers in the 2013 Curriculum are low teacher quality, inadequate infrastructure, and a curriculum that has not been standardized. In results, Pantiwati et al. (2023) summarized Kurka (2022) and Nasution (2022), "Four main education policy programs will become future learning directions: the school-based national exam, the national exam replaced with a minimum competency assessment and character survey, a learning implementation plan, and atoning for new student admissions". This is in line

with the paradigm of the *Kurikulum Merdeka*.

Naturally, the curriculum reforms would affect the implementation of teaching and learning in classrooms, one of them is assessment. Widodo (2016) defines assessment as a way to recognize students' agency and provide useful input for innovating or reframing the existing language pedagogy. Regarding the curriculum policy, Alsaadi (2021) states that successful ways of assessment can lead to radical improvements in the curriculum and teaching methodologies. In the dynamic process of foreign language learning and teaching as a crucial educational field of study, language assessment and testing can be described as an indispensable part. Further, language policy in the *Kurikulum Merdeka* suggests teachers to adopt Genre-Based Text approach and other approaches to teach with B1 level CEFR goals for senior high school students and language assessment in this curriculum is supposed to be dynamic; meaning there should be an integration between learning and assessment by emphasizing mostly on process or formative assessments.

Several previous studies showed that there were significant relationships between teachers' beliefs/conceptions of assessment on their practices in the classroom. Teachers viewed assessment positively (Roslan et al., 2022; Unal & Unal, 2019) and that teachers' beliefs about assessment influenced their practices in the classroom, including instructional and grading decisions (Barnes et al., 2015; Cheng & Sun, 2015; Daniels & Poth, 2017; Takele & Melese, 2022). This is proven by the conformity between teachers' beliefs and practices in the classroom (Mamad & Vigh, 2021; Mamad & Vigh, 2023). The positive conceptions of assessment by teachers are related with assessment as tools for improving teaching and learning or improvement conception, accountability (school and student), and also

assessment as irrelevant (Azis, 2015; Barnes et al., 2017; Brown et al., 2011; Calveric, 2010; Elshaw, 2016; Gurmesa et al., 2022; Kunnath, 2017; Roslan et al., 2022; Takele & Melese, 2022). Brown et al. (2009) found that there are statistically and practically significant relations between teachers' belief and thinking about the nature and purpose of assessment and the practices they conduct.

On the other hand, some studies revealed that teachers' practices in the classroom do not reflect their beliefs or conceptions of assessment (Baidoo-Anu et al., 2023; Chew & Lee, 2013; Gurmesa et al., 2022; Ha et al., 2021; Jannati, 2015; Muliya & Aminatun, 2020; Narathakoon et al., 2020; Widiastuti et al., 2020). These incongruences were due to the teachers' different perspectives on assessment (Berry et al., 2019; Dashti, 2019), which can be caused by various factors and challenges, such as (a) language assessment literacy or knowledge (Fang & Yu, 2022; Farhady & Tavassoli, 2021; Jannati, 2015; Moges, 2018; Narathakoon et al., 2020; Wu et al., 2021); (b) educational background and professional development (Battistone et al., 2019; Coombs et al., 2022; Dashti, 2019; Widiastuti et al., 2020); (c) culture and policies; (d) students' characteristics and needs (Muliya & Aminatun, 2020; Wu et al., 2021). These factors can be summarized into what Chen and Bonner (2017) called as 'individual, institutional, and societal level' factors. This concept is in line with Barnes et al. (2015) and Brown et al. (2019) who suggested that teacher assessment conceptions are not universal, they are complex and multifaceted constructs, shaped by historical, social, and educational factors, in which teachers conduct their teaching and learning practices.

The literatures have identified several key points regarding language assessment practices by either teachers or teacher candidates. The areas include the relationship between teachers' perception with their

practices of assessment, teachers' assessment literacy (knowledge) with their practices, teachers' grading decisions, practices of teacher-based assessment (TBA), and teachers' perceptions of policy/curriculum reform. The previous researchers have identified that there are relationships between teachers' assessment literacy and their practices, be it linear or contradictory.

However, the investigation on teacher assessment literacy and practice in Indonesia within the framework of the *Kurikulum Merdeka*, especially for English or language subjects are still scarce. To answer that, this research attempts to discuss the EFL teachers' assessment literacy in practice in regards to the *Kurikulum Merdeka*, including teachers' conceptions, knowledge, practices, and the contextual factors affecting their practices in the context of senior high schools in Temanggung Regency, Central Java. The context of Temanggung Regency was chosen because many EFL teachers in this area –in which the first researcher is currently teaching– confessed that they haven't fully implemented the curriculum paradigm in their teaching practices. Besides, the English teacher community haven't been activated after COVID-19 pandemic that we thought this study could be beneficial to investigate the representatives of these teachers as reflection and further necessary professional development for the optimal implementation of the curriculum. Further, the theoretical framework employed in this study was based on the teacher assessment literacy in practice (TALiP) model by Xu and Brown (2016).

METHODOLOGY

This design of this study is qualitative research. The nature of this study is interpretative, naturalistic and exploratory because the study investigates practitioners' experiences in a specific setting. Therefore, this study employs exploratory case study

design as a means to define the necessary questions and hypotheses for developing consecutive studies (Yin, 2018). He further highlights case study research's supportive role in developing continuative social research in general. However, due to its intuitive approach, to triangulate the results obtained from the qualitative data, to validate and to explore the different aspects of the phenomenon to get more detailed information about the issue (Creswell, 2014), quantitative data were collected. Thus, this study makes use of methodological triangulation. The questionnaire was used for collecting quantitative data, whereas the interviews, observations, and document analysis for qualitative data.

The participants included in the questionnaire stage was thirty EFL teachers joining the teaching community (MGMP) of Temanggung Regency. After the specified time provided for responding, nine teachers submitted their responses. From these nine teachers, three teachers were selected as the final participants for interviews and classroom observations as below.

Table 1. List of Final Participants

Teacher Code	School Type	Sex	Education- al Back- ground	Years of Teaching	Training Experi- ence
T1	Public	Female	Bachelor	Between 2 and 5	None
T4	Public	Male	Bachelor	More than 10 years	None
T6	Public	Female	Master	More than 10 years	½ to 1 day Workshop or Seminar

Along with these three teachers, three principals where these teachers teach also automatically became the participants for principal interviews. These participants were coded as T1, T4, and T6 (teachers), as well as P1, P4, and P6 (principals).

The questionnaire responses were analyzed with descriptive statistics, where frequency counts, percents, and means.

Thereafter, the questionnaire items were analyzed twice, using the SPSS 30 software, and then the collated data was verified twice. Next, thematic analysis was used to analyze both the qualitative data (interviews and observations) in order to uncover teachers' knowledge, thoughts, and issues regarding testing and assessment in their curriculum innovation context. An inductive approach to analyzing both interview and observation transcripts was adopted. The collected data were further coded into six themes to answer the research questions including conceptions, knowledge base, self-perceived practices of assessment, perception on the *Kurikulum Merdeka*, assessment practice in regards to the curriculum instructions, and contextual factors consisting of challenges and institutional support.

FINDING AND DISCUSSION

Simultaneous Multiple Conception of Assessment (CoA)

As stated by Brown (2004, 2008), the conception of assessment framework has three major purposes of assessment and one counter-purpose: (a) improvement (IMP), (b) school accountability (SQ), (c) student accountability (SA), and (d) assessment is irrelevant (IRREL). This component was measured using the TCoA-IIIA Inventory from Brown, Lake, and Matters (2011). Table 1 shows the statistics of the purposes, while Table 2 the counter-purpose of this component.

Table 1. Statistics of Purposes of CoA

		SQ	SA	IMP
N	Valid	9	9	9
	Missing	0	0	0
Mean		11.78	12.33	48.78
Std. Deviation		1.302	1.500	3.898
Variance		1.694	2.250	15.194
Range		3	4	10
Minimum		10	10	44
Maximum		13	14	54

Table 2. Statistics of Counter Purpose of CoA

N	Valid	IRREL
	Missing	
		9
		0
Mean		21.1111
Std. Deviation		2.84800
Variance		8.111
Range		8.00
Minimum		16.00
Maximum		24.00

Improvement

There are four aspects within the improvement purpose: diagnose, reliability, student improvement, and teacher improvement. The responses from the three respondents were ranging from 'moderately agree' to 'strongly agree', except two responses with 'mostly disagree' (one from T1 regarding teacher improvement and the other from T6 regarding reliability). Overall, it can be said that all respondents possessed positive conceptions about the improvement purpose of assessment, especially about assessment helping students' development and the learning process, as well as the assessment accuracy and reliability.

School Accountability

The participants' conception on the school accountability purpose were mostly positive for the responses were 'mostly agree' to 'strongly agree' to all statements, except for T4 who responded with 'moderately agree' to the statement 'Assessment is a good way to evaluate schools'. Therefore, with no negative response to this purpose, it can be said that the teachers perceived that students' assessment results could become a parameter to evaluate and determine the school quality.

Student Accountability

It was found that two of the final respondents obtained the lowest score in this aspect, T1 and T4, while T6 had the highest score compared to the nine initial respondents. T1

mostly disagreed with the statement of categorizing students based on the assessment result, T4 moderately agreed, while T6 strongly agreed with this statement. Further, T4 moderately agreed with the statement of assigning grade or level to student work, while the other two mostly agreed. Lastly, the three respondents mostly agreed with the statement of assessment determines if students meet qualifications standards. From this description, it can be inferred that there are slightly different conceptions between the respondents regarding the student accountability purpose.

Assessment is Irrelevant

In this regard, assessment is viewed as activity that interferes with teaching, has little impact on teaching, and should be taken into account for any imprecision and error. Further, it views assessment as unfair to students, the results are ignored, and teachers make little use of them. From the analysis, the three respondents' conceptions were generally negative on this aspect because most responses were 'strongly disagree'. Hence, for statements 'Assessment forces teachers to teach in a way against their beliefs' and 'Assessment results should be treated cautiously because of measurement error', the respondents answered with 'mostly agree'. Then, a wide gap appeared for statement 'Assessment is unfair to students', in which T1 and T4 strongly disagreed, whereas T6 -the most senior teacher- strongly agreed with it. To conclude, the respondents believed that assessment do not interfere with teaching, it has impacts on students' learning, it is not an imprecise process, and that assessment results are to be made use of instead of filed and ignored. However, there was also a perception that assessment is unfair for students.

To conclude, it was found that the teachers' conceptions of assessment varied.

Positive conceptions were found in improvement and school accountability, mixed conceptions in student accountability, and mostly negative conceptions in irrelevance. This result is consistent with the previous studies stating that teachers can have differed as well as multiple beliefs or perceptions about assessment simultaneously (Azis, 2015; Barnes et al., 2017; Brown et al., 2011; Calveric, 2010; Gurmesa et al., 2022; Monteiro, Mata, & Santos, 2021; Muianga, 2023; Takele & Melese, 2022, Wicking, 2017).

Differing Levels of Teachers' Knowledge Base of Assessment and Its Underlying Factors

The questionnaire items for this component is adapted from the Language Assessment Literacy (LAL) scale developed by Mohammadkhah et al. (2022) aligned with the concept of teacher's knowledge base by Xu and Brown (2016). The developed scale includes four distinct aspects including: (a) language assessment in pedagogical content knowledge (PCK); (b) assessment theories and principles (ATP); (c) scoring and interpretations (SI); and (d) technical skills (TS). Table 3 shows the statistics of the responses on teacher knowledge base component.

Table 3. Statistics of Teacher Knowledge Base

		PCK	ATP	SI	TS
N	Valid	9	9	9	9
	Missing	0	0	0	0
Mean		28.6667	17.8889	14.7778	32.8889
Std. Deviation		4.06202	3.51584	1.85592	4.31406
Variance		16.500	12.361	3.444	18.611
Range		13.00	11.00	4.00	12.00
Minimum		21.00	10.00	12.00	25.00
Maximum		34.00	21.00	16.00	37.00

Language Assessment in Pedagogical Content Knowledge

This aspect is related to how teachers must possess the knowledge of the disciplines and how to teach or deliver the content to the students. In short, teachers must know how to assess the content of the school curriculum. Out of the three respondents, T6 obtained the highest score with 'knowledgeable' responses to all items of questionnaire, meanwhile T4 had the lowest score with 'slightly knowledgeable' and 'moderately knowledgeable' responses. The contrast of responses appeared in the statement about 'knowing how major language learning theories and trends affect selection of assessments trends', T2 and T4 were slightly knowledgeable, while T6 knowledgeable. Further, in statement 'training my students how to peer-assess each other's performance', T1 and T6 were knowledgeable, while T4 was slightly knowledgeable.

Lastly, in the statement 'knowing how to statistically evaluate language assessments regarding their reliability or validity', the three respondents responded differently. T1 with 'slightly knowledgeable', T4 'moderately knowledgeable', while T6 'knowledgeable'. Therefore, in the aspect of pedagogical content knowledge, it was found that the three teachers have different levels of knowledge. Relating the responses with the participants' years of teaching, it showed that the senior teacher (T6) had more knowledge in this aspect rather than the junior ones (T1 and T4). Nevertheless, a gap appeared in T4's responses, despite his experience (more than 10 years of teaching as well as having the deputy of curriculum board), he was slightly knowledgeable in this area.

Assessment Theories and Principles Knowledge

This aspect concerns with teachers' knowledge about the theories and principles

of assessment such as assessment bias, assessment protocols, and what a variety of relevant assessment strategies are. The score obtained by the three respondents also varied in responding to five items or statements to measure this aspect. T1 had the lowest score, followed by T4, and the highest score obtained by T6. The responses by T4 were all 'moderately skilled' and T6's responses were all 'skilled/knowledgeable'. Meanwhile, T1 had two items with 'not skilled at all' responses, in which the items are concerned with assessment bias and how to accommodate candidates with disabilities or other learning impairments. From this result, it can be seen that teachers have different levels of knowledge and/or skills regarding the theories and principles of assessment. Further, the years of experience also plays a role in determining the teachers' knowledge in this aspect.

Scoring And Interpretations Knowledge

This is an understanding of scoring techniques for objectively-marked testing and principles of consistency and moderation for judgment-based assessments, ways of interpreting evidence generated from assessment, as well as communicating assessment results to stakeholders (Xu & Brown, 2015). Out of the three respondents who responded to four items in the questionnaire, T6 had the highest score followed by T1 and T4 who had similar scores. T6's responses were all 'skilled/knowledgeable', T4's with all 'moderately skilled/knowledgeable' responses, while T1 had one response of 'slightly skilled/knowledgeable' regarding statement 'understanding concepts of technical reports and terminologies of assessments'. In this respect, it can also be seen that the teachers have different levels of knowledge and/or skills in assessment scoring and interpretation. Similar to the

previous aspect, the years of experience also takes part in this aspect.

Technical Skills

This aspect deals with the methods and variations of language assessments in which teachers must possess to provide quality assessment for students. In tune with the previous aspect, in this aspect the respondents also showed different levels of knowledge/skills. T6's responses were all 'skilled/knowledgeable', while T4's responses were mostly 'moderately skilled/knowledgeable'. Hence, T1's responses were more varied with one 'not skilled at all' response regarding training others to write good quality items (questions) or tasks for language assessment' and three responses of 'slightly skilled' related to developing quality items and tasks for particular assessment purposes and using rating scales to score oral skills (e.g., speaking) as well as written skills (e.g. writing). Therefore, it can be concluded that there are different levels of knowledge possessed by the three teachers with the most senior teacher (T6) is more knowledgeable rather than T4 and T1, with T1 with the least skills regarding the technical skills.

All in all, in the knowledge base of assessment, teachers have different levels of knowledge/skills, but overall can be said to be in 'fair' level. From four aspects in this component, T1 the one with the least teaching years always placed last in the final score, while T6 with the most teaching years and a master's degree always placed first. Therefore, it can be inferred that years of experience, educational background, and training programs play a part in determining the teachers' knowledge/skills of assessment. This result is in line with previous studies stating that several factors affecting teacher's knowledge or literacy of assessment include years of experience educational background, and professional development program (Angraeni & Yusuf, 2022; Chen & Bonner,

2017; Jannati, 2015; Hutami & Putro, 2023; Sun & Zhang, 2022; Unal & Unal, 2019).

Incongruencies Between Teachers' Self-Perceived Practices and Their Conceptions of Assessment

The questionnaire items for this component is adapted from Teachers Practice of Assessment Inventory (TPrAI) India by Brown, Chaudhry, and Dhamija (2015). The questionnaire was developed to identify the degree to which teachers agreed with assessment practices that (a) diagnose student learning needs (Diagnostics-PrAI-D), (b) prove school quality (School Evaluation-PrAI-S), (c) prepare students for high-stakes examinations (Teaching for Exams-PrAI-E), and (d) ignore or treat assessment as irrelevant information (Ignore Exams-PrAI-IE). There are 29 items included in the questionnaire which results is presented in the table below.

Table 4. Statistics of Teacher Practices of Assessment

	PrAI-D	PrAI-S	PrAI-E	PrAI-IE
N Valid	9	9	9	9
Missing	0	0	0	0
Mean	24.4444	22.6667	49.2222	13.2222
Std. Deviation	1.74005	1.65831	3.45607	2.90593
Variance	3.028	2.750	11.944	8.444
Range	6.00	5.00	10.00	9.00
Minimum	23.00	20.00	43.00	7.00
Maximum	29.00	25.00	53.00	16.00

Diagnostics (PrAI-D)

The first aspect to be examined in this component is how teachers' assessment practices diagnose student learning needs by identifying students' strengths and weaknesses, how much they have learnt from the teaching, as well as their learning needs. The three respondents' scores were the same, which placed them in the 'medium' category. Most of their responses were 'mostly agree'

and 'strongly agree'. The slight differences appeared in the statement regarding 'using assessment results to predict future student performance' and 'using assessment to determine how much students have learnt from teaching'. T4 responded to those two statements with 'moderately agree', while T1 and T6 with 'mostly agree'. In this regard, it can be seen that the three respondents' assessment practice show positive alignment with the aspect of diagnosing student learning needs.

School Evaluation (PrAI-S)

Secondly, school evaluation aspect shows how school treats assessment or examination in connection with its quality. The responses to the six items in this aspect were mostly 'mostly agree'. The slight differences were visible in the responses for statement 'My school evaluates its performance mainly by public examination results' and 'Our school puts most effort in preparing students for public examinations'. Similar to the previous aspect, the 'moderately agree' responses came from T4 for both statements, while T6 responded with 'mostly agree' for both statements, and T1 with 'mostly agree' and 'strongly agree'. All in all, the teachers agree that their schools view assessment or examination as important matter related to school evaluation and quality, thus the schools through teachers are making sure that students are prepared for such assessments.

Teaching for Exams (PrAI-E)

In this aspect, there are thirteen items to be responded by the participants regarding their assessment practices which align with their teaching purposes in their classrooms, such as referring to the public examinations when conducting the assessment, taking into account students' test results to better the teaching process, as well as considering using alternative assessments to assess students with different abilities. The highest

score for this aspect, out of the three final respondents, is obtained by T1, followed by T6, and lastly T4. The responses varied from 'moderately agree' to 'strongly agree'.

However, one response from T4 for statement 'I take into account error and imprecision when using assessment results' was 'strongly disagree', while the other two with 'moderately agree'. For this reason, T4 belongs to 'low' category, while T1 and T6 belong to 'medium category'. Thus, it can be said that T1 and T6 align their assessment practice with their teaching purposes better than T4. Nevertheless, the three teachers had the same level of agreement on two statements, namely 'I use alternative assessments to assess different student abilities' and 'On discussing any inconsistency in students' assessment results, I will review their exam papers' with 'mostly agree' responses. Therefore, the teachers take into account students' different abilities and make sure that assessment results are consistent.

Ignore Exams (PrAI-IE)

In this aspect, teachers deem assessment or examination as something irrelevant to their pre-existing teaching plans because they have little time to pay deeper attention to the assessment results as well as focusing on completing the curriculum as their priority. From the analysis, it was found that there was a quite gap of difference in the responses. T1's responses were all 'moderately agree', T4's responses were more varied from 'moderately agree' to 'strongly agree', while T6 responded with one 'mostly agree' and three 'strongly disagree' responses; thus, placing T6 in the 'low' category. However, this shows that T6 treat assessment more deeply compared to T1 and T4 because she has more time to explain items tests after the exam, as well as do not make students passing examination and completing curriculum her top priorities.

Thus, it can be said that T1 and T4 had more tendency to ignore exams rather than T6. To repeat, T6 has the superior teaching experience rather than T1 and T4.

In conclusion, the teachers' practices of assessment show varied or mixed results. Teachers show positive responses for diagnose student needs and school evaluation, mixed response for teaching for exams in which the similarity found in taking into account student different abilities and assessment consistencies, and mixed response also for ignoring exams in which one teacher show less ignorance than the other two. The analysis result for this theme is congruent with the previous studies finding one of the factors that affects teacher assessment practices is years of experience (Calveric, 2010; Chen & Bonner, 2017; Unal & Unal, 2019; Yan & Fan, 2021).

Teachers' Mixed Perception on Assessment according to the *Kurikulum Merdeka*

On their perception of assessment according to the *Kurikulum Merdeka* instructions, the teachers came up with the similarities and differences between the current curriculum with the previous ones.

First of all, the similarity lies in the concept of assessment that had already been introduced even in the previous curricula.

Teacher 4: "*Assessment in the sense of making grades in my opinion in its implementation is no different. The assessment process is the same. ... If I observe from several curricula that I have experienced, the theory is the same and the process is the same. However, now it is rather emphasized that when teaching there must be a diagnostic assessment, actually it used to be done too, but it was not trumpeted.*" (Excerpt 1)

Meanwhile, some of the differences expressed by these teachers are the assessment aspects that are more emphasized, where the *Kurikulum Merdeka* emphasizes competency-based assessment and formative assessment, while the previous curriculum focused more on academic assessment with three aspects: cognitive, affective, and skills. The previous curriculum also focused on the standardizations, such as the test items and the minimum criteria for students' achievement.

Teacher 1: *"In this Kurikulum Merdeka, the emphasis or assessment is on competency-based assessment, whereas in the previous curriculum, the assessment was only based on tests given to students at the end (of a material), for example final tests or class promotions and so on."* (Excerpt 2)

Teacher 6: *"The Kurikulum Merdeka is more on the process, while the 2013 Curriculum is more on academic assessments, yes; structured assessments through written and oral exams. Generally speaking, the goals are automatically different. For the 2013 Curriculum, the goal of the assessment is more to improve students' academic abilities, while the Kurikulum Merdeka is more on developing students' character and morals because it is more on the process. ... The assessment results (in Kurikulum Merdeka) are used to design learning that is appropriate to the student's achievement stage, more focused on the needs of the students. While in the 2013 curriculum, the assessment results are used to determine learning progress and learning achievements that are in accordance with the minimum achievement criteria."* (Excerpt 3)

In addition, another difference is the format of official learning outcome reporting documents along with their descriptions.

Teacher 4: *"The only difference is the report formats, such as report card formats, diplomas, for example (differences in) descriptions, so it's only the technicalities (that are different)."* (Excerpt 4)

Furthermore, on the effectiveness of assessment according to the *Kurikulum Merdeka*, teachers expressed mixed perceptions. T1 felt that the assessment is effective because it can be implemented in the classroom and show the students' competencies and skills deeper with the correct implementation and planning.

Teacher 1: *In my opinion, the assessment framework in the Kurikulum Merdeka is clear, Ma'am, because I have also implemented it and it is effective. So in the assessment in the Kurikulum Merdeka we implement (assessment) as, for, and of learning, so that we can find out students' competencies more deeply, we know what students' abilities are like, so that we can meet the criteria or what is expected in the learning process."* (Excerpt 5)

T4 and T6 conveyed that the assessment can be either effective and ineffective. It can be effective if done correctly and fairly; however, it can be ineffective if the implementation is incorrect or the results of assessment interfered by third party so that the final score on the report do not reflect the students' real competencies.

Teacher 4: *"As far as I know, it can be said both effective and ineffective. It could be effective if the reporting can be read by a third party. Ineffective because what is in the writing or the numbers do not actually reflect the real students. So, I think it is ineffective"*

because it is not in accordance with the real-facts.” (Excerpt 6)

Teacher 6: Many factors influence whether it can be effective or ineffective. With diverse students, if we do not conduct an initial assessment, for example, then learning tends to be ineffective, because we will often reprimand children, it will be more on class management. However, if we have designed the assessment according to the Kurikulum Merdeka, aligned with the learning objectives that we will achieve, I am sure it will be more effective. The key here is because the Kurikulum Merdeka focuses more on the students’ needs, so we should inevitably design assessment that suits those needs. Thus, we will get the effectiveness of this assessment. So, for this effectiveness it depends on how we design it.” (Excerpt 7)

In conclusion, the three teachers did not see the assessment reform negatively. T1 as the most junior teacher left a positive note, T6 -with the most numbers of curricula in her past teaching experiences- expressed a neutral note, whereas T4 showed a typically coercive fidelity (Iskandar, 2020) by thinking that the curriculum is a mandatory top-down practice from the government.

The Alignment between EFL Teachers’ Practices of Assessment and Assessment Reform in the *Kurikulum Merdeka*

The last discussion on this study is the teachers’ assessment practices in real situations. Three teachers were observed during their teaching and learning activities in their classrooms and second interview. The observation guide was adapted from World Languages Office of University of Virginia’s Curry School of Education (2011).

After analysing these instruments and referring to the curriculum documents, we came up with six aspects of assessment according to the *Kurikulum Merdeka*, including: (a) Formative assessment; (b) Summative assessment; (c) Differentiation; (d) Teacher’s reflection for Improving learning process; (e) Conformity of assessment purpose with the technique; and (f) Scoring/Grading.

Table 5. EFL Teachers’ Assessment Practices according to the Implementation of the *Kurikulum Merdeka*

No.	Component	T1	T4	T6
1.	Formative Assessment	√	√	√
2.	Summative Assessment	√	√	√
3.	Differentiation	√	√	√
4.	Teachers’ Reflection for Improving Learning Process (Feedback)	√	√	√
5.	Conformity of Assessment Purpose and Design	√	√	√
6.	Scoring/Grading	√	√	√

Planning and Instructions

In this respect, T1 taught narrative texts, T4 exposition text, while T6 taught argumentative text. The three teachers showed good planning and delivery during the teaching and learning activities. This includes how the students experienced multiple language skill modalities and aspects in each meeting. For example, teaching grammar by giving contextual examples, explaining the expected outcome of the meeting, students experience multiple language skills and aspects in one meeting, using different strategies such as note-taking, groupings, gamification, etc. Thus, the flow of the teaching and learning activities were already planned well after doing a diagnostic assessment.

Assessment Practice

The three teachers were observed to have implemented the formative and summative assessments in their classrooms. The formative assessment conducted by teachers were in form of diagnostic assessment or pre-assessment and assessment throughout the learning process such as observation and quizzes. The media used for formative assessment also varied, from contextual, digital, to interactive ones. However, T4 relied mostly on the digital media for the activities, so that there were times when students felt uninterested when T4 asked them to have similar activities, even though they became motivated again at the process. The three teachers also showed the use of feedback for students, although mostly on the praises and encouragement. However

The summative assessment included performance task, such as writing and speaking as well as written test such as test for material comprehension at the end of the chapter. The types of assessment used by the three teachers were (1) product assessment, such as writing argumentative text by T4 and T6; (2) written test (objective test) for narrative text by T1, and (3) oral test in form of presentation by T6. During the interview, the teachers conveyed that the summative assessments would not be the only source of consideration for deciding on the students' final scores. They would also pay attention to the process during when the students were learning in the classrooms. Therefore, it can be said that the three teachers had practiced both formative and summative assessments, as what the *Kurikulum Merdeka* instructed.

The three teachers mentioned the importance of formative assessment (diagnostic assessment, the concept of assessment as/for/of learning) to aid them in improving the upcoming learning activities and to know what they should provide their students with during their first interview. Nevertheless, it was not explicitly

documented and intensely done in the real classroom. Lastly, the scoring and grading conducted by the teachers for any assessment teachers give to students either formative or summative do not only focus on exact numbers/scores, but also on the descriptions and notes on students' progress. This is done in order to find the students' strengths and weaknesses for the purpose of improving the learning process. To sum up, the three teachers have demonstrated the implementation of assessment according to the *Kurikulum Merdeka* instruction, even if not in its entirety.

Treating Students with Different Abilities

In regards to this activity, the three teachers stated that the students' abilities in their classroom were diverse or heterogenous. Thus, teaching or learning process that involved treating students with different levels of abilities or called as 'differentiated/individualized learning' in the *Kurikulum Merdeka* was observed. T4 attempted to teach the students using middle-level of materials after identifying the students' ability levels at the beginning, T1 and T6 attempted to do groupings based on the students' level and progress during the learning process, and T1 made students have final product of the chapter based on their preferences and abilities (product differentiation). However, the use of different level of materials or different activities for different level of students were not observed.

Challenges and Institutional Support in Implementing Assessment according to the Kurikulum Merdeka

Challenges According to Teachers

From the teachers' point of view, there were several challenges that they faced for implementing the ideal concept of assessment according to the *Kurikulum Merdeka*. First, the diverse abilities and

characters of the students made teachers work extra to provide the best individualized learning experience for their students. Second, time constraints meant that some learning plans could not be implemented optimally due to activities outside the classroom, such as the *Pancasila* student profile strengthening project (P5). Third, limited internal and external resources to study assessment instructions based on the *Kurikulum Merdeka*. To overcome these challenges, teachers expressed their eagerness to learn and practice more on the *Kurikulum Merdeka* framework by joining seminar or trainings, collaborating with other colleagues, exploring the development of technologies that relate to the students in supporting the assessment practices, as well as learning how to manage students with different abilities through class management.

Challenges according to Principals

After conducting the interview with the principals in their respective schools, it was found that they had several concerns regarding the implementation of the *Kurikulum Merdeka*. These include teachers' readiness, teachers' proficiencies, teachers' commitment, and resistance in implementing the curriculum.

The teachers' readiness in implementing the curriculum was deemed adequate in general by the three principals. However, there were still teachers who needed more guidance and assistance on the practices. This also reflected on the teachers' proficiencies and commitment in the real-classroom practices. P6 stated that almost 80% of the teachers were proficient and committed in the implementation of the curriculum, while P1 and P4 categorized the teachers into two categories, namely the fast-runner and the slow-runner. The fast-runner included teachers who had been actively involved in the curriculum implementation through trainings, seminars, and learning

community, thus committed to implement the curriculum. Meanwhile, the slow-runner generally involved senior teachers and those with low motivation to improve their competencies and resistant to the new curriculum. Lastly, the principals on the challenges that hinder the ideal implementation of the curriculum, mentioned several factors. Some of them were time constraints, teachers' motivation and mindset, students' conditions after Covid-19, and the overwhelming administrative task.

Teachers and schools are expected to provide learning experiences for students according to their individual characteristics and backgrounds. This is considered a difficult task for teachers and schools to realize. Meanwhile, the least significant factors mentioned by the respondents were teacher readiness and commitment (Arifa, 2022), teacher proficiency (Barnes, Fives, & Dacey, 2015; Narathakoon et al., 2019; Nguyen & Mai, 2015), teacher resistance to change (Aytaç, 2023), and teacher motivation (Xu & Brown, 2016; Wu, Zhang, & Liu, 2021). In conclusion, it can be inferred that the factors affecting the implementation of assessment according to the *Kurikulum Merdeka* mostly came from outside of the teachers themselves or called as 'external factors' which include students (diverse students) and educational policy (curriculum structure and workload) as reported by previous studies (Narathakoon et al., 2019; Wiese & Nortvedt, 2023).

Institutional Supports

The attempts to overcome these challenges explained by the principals and teachers were Teachers and principals revealed how they attempt to overcome the factors as mentioned in the previous sub-section conducted by the school. There are six attempts that schools provide as a support for the successful implementation of the curriculum, namely (a) professional development program; (b)

optimizing school community of practice; (c) school assessment strategy; (d) communication and collaboration; (e) building reflection culture; and (f) increasing teacher and student motivation. The most mentioned support was professional development program. This shows how important development program for teachers to prepare and support them in conducting the assessment and teaching process in general (Azis, 2015; Battistone, Buckmiller, & Peters, 2019; Brown, 2014; Chew & Lee, 2013; Ha, Nam, & Hai Tran, 2021; Kunnan, Mathew, & Zhang, 2021; Monteiro, Mata, & Santos, 2021). The attempts for optimizing school community of practice for internalizing the curriculum (Ryan et al., 2021), encouraging communication and collaboration (Berry et al., 2019; Luthfiyyah, Basyari, and Dwiniasih, 2020), as well as increasing both teacher and student motivation also become the provided supports by school leaders. Reportedly, one principal also suggested the need for building reflection culture for both teachers and students (Lan & Lam, 2020; Wiese & Nortvedt, 2023).

CONCLUSION

Four components of TALiP investigated in this study include: teacher conceptions of assessment, knowledge base, practices of assessment; and contextual factors. First, teacher had mixed conceptions of assessment in general and assessment according to the *Kurikulum Merdeka*. Second, the teachers' knowledge base of assessment varied but generally belonged to 'medium to fair' category. Third, the teacher self-perceived practices of assessment had varied results, but the practices of assessment according to the *Kurikulum Merdeka* showcased that they had implemented various aspects of assessment as the curriculum instructed, except for alternative assessments. Fourth,

teachers hold mixed perception on the assessment paradigm according to the *Kurikulum Merdeka*, but not negatively. Fifth, the contextual factors comprise the challenges and supports of the assessment implementation. The factors affecting the implementation of assessment according to the *Kurikulum Merdeka* mostly came from outside of the teachers themselves or can be called as 'external factors' which include students (diverse students) and educational policy. Principals also added teachers' resistance and low motivation. To overcome the challenges, professional development program, optimizing school community of practice, school assessment strategy, communication and collaboration, building reflection culture, and increasing teacher and student motivation. To conclude, the conceptions, knowledge based, and practices of assessment of EFL senior high school teachers in Temanggung Regency are positively related to each other and influenced by contextual factors, such as years of experience, educational background, and training programs. This study is expected to provide additional insights for all educational stakeholders on the implementation of the *Kurikulum Merdeka*, especially regarding assessment, as well as building awareness on the challenges to overcome for the successful implementation. However, due to the limitation of the study, it is hoped that further research under the same field can be conducted in a larger context.

REFERENCES

- Alsaadi, H. M. A. (2021). Dynamic assessment in language learning; an overview and the impact of using social media. *English Language Teaching*, 14(8), 73-82. <https://doi.org/10.5539/elt.v14n8p73>
- Angraeni, A. & Yusuf, F. N. (2022). Psychologically ready? Cases of EFL

- elementary school teachers' readiness to teach English in Indonesia. *Mimbar Sekolah Dasar*, 9(2), 352-366. <https://doi.org/10.53400/mimbar-sd.v9i2.49815>
- Anugrahwati, Y. & Agustien, H. I. R. (2015). The integration of second core competence (KI 2) of curriculum 2013 in English classes. *Journal of English Language Teaching*, 4(1), 1-8. <http://journal.unnes.ac.id/sju/index.php/elt>
- Astuti, P., Kasprabowo, T., Anam, Z., & Saraswati, G. P. (2023). Staying agentic in times of crisis: A literature review for a narrative inquiry on female EFL teachers. *Indonesian TESOL Journal*, 5(2), 148-168. <https://doi.org/10.24256/itj.v5i2.3907>
- Azis, A. (2015). Conceptions and practices of assessment: A case of teachers representing improvement conception. *TEFLIN Journal - A Publication on the Teaching and Learning of English*, 26(2), 129-154. <https://doi.org/10.15639/teflinjournal.v26i2/129-154>
- Baidoo-Anu, D., Rasooli, A., DeLuca, C., & Cheng, L. (2023). Conceptions of classroom assessment and approaches to grading: teachers' and students' perspectives. *Education Inquiry*, 1-29. <https://doi.org/10.1080/20004508.2023.2244136>
- Barnes, N., Fives, H., & Dacey, C. M. (2015). Teachers' beliefs about assessment. In H. Fives, & M. G. Gill (Eds.), *International handbook of research on teacher beliefs* (pp.284-300). Routledge.
- Battistone, W., Buckmiller, T., & Peters, R. (2019). Assessing assessment literacy: Are new teachers prepared to assume jobs in school districts engaging in grading and assessment reform efforts? *Studies in Educational Evaluation* 62(2019), 10-17. <https://doi.org/10.1016/j.stueduc.2019.04.009>
- Brown G.T.L., Gebril A., and Michaelides M.P. (2019). Teachers' conceptions of assessment: A global phenomenon or a global localism. *Frontiers in Education*, 4(16), 1-13. <https://doi.org/10.3389/educ.2019.00016>
- Brown, G. T. L., Hui, S. K. F., Yu, F. W. M., & Kennedy, K. J. (2011). Teachers' conceptions of assessment in Chinese contexts: A tripartite model of accountability, improvement, and irrelevance. *International Journal of Educational Research*, 50(5-6), 307-320. <https://doi.org/10.1016/j.ijer.2011.10.003>
- Brown, G.T.L., Kennedy, K.J., Fok, P.K., Chan, J.K.S., and Yu, W.M. (2009). Assessment for student improvement: understanding Hong Kong teachers' conceptions and practices of assessment. *Assessment in Education: Principles, Policy & Practice*, 16(3), 347-363. <https://doi.org/10.1080/09695940903319737>
- Brown, G. T. L., Lake, R., & Matters, G. (2011). Queensland teachers' conceptions of assessment: The impact of policy priorities on teacher attitudes. *Teaching and Teacher Education*, 27, 210-220. <https://doi.org/10.1080/02607470802587152>
- Calveric, S. (2010). *Elementary teachers' assessment beliefs and practices*. A Dissertation. Graduate School. Virginia Commonwealth University.

- Cheng, L. & Sun, Y. (2015). Teachers' Grading Decision Making: Multiple Influencing Factors and Methods. *Language Assessment Quarterly*, 12(2), 213-233. <https://doi.org/10.1080/15434303.2015.1010726>
- Chew, A. & Lee, I. (2013). *Teachers' Beliefs and Practices of Classroom Assessment in Republic Polytechnic, Singapore*. 39th Annual Conference of the International Association of Educational Assessment (IAEA) on "Educational Assessment 2.0: Technology in Educational Assessment", Tel Aviv, Israel, 20 – 25 October 2013.
- Coombs, A., Rickey, N., DeLuca, C., & Liu, S. (2022). Chinese teachers' approaches to classroom assessment. *Educational Research for Policy and Practice*, 21, 1–18. <https://doi.org/10.1007/s10671-020-09289-z>
- Daniels, L. M., & Poth, C. A. (2017). Relationships between pre-service teachers' conceptions of assessment, approaches to instruction, and assessment: an achievement goal theory perspective. *Educational Psychology*, 37(7), 835–853. <https://doi.org/10.1080/01443410.2017.1293800>
- Dashti, S. (2019). *EFL Teachers' Beliefs and Practices about Classroom Assessment: A Multiple Case Study in the Context of Kuwait*. A Dissertation. Faculty of Graduate Studies. York University.
- Elshawa, Niveen. (2016). *Universiti Putra Malaysia English language instructors' assessment beliefs and practices in the Malaysian tertiary context*. A Thesis. Faculty of Modern Languages and Communication, Universiti Putra Malaysia. <http://dx.doi.org/10.13140/RG.2.2.18241.79202>
- Fang, L. & Yu, G. (2022). Language assessment literacy of EFL teachers in East Asia: From teachers' conceptions of language assessment to their practices in classrooms. In: Lee, W.O., Brown, P., Goodwin, A.L., Green, A. (eds). *International Handbook on Education Development in Asia-Pacific*. Springer. https://doi.org/10.1007/978-981-16-2327-1_31-1
- Farhady, H. & Tavassoli, K. (2021). EFL teachers' perceptions and practices of their language assessment knowledge. *Language Testing in Asia*, 11(17), 1-19. <https://doi.org/10.1186/s40468-021-00135-4>
- Gurmesa, D. B., Birbirs, D. T., Hussein, J. W., & Tsegaye, A. G. (2022). Ethiopian secondary school EFL teachers' classroom assessment conceptions and practices from an activity theory perspectives. *East African Journal of Education Studies*, 5(1), 105-116. <https://doi.org/10.37284/eajes.5.1.574>
- Ha, X., Nam, T., & Hai Tran, N. (2021). Teachers' beliefs and practices regarding assessment in English as a foreign language classroom in Vietnam. *The Qualitative Report*, 26(11), 3457-3475. <https://doi.org/10.46743/2160-3715/2021.5063>
- Hutami, D. T., & Putro, N. H. (2023). Investigating teachers' language assessment literacy in the implementation of the Merdeka Belajar (freedom of learning) curriculum. *VELES: Voices of English*

- Language Education Society*, 7(1), 44-56.
<http://dx.doi.org/10.29408/veles.v7i1.7624>
- Jannati, S. (2015). ELT Teachers' Language assessment literacy: Perceptions and practices. *The International Journal of Research in Teacher Education*, 6(2) (2015) 26-37.
<http://www.eab.org.trhttp://ijrte.eab.org.tr>
- Kunnath, J. P. (2017). Teacher grading decisions: Influences, rationale, and practices. *American Secondary Education*, 45(3), 68-88.
<http://www.jstor.org/stable/45147907>
- Mamad, A., & Vigh, T. (2023). Moroccan EFL public university instructors' perceptions and self-reported writing practices. *Heliyon*, 9(2023), 1-12.
<https://doi.org/10.1016/j.heliyon.2023.e15626>
- Moges, B. (2018). The implementations and challenges of assessment practices for students' learning in public selected universities, Ethiopia. *Universal Journal of Educational Research*, 6(12), 2789-2806.
<https://doi.org/10.13189/ujer.2018.061213>
- Mohammadkhah, E., Kiany, G.R., Tajeddin, Z., & ShayesteFar, P. (2022). EFL teachers' assessment literacy: A contextualized measure of assessment theories and skills. *Language Teaching Research Quarterly*, 29, 102-119.
<https://doi.org/10.32038/ltrq.2022.29.07>
- Monteiro, V., Mata, L., and Santos, N.N. (2021). Assessment conceptions and practices: Perspectives of primary school teachers and students. *Frontiers in Education*, 6, 631185.
<https://doi.org/10.3389/feduc.2021.631185>
- Muianga, F. (2023). English language teachers' conceptions of assessment. *Frontiers in Education*, 7:972005, 1-15.
<https://doi.org/10.3389/feduc.2022.972005>
- Muliyah, P. & Aminatun, D. (2020). Teaching English for specific purposes in vocational high school: Teachers' beliefs and practices. *Journal of English Teaching*, 6(2), 122-133.
<https://doi.org/10.33541/jet.v6i2.1756>
- Narathakoon, A., Sapsirin, S., & Subphadoongchone, P. (2020). Beliefs and classroom assessment practices of English teachers in primary schools in Thailand. *International Journal of Instruction*, 13(3), 137-156.
<https://doi.org/10.29333/iji.2020.13310a>
- Pantiwati, Y., Chamisijatin, L., Zaenab, S., & Aldya, R. F. (2023). Characteristics of learning assessment towards implementation of Merdeka learning curriculum. *Jurnal Penelitian Dan Pengkajian Ilmu Pendidikan: E-Saintika*, 7(1), 115-128.
<https://doi.org/10.36312/esaintika.v7i1.1125>
- Roslan, R., Nishio, Y., & Jawawi, R. (2022). Analyzing English language teacher candidates' assessment literacy: a case of Bruneian and Japanese universities. *System*, 111(2022), 1-9.
<https://doi.org/10.1016/j.system.2022.102934>
- Sun, H. & Zhang, J. (2022). Assessment literacy of college EFL teachers in China: Status quo and mediating factors. *Studies in Educational Evaluation* 74(2022), 1-10.

-
- <https://doi.org/10.1016/j.stueduc.2022.101157>
- Takele, M. & Melese, W. (2022). Primary school teachers' conceptions and practices of assessment and their relationships. *Cogent Education*, 9(1), 2090185. <https://doi.org/10.1080/2331186X.2022.2090185>
- Unal, A. & Unal, Z. (2019). An examination of K-12 teachers' assessment beliefs and practices in relation to years of teaching experience. *Georgia Educational Researcher*, 16(1), 1-21. <https://digitalcommons.georgiasouthern.edu/gerjournal/vol16/iss1/2>
- Wicking, P. (2017). The assessment beliefs and practices of English teachers in Japanese universities. *JLTA Journal*, 20, 76–89. https://doi.org/10.20622/jltajournal.20.0_76
- Widiastuti, I. A. M. S., Mukminatien, N., Prayogo, J. A., & Irawati, E. (2020). Dissonances between teachers' beliefs and practices of formative assessment in EFL classes. *International Journal of Instruction*, 13(1), 71-84. <https://doi.org/10.29333/iji.2020.1315a>
- Widodo, H. P. (2016). Language policy in practice: Reframing the English language curriculum in the Indonesian secondary education sector. In R. Kirkpatrick (Ed.), *English education policy in Asia* (pp. 127–151). Springer.
- Wu, X., Zhang, L. J., & Liu, Q. (2021). Using assessment for learning: Multi-case studies of three Chinese university English as a foreign language (EFL) teachers engaging students in learning and assessment. *Frontiers in Psychology*, 12, 1-15. <https://doi.org/10.3389/fpsyg.2021.725132>
- Xu, Y., & Brown, G. T. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education*, 58, 149–162. <https://doi.org/10.1016/j.tate.2016.05.010>
- Yan, X. & Fan, J. (2020). “Am I qualified to be a language tester?”: Understanding the development of language assessment literacy across three stakeholder groups. *Language Testing*, 38. <https://doi.org/10.1177/0265532220929924>
- Yin, R. K. (2018). *Case Study Research and Applications: Design and Methods* (6th ed.). Sage.