

**INVESTIGATING SCIENCE TEACHER
ASSESSMENT LITERACY - *UNDERSTANDING,
CLASSROOM ASSESSMENT PRACTICES,
CHALLENGES AND OPPORTUNITIES*: A CASE
STUDY IN A MADRASAH SCHOOL IN LOMBOK**

A Thesis

**Submitted to the Master's Study Program of Education at the Faculty of
Education in partial fulfillment of the requirements for the degree of**

Master of Arts (M.A.)



by:

MAHURNI

04212210006

UNIVERSITAS ISLAM INTERNASIONAL INDONESIA

DEPOK

2024

**INVESTIGATING SCIENCE TEACHER
ASSESSMENT LITERACY - UNDERSTANDING,
CLASSROOM ASSESSMENT PRACTICES,
CHALLENGES AND OPPORTUNITIES: A CASE
STUDY IN A MADRASAH SCHOOL IN LOMBOK**

A Thesis

**Submitted to the Master's Study Program of Education at the Faculty of
Education in partial fulfillment of the requirements for the degree of**

Master of Arts (M.A.)



by:

MAHURNI

04212210006

UNIVERSITAS ISLAM INTERNASIONAL INDONESIA

DEPOK

2024

ABSTRACT

Mahurni

04212210006

MA in Education

Universitas Islam Internasional Indonesia

Corresponding email: mahurni@uiii.ac.id

Over the past two decades, there has been a mounting concern about the comparison of scholastic achievement of students worldwide, including Indonesia. OECD has disclosed PISA assessment results showing a global decline in student performance in the domains of reading, mathematics, and science. In an educational index, 21 provinces in Indonesia, including NTB is the position of under 50% of educational attainment index, even worse in some districts if looking at the results of ANBK and UTBK, a kind of national public examination, for some latest years. These low levels of performance have been attributed to various factors: teacher-student relationships, quality teachers, the socioeconomic status of students, the use of information and communication technology, and the implementation of metacognitive methods. Quality teachers are really matters. This is in line with syntheses of 500,000 research articles asserting that one of many factors contributing to the success of students' performance is teachers' quality in elements of teaching and assessment competency. Regarding the teacher's assessment competency, there are limited numbers of study dealing with this concern in Indonesia, particularly in science teachers. Thus, this study wants to investigate Madrasah School science teachers' assessment literacy. This research employed a qualitative approach using descriptive case study design in one of private madrasah schools in West Lombok, NTB. It involves three teacher participants from three different subjects to get their conception, assessment practice, factor of the designing and implementing assessment, along with their constraints and opportunity. Observation along with documentaries and interviews were used to collect the data, and the analysis of the data using inductive and deductive thematic analyses. Findings depict that the teachers have (1) sufficient conception regarding the assessment literacy – level two and have potential to develop further; (2) some considerations in designing and implementing the formative assessments; (3) two to three types of assessment used in their classroom teaching based on the assessment principles and approaches; and (4) challenges on students' readiness and learning sources. Moreover, the finding revealed that the school has provided a regular, annual training to improve teacher assessment literacy, followed by the government training in the form on an online platform.

Keywords: Science teachers, understanding, assessment practice, challenge, and opportunity.

ACKNOWLEDGMENT

Bismillahirrahmanirrohin

The successfulness of this thesis writing completion would not be possible without Allah (the God's) assistance along with all contributing supports from any individuals and institutions.

First and foremost, I would like to sincerely thank my two outstanding supervisors, *Bapak Bambang Sumintono, Ph.D.*, and *Ibu Tati Lathipatud Durriyah, Ph.D.*, for their invaluable guidance and support. Throughout the consultation period, their assistance, knowledge, and steadfast encouragement played a crucial role. The thesis writing project quality is much improved by the provision of constructive and insightful criticism, as well as the desire to invest their time. I have acquired invaluable research skills, particularly in the areas of research methodology and data reporting.

Moreover, I am profoundly grateful to the educational institution that granted me the opportunity to conduct the research, as well as to the participants who graciously dedicated their precious time and ideas. Also, I am deeply grateful to my academic advisor – Dr. Lukman Nul Hakim, M.Sc. who consistently reminded me to study effectively, and finish my study. Furthermore, I thank to the Ministry of Religious Affairs and Universitas Islam International Indonesia (UIII) for providing me with the opportunity to pursue my second master's degree through a fully funded scholarship that covers a two-year study duration.

Finally, I would want to convey my profound gratitude to my beloved wife - Sanita Permatasari and four children (*Shofiyah Azzahra, Utsman Khalid, Rumaysho Khumaira, and Aisyah Nur Amaliya*) who demonstrate remarkable patience and continue to offer their prayers for me while being physically apart. Further, I would like to express my gratitude to my *Inak* - mother, as well as my colleagues at the Institute of Education Development NTB (LPPNTB) and the Regional Research and Innovation Agency (BRIDA NTB), for their invaluable support.

Despite already having obtained my first master's degree in Australia, this research project at UIII has proven to be an invaluable experience. Without using any words, I express my gratitude to Allah and extend my thanks to everyone.

Alhamdulillahirobbilalamiin

TABLE OF CONTENTS

TITLE PAGE.....	i
STATEMENT OF AUTHENTICITY.....	ii
ANTI-PLAGIARISM STATEMENT.....	iii
THESIS ATTESTATION.....	iv
THESIS DEFENSE APPROVAL.....	v
ABSTRACT.....	vi
ACKNOWLEDGMENT.....	vii
TABLE OF CONTENTS.....	viii
ABBREVIATION.....	x
CHAPTER I.....	1
INTRODUCTION.....	1
1.1. Research Background.....	1
1.2. Research Questions.....	10
1.3. Research Objective.....	10
1.4. Significance of the Study.....	11
CHAPTER II.....	12
LITERATURE REVIEW.....	12
2.1. Theory of Learnings.....	12
2.2. Theory of Assessment Literacy.....	20
2.3. Standards and Principles of Assessment Literacy.....	23
2.5. Levels of Assessment Literacy.....	26
2.6. Procedure and Methods of Assigning Assessment.....	29
2.7. Assessment Approaches.....	33
2.8. Previous Study.....	36
CHAPTER III.....	39
RESEARCH METHODOLOGY.....	39
3.1. Research Approach.....	39
3.2. Research Design: Descriptive Case Study.....	40
3.3. Research Context.....	40
3.4. Research Subjects.....	42
3.6. Data Collection Methods.....	44
3.7. Data Analysis Methods.....	48
3.8. Ethical Issue of the Research.....	49

3.9. Limitation of the Research	50
CHAPTER IV	51
FINDING AND DISCUSSION.....	51
4.1. Findings	51
4.1.1. Teachers’ Understanding on the Assessment Literacy	51
4.1.2. Teachers’ Classroom Assessment Practice.....	57
4.1.3. Factors in Designing and Implementing Assessments	61
4.1.4. Teachers’ Constraints and Opportunities	63
4.2. Discussion.....	67
4.2.1. Discussion on RQ1: Teachers’ Understanding on the Assessment.....	67
4.2.2. Discussion on RQ2: Teachers’ Classroom Assessment Practice	71
4.2.3. Discussion on RQ3: Designing and Implementing Assessments.....	72
4.2.4. Discussion on RQ4: Teachers’ Challenges and Opportunities.....	74
CHAPTER V	77
CONCLUSION AND RECOMMENDATION	77
5.1. Conclusion.....	77
5.2. Recommendation	82
REFERENCES	84
APPENDICES	94
Appendix 1	94
Classroom Observation Sheet.....	94
Appendix 2	95
Semi-Structure Interview Questions.....	95
Appendix 3	99
Interview Transcripts.....	99
Appendix 4	108
Data Coding Analysis.....	108
Appendix 5	109
Documents of Classroom Observation and Interviews	109

ABBREVIATION

PISA	: <i>The Program for International Student Assessment</i>
OECD	: <i>Organization for Economic Cooperation and Development</i>
AN	: <i>Asesmen Nasional (National Assessment)</i>
MoEC	: <i>Ministry of Education and Culture</i>
MoERCT	: <i>Ministry of Education, Culture, Research and Technology</i>
UN	: <i>Ujian Nasional (National Examination)</i>
UTBK	: <i>Computer-Based Written Examination</i>
LPMP	: <i>Lembaga Penjamin Mutu Pendidikan</i>
NTB	: <i>Nusa Tenggara Barat (West Nusa Tenggara)</i>
SMA	: <i>Sekolah Menengah Atas (Senior High School)</i>
MA	: <i>Madrasah Aliyah (Islamic Senior High School)</i>
AL	: <i>Assessment Literacy</i>
TAL	: <i>Teacher Assessment Literacy</i>
CAL	: <i>Classroom Assessment Literacy</i>

CHAPTER I

INTRODUCTION

Teaching effectiveness is contingent upon assessment effectiveness. However, research indicates that many teachers need more vital assessment literacy to accurately assess their students' learning. This chapter describes the present status of student learning performance outcomes in both international and national assessment results, along with the government's options to improve teacher assessment literacy. Following these, this chapter examines the scope of the problem and its potential repercussions. It also discusses research questions, objectives, and the significance of this research.

1.1. Research Background

Over the past decade, there has been a mounting concern about the academic achievement of students worldwide. The Program for International Student Assessment (PISA), conducted by the Organization for Economic Co-operation and Development (OECD), has disclosed a global decline in student performance in the domains of reading, mathematics, and science (OECD, 2018, 2023). The decline has been attributed to various factors such as the teacher-student relationships, the socioeconomic status of students, the use of information and communication technology, and the implementation of metacognitive methods (Vazquez-Lopez & Huerta-Manzanilla, 2021). This decline is not limited to any specific area or country; instead, it is a widespread phenomenon that has been recorded in several countries. Within the Indonesian context, a study revealed a decline in the overall performance of Indonesian pupils on the PISA assessment (Ningsih et al., 2021; OECD, 2023). This pertained to the academic achievement of students in Indonesia as measured by the PISA assessments.

More specifically, the OECD (2023) report found that the average score for reading ability among Indonesian students was 359, which is 117 points lower than the average score of 476 for students around the world. In comparison to the previous edition, this score represents a decrease of 12 points, which is also a significant decline. Furthermore, the fall in marks in this particular area has been especially noticeable in the most recent five editions of the test. In the meantime, Indonesia's average score in mathematical proficiency, which is the primary area of interest in PISA 2022, dropped by 13 points to 366, compared to the score of 379 in the previous edition of the test. According to the OECD (2023), this result represents a departure of 106 points from the average score for the entire world. In addition to this, the average score for academic competence in the subject of science dropped by thirteen points. When it came to this particular aspect of the PISA 2022 examination assessment, Indonesia's performance was characterized by a mean score of 383, which was 102 points lower than the average score for the entire world.

Apart from PISA test results, the Indonesian students' learning performance can also be traced from the results of an *Asesmen Nasional*/National Assessment (*henceforth*: AN) test program. The AN program has been administered by central government since 2021. According to MoEC (n.d), the AN, a student learning outcomes assessment program conducted by the Ministry of Education, Culture, Research and Technology of Indonesia (MoECRT), aims to assess the caliber competency of primary to secondary education at a national level and produce accurate data to improve the overall educational benchmarks. MoEC (2023) reports that the 2022 AN's findings indicate that Indonesian students, on average, have a satisfactory level of literacy, ranging from 40% (low

performing regions) to 70% (highest performing regions). These findings show that Indonesia education quality needs significant improvements.

If we refer to an educational index throughout 38 provinces across Indonesia, Aldini (2020) asserted that the provinces with above 70% of the education attainment index are Yogyakarta, Bangka Belitung, and Jakarta, while those with the educational index ranging from 50%-70% are 13 provinces including Gorontalo, Kepulauan Riau, South and West Sumatra, Central, East and West Java, South Kalimantan, and Bali. However, the rest 21 provinces, including the province of West Nusa Tenggara (henceforth: NTB), are still in the position of under 50% of the education attainment index (Aldini, 2020). Dealing with NTB students performances since the last test years, the aforementioned education index can be correlated with the results of the previous *Ujian Nasional/UN* (National Examination), the first AN, and two latest *Ujian Tulis Berbasis Komputer/UTBK* (Computer-Based Written Examination) which showing not so different result.

The National Examination result, which dates back to 2019, revealed that none of the NTB schools deserved to be among the 100 best-performing schools (Fahrunnisa, 2021). Most NTB students attained a mean score of 55 to 70 (MoEC, 2019). In more detail, the Center for Educational Assessment (MoEC, 2019) reported that the total academic achievement of senior high school students in the NTB province showed a minor decrease in the 2019 UN, as compared to the previous year of 2018. This phenomenon was noted in the *Ujian Nasional Berbasis Komputer/UNBK* (Computer-Based National Examination), which included a wide range of courses including both science and non-science subjects. The disparity between high- and low-performing schools has increased, highlighting the necessity for focused interventions to overcome

these discrepancies. On average, public schools had superior performance compared to private Islamic or non-Islamic schools in terms of scores.

In addition, the initial AN result from NTB in 2021 revealed unsatisfactory outcomes, particularly in the literacy and numeracy test scores of school students (Ron, 2022). According to the NTB Institute of Education Assurance (Lembaga Penjamin Mutu Pendidikan NTB/LPMPNTB) official (2022) report, the result of the 2021 national assessment illustrated that NTB's education report is quite poor, accounting for a mean score of 50, ten points below the national average score of 60, from the maximum score of 100. Furthermore, as stated by the Head of the High School Division of the NTB Provincial Education Office, the current state of literacy and numeracy is not promising since both literacy and numeracy rates remain lower than the national average (Ron, 2022). Over half, or more than 50%, of NTB senior high school students possess reading levels that fall below the expected requirement.

Furthermore, the two latest *Ujian Tulis Berbasis Komputer/UTBK* (Computer-Based Written Examination), a test which has quite similar to National Assessment tests that is used for public university entrance selection, illustrated that only a few numbers of schools in NTB achieved a position within the top 1000 Senior High Schools (*henceforth*: SMA) in Indonesia in terms of performance (Dikbudntb, n.d). In more detail, the 2021 students' UTBK result depicted that only six NTB schools were classified into 1000 best-performance schools. This includes Madrasah Aliyah Negeri/MAN (State Islamic Senior High School) 1 Selong, state SMA 1 Selong, SMA 1 Mataram, SMA 5 Mataram, SMA 1 Praya, and SMA 1 Sumbawa. Regrettably, the results of the 2022 students' UTBK results were even worse or more unfavorable. There are only two schools (state SMA 1 Mataram and SMA 1 Sumbawa) in NTB achieving a ranking among the top

1000 performing schools across the country. Paradoxically no students from either state schools or private school-based Islamic boarding schools in the regencies of West and North Lombok, West Sumbawa, Dompu, and Bima met the aforementioned UTBK criteria.

Some previous studies have actually mentioned factors affecting students' test performances in various assessments. According to Akay and Oskonbaeva (2024), some variables like economic and socio-economic status significantly contribute to the students' learning outcome in Asian countries. Additionally, factors such as metacognition, learning time, and the index of economic, social, and cultural status (ESCS) have been identified as key predictors of mathematics achievement in Indonesia (Anggraheni et al., 2022). Furthermore, Santi et al. (2023) mentioned factors influencing Indonesian students' PISA scores are class taken, parental education, home facilities, student discipline, teacher feedback, kindergarten entry age, and grade retention.

There are still very limited numbers of studies in Indonesia focusing on the relationship of teachers' quality on the aforementioned assessments. However, regarding students' learning performance, Hattie (2003), who synthesized 500,000 journal research articles, found that quality teacher also plays a significant role in improving students' performance, which is accounting for 30%. I would argue that the 30% is a reasonably substantial quantity that can have a significant impact on student achievement. Influences of parents, school headmaster, and pairs on students' success are ranging about 5 to 10 percent each. Tapung (2024) also found that there is significant correlation between teacher quality and students' learning process which impact towards students' learning activeness. Thus, Hattie (2003) proposes that enhancing the teachers' quality can have a substantial impact on student achievement.

To increase teachers' competency across Indonesia, including NTB, the central government has long been trying some alternatives. One of which was through support for the conduct of teacher certification program which aims at ensuring the teachers are qualified in their fields and are expected to provide better-quality services which subsequently affect their student's learning performance. However, it is unfortunate that research findings asserted that the certification program possibly failed and probably wasted a lot of state funds (Sulaksono *et al.*, 2009; World Bank, 2012; Kusumawardhani, 2017; Ree et al., 2018;). Sulaksono *et al.* (2009) opined that the program failed to ensure the teachers' teaching capabilities in their field, and its failure was linked with deceptions in process of the portfolio preparation wherein almost the entire teachers attending the certification program in 2007 would exactly be given a pass certificate, as long as they completed the requirements and attended remedial trainings. Furthermore, Kusumawardhani (2017) and Ree et al. (2018) in their study discovered that providing double payment to certified in-service teachers does not result in skill upgrading to enhance teachers' productivity or quality, nor does it lead to any improvement in students' learning test results. If more teachers are increasingly qualified, this will ultimately have an impact on the student learning process, and they could significantly contribute to students' achievement (Darling-Hammond, 2000).

Another alternative in ensuring teachers will perform effectively is through a-one-year *Program Profesi Guru/PPG* (professional teacher training program) for both pre-service and in-service teachers. This program aims to adequately equip teachers with the necessary skills and knowledge required to meet national standards and effectively carry out teaching responsibilities. Through this program, teachers are highly expected to enhance their competencies in both assessment practices and instructional delivery or

teaching professionalism (MoEC, n.d.). However, it is unfortunate that PPG curriculum does not place significant focus on the testing and evaluation abilities of teachers (Sulistiyowati et al., 2021). Meanwhile, the Ministry of Education (n.d.) asserted that teachers are highly expected to enhance their competencies in both assessment practices and instructional delivery or teaching professionalism.

However, it is unfortunate that the government in the PPG curriculum allocates only three credits for teaching and assessment course or about 1.5 credits for teaching and the rest for assessment (MoEC, n.d.). Indeed, both competences are pivotal since half of the learning is about assessment (Girgla et al., 2021). Also, Anggraena et al. (2022) asserted that pedagogical skills and assessment literacy are inseparable components that form a cohesive entity in the education system. The assessment literacy is defined as the comprehensive knowledge, skills, and attitudes that teachers must possess to effectively assess and support student learning (Pastore, 2023). According to studies of Davison (2012) and, Lian and Yew (2016), low level of teacher assessment literacy not only leads to insufficient preparation in assessing the students, but it also can create issue inability in assessing and evaluating the student accurately. However, if the teachers are assessment literate, they can describe the target of achievements of the students in any specific context and terms (Stiggins, 1991; Hattie & Brown, 2010).

Given that having expertise in the assessment literacy is an essential part of the current national education curriculum, it is undoubtedly that the teachers' conceptions and implementation of solid assessment techniques along with the properly classroom assessment practice play a crucial role in facilitating successful appropriate teaching and learning (DeLuca & Klinger, 2010). However, it is unfortunate that it is likely the current PPG program does not offer sufficient credits for assessment literacy courses (MoEC,

n.d.). Indeed, Baird and Reid (2023) asserted that assessment is an indispensable component of the teaching and learning process. The higher the level of the teacher's assessment literacy, it can make the greater teaching efficacy and commitment they will perform (Zhang & Burry-Stock, 2003). Also, Wang et al. (2022) revealed their finding which shows that teachers' level of understanding and competence in assessing students' performance had a strong and favorable correlation with their effectiveness and dedication in teaching. In other words, assessment literacy is crucial for successful teaching profession, as it enables teachers to proficiently create, execute, and analyze evaluations from the students' test results (Michigan Assessment Consortium, 2015).

Furthermore, it is noted that an effective assessment is a cornerstone of quality education which could provide valuable information about student learning, inform instructional decisions, and motivate students (DeLuca, et al., 2016). DeLuca et al. (2015) and Giraldo et al. (2018) contended that classroom teaching along with quality of assessment is up to fifty percent about assessment activities which will significantly affect the students' learning outcome. This means that half of the learning process in the classroom is actually about assessment. In other words, teachers throughout the learning will daily monitor the student's learning performance. In consequence, it is crucial to fully understand and effectively use the principles of teaching and evaluation to successfully meet educational objectives (Anggraena et al., 2022).

There are also some other factors which influence the teachers' expertise in assessments. These include personal factors, contextual factors, views on social studies, and competency (Yan et al., 2022; Bourke et al., 2018; Alkharusi et al., 2014; Girgla et al., 2021; Paramo, 2008). The first factor is personal factors such as teacher rank, years of teaching experience, and self-efficacy have been influential factors in teacher

formative assessment procedures (Yan et al., 2022). Another factor that has a substantial impact on the formation of teachers' formative assessment practice (FAP) is contextual factors (Bourke et al., 2018). This includes the level of support provided by the school, the overall number of teachers in a school, and the specific grade level being taught. The third factor is related to the teachers' views (Alkharusi et al., 2014; Girgla et al., 2021). They argued that teachers' views towards social studies, gender, teaching experience, and educational degree significantly impact their perceived assessment procedure needs. Finally, competency plays a crucial role in explaining the differences in their practices when it comes to interpreting classroom assessments (Paramo, 2008).

Related to these concerns, one of ways to effectively evaluate students' abilities and character development within framework of the current national curriculum is to fully understand and effectively use the principles of teaching and evaluation to successfully meet educational objectives (Anggraena et al., 2022). Also, the current high school national curriculum encompasses the comprehension of crafting assignments that foster critical thinking, analyzing outcomes from evaluations, and utilizing the data to shape instructional and learning approaches (Hutami & Putro, 2023). I argue that students will be much more familiar with a variety of test difficulties either in the national assessment or in PISA tests, if their teachers are capable of creating and assigning them with such test questions.

I found many some researches dealing with English as a Foreign Language (EFL) teachers' assessment literacy within 2013 curriculum (Suherman, 2022; Mirizon, 2021; Wardani et al., 2021; Umam & Indah 2020; Zulaiha et al., 2020; Pupawati, 2019; Prasetyo, 2018; Arsyad & Sumarni, 2018). Nevertheless, I only found one Indonesian research dealing with the teacher assessment literacy dealing with training improve

teachers' competency in creating e-students' worksheet (Iswatiningsih et al., 2022). In other words, a very limited research delved into either the science teachers' assessment literacy in terms of their understanding along with their classroom practice in accordance with the assessment principles in West Lombok NTB context. Considering the essence of possessing assessment literacy both in concepts and practices, researchers argue that it is pivotal to investigate teachers' assessment literacy within this context. Thus, this study wants to investigate deeper Madrasah School science teachers' assessment literacy, understanding, classroom assessment practice, challenges, and opportunities.

1.2. Research Questions

In this study, I formulate research questions to investigate the proposed topic. The research questions are:

1. To what extent do private Madrasah Aliyah (MA) science teachers in West Lombok understand their current assessment literacy?
2. What are the current specific classroom assessment practices utilized by MA science teachers in West Lombok which are align with the assessment approaches, principles, or procedures?
3. How do MA science teachers in West Lombok design and implement classroom assessments?
4. What challenges and opportunities do MA science teachers in West Lombok face in implementing assessment practices?

1.3. Research Objective

The objectives of this research are:

1. To explore the understanding of MA science teachers regarding their current assessment literacy.

2. To identify and explore the current specific classroom assessment practices utilized by the MA science teachers which are aligned with the assessment principles or procedures.
3. To explore MA science teachers' ways of designing and implementing classroom assessments.

To identify and explore challenges and opportunities do the MA science teachers face in implementing assessment practices.

1.4. Significance of the Study

This study has the potential to provide useful insights into the field of teacher education in NTB, Indonesia. Through an examination of teachers' perceptions, classroom assessment practices, ways of designing and implementing assessment along with their challenges they might encounter, and opportunities offered, teachers can utilize forthcoming findings as reflections to improve their current assessment literacy. This will enable them to adopt more effective assessment approaches, thereby enhancing students' learning performance in any tests. Additionally, the upcoming discoveries can be used as a reference to inform the development of professional development programs designed to improve the assessment literacy of school instructors. This can provide teachers with the expertise and skills to create assessments that accurately measure student progress towards specific learning objectives.

CHAPTER II

LITERATURE REVIEW

This chapter highlights the current studies about assessment literacy for teachers. Comprehending the current state of knowledge regarding teachers' proficiency in designing, implementing, and interpreting tests is vital of important. Thus, a thorough analysis of academic literature, I intend to discern the very fundamental elements of teacher assessment literacy and to investigate its influence on student learning. Thus, I present a comprehensive examination and evaluation of the literatures in five primary domains (Shepard, 2000; Mertler, 2004; James, 2006; James 2008; Schunk, 2012; Deluca, 2016; Xu & Brown, 2016; Santrock, 2018; Girgla, et al., 2021; Popham, 2021; Anggraena et al., 2022; Pastore, 2023): a) the theory of learning and assessment which serves as the theoretical foundation for this research and is linked to theories of teaching and learning, b) standards and principles of assessment literacy which teachers should be familiar with, c) levels of assessment literacy that show which levels of teachers possess, d) procedure and methods of assigning assessment which explain the standard procedures and method teachers can practice in classroom assessment, and e) afterwards, I identify the shortcomings that exist in the current corpus of knowledge in these particular areas from some previous studies.

2.1. Theory of Learnings

Learning theory or approach developed by some educational psychology experts like Ivan Pavlov, Bandura, Skinner, or Vygotsky has shed light on our understanding on how individuals acquire knowledge and skills. This research refers to three theories of learning which teachers can use as reference as they design teaching and assessments for their students. Firstly, if teachers adopt the learning theory of behaviorism which was

popularized by Ivan Pavlov and Skinner, they view student's behaviors are formed by environment and reinforcement. Secondly, if teachers adopt the theory of constructivism, they will see that how learners themselves construct their own knowledge and skills to acquire new understanding. Meanwhile, if they adopt the third theory - a socio-cognitivism, they will view the learners acquire knowledge through observation and modelling. The following are the further explications.

2.1.1. Learning Theory of Behaviorism

Behaviorism, a learning theory developed by American psychologist John B. Watson, influenced teachers' methods of instruction throughout processes of learnings (James, 2006; Schunk, 2012). In behaviorism theory, learners create connection between what they experience which are related reinforcements and punishment. These can become determinants of students' learning and behaviors. Schunk (2012) mentioned five essential approaches in behaviorism theory consisting of connectivism, classical conditioning, contagious conditioning, operant conditioning, and instructional application. However, Santrock (2018) mentions two main approaches in this behaviorism theory: classical conditioning and operant conditioning.

Classical conditioning, which was developed by Ivan Pavlov in 1927, can be noted as the behaviorism theoretical cornerstone in which it explicates how people's environments and experiences, or stimuli, influenced learning (Schunk, 2012; Santrock, 2018). Santrock (2018) argues that to understand this theory, teachers have to understand types of stimuli (unconditioned and conditioned stimuli) and responses (unconditioned and conditioned responses). When the unconditioned stimuli can trigger a reflection (unconditioned response) along with a neutral stimulus that initially has response. If these two stimuli are repeatedly conducted, it will condition the neutral stimulus. On the other

hand, operant conditioning, also known as instrumental conditioning, is another learning approach that emphasis on behavior consequences impact changes as a result of adding or removing rewards and punishment (Santrock, 2018). Rewards have probability to increase learners' behavior, in contrast to punishment might decline the occurrence of the behaviors.

In terms of students' learnings, behaviorists held that analysis of repeated interactions (stimulus-response) with the environment was the source of all knowledge and that behaviors could be explained by this relationship. This perspective, as expressed by Shepard (2000) and Schunk (2012), saw learning as a straightforward, sequential process that was assessed by students' responses, and knowledge as a collection of facts that teachers imparted to their charges. In assessing students, teachers test students on material they have learned by memorization. But there were drawbacks to this strategy. Originally, knowledge was viewed as being at strict levels, with flawless memorization being the only criteria for being deemed knowing (Wolf et al., 1991). Secondly, learning potential and intellect were considered fixed attributes that could not be acquired via education.

Furthermore, teachers who adopt behaviorism will transform classrooms into settings where they wield complete authority, determining the dissemination of knowledge and exerting control over the speed and substance of learning (James, 2008). In other words, students were not anticipated to possess autonomous thinking skills; instead, they acquired knowledge by engaging in repetitious exercises and simply replicated the teacher's statements (James, 2008; Schunk, 2012). The intricate skills were deconstructed into minuscule stages, and pupils were categorized according to their aptitude, presupposing that all individuals acquired knowledge at the same pace. This

approach had an impact on the ways the teachers evaluate their students. Tests were traditionally regarded as distinct from the process of learning, serving just as a means to assess whether students had memorized particular information (Schunk, 2012). These assessments frequently consisted of multiple choice or fill-in-the-blank questions to assess memorization and anticipated conduct. According to Buhagiar (2007) and Schunk (2012), good grades become the only criterion for determining success. Both students and teachers are driven by extrinsic incentives and consequences, with teachers using frequent assessments to exert pressure on pupils for improved performance.

2.1.2. Learning Theory of Constructivism

Contrary to behaviorism, which emphasizes conditioned reactions, constructivism theory emphasizes the process by which learners construct their own understanding (Schunk, 2012). Schunk mentioned that if teachers adopt the constructivism, they will reject the assumption “scientific truths exist and awaits discovery and verification (p. 230).” They believe that knowledge is imposed from inside, not outside people, meaning that constructivism underlines how individual students interact with situations to acquire and refine the knowledge and skills. Thus, the key point of this view is that teachers should design learnings which encourage students to be active to develop their own knowledge. Teachers should design the learning where students are actively involved with course materials in social interactions, such as phenomena observation, data collection, hypothesis generation and testing, and working collaboratively.

In other words, students engage in the process of actively constructing meaning by establishing connections between new material and their preexisting knowledge and experiences (James, 2008; Schunk, 2012). In this theory, learning is perceived as an individualistic expedition, wherein every student assumes accountability for expanding

upon their existing knowledge. Schunk (2012) stated that constructivists assert that our learning process involves actively constructing our understanding of the world by drawing upon our pre-existing knowledge and cognitive frameworks. This process is highly dependent on our existing knowledge and our capacity to engage in metacognition, which involves reflecting on and controlling our own learning. The purpose of this is to comprehend and assimilate new information (Shepard, 2000; Schunk, 2012). Schunk (2012) mentioned that learning occurs when students actively examine and understand new experiences utilizing a variety of resources and activities, which serve as a support system for constructing knowledge.

Further, in constructivism, James (2006) stated that teachers act as bridges. They help students connect what they already know to new ideas. The aim is to turn students from beginners into experts by providing them with support and problem-solving techniques. This involves using both logical reasoning (deductive) and figuring things out from examples (inductive) (James, 2006). Instead of simply following a set plan, teachers encourage discussions, sharing of ideas, and even venturing beyond the curriculum. To do this well, teachers need to truly understand the subject, how it builds on itself, and the common mistakes students make. Additionally, in assessing students' performances in the discussions or idea sharing, teachers can ask open-ended questions and design activities that make students think critically, argue their points, and connect different ideas (Nuthall, 1997). Students, on the other hand, are expected to be actively involved, think about their learning, and be self-motivated. They take charge of their learning and adjust their approach as needed. They get to use what they've learned in new situations to solidify their understanding (Corden, 2000). A key skill for students is self-regulation, which allows them to use all available resources to solve problems and understand the

world around them. However, implementing curriculums and assessments that encourage this self-regulation can be challenging, even difficult for teacher educators.

2.1.3. Learning Theory of Socio-Cognitive

Different from the two above-mentioned theories of learning, a socio-cognitive theory developed by Bandura asserts that humans learn through observations of what is happening in their surrounding environment (Schunk, 2011; Santrock, 2018). In the context of education, teachers should understand that environments of classroom and behaviors of students are the two interconnected elements which influence each other. Bandura (1986, cited in Schunk, 2011) asserted that learning is an activity to process information where both behavioral structures and environmental events are developed to form symbolic representation as an action reference. Take for example, when students respond to a question incorrectly, the teacher might take an action to repeat the explanation where the student made incorrect answers. Another analogy is “when people fail or are punished, they know that they are doing something wrong and may try to correct the problem (Schunk, 2011, p. 134).” This means that such individuals’ behavior influence environment, nor the cognitions.

According to this theory, teachers should see and assess students as humans who acquire knowledge through direct experience (enactive learning) or observing others (vicarious learning). The enactive learning refers to acquiring knowledge through experiencing the consequences of one's actions; meanwhile, the vicarious learning involves humans acquiring knowledge by listening to or witnessing models, whether they be human or non-human. In traditional classroom settings, students acquire knowledge by seeing their teachers' explanations and demonstrations of the subjects, as discussed by Schunk (2011) and Santrok (2018). Meaningful learning in complicated skills occurs

when students can watch and then do the task. As an example of this principle, a genuine golfer observes themselves while practicing and receive constructive comments from the trainers.

Put simply, the social cognitive theory is a learning notion that is based on the process of seeing and imitating others. Nevertheless, this observational learning paradigm encompasses several crucial elements, namely attention, retention, production, and motivation (Bandura, 1986, as cited in Santrock, 2018). When it comes to the main determinant, the attention, students show a greater inclination towards learning from what they prefer or consider reliable models, rather than those who lack credibility. These could include their teachers, parents or peers who have prestige and competence. Retention refers to the ability of students to store seen actions and behaviors in their memory for later retrieval. Verbal actions or visual representations can enhance pupils' ability to remember information. During the production stage, students have an opportunity to transform their stored information into tangible action or performance, similar to how dancers require practice. Ultimately, motivation, such as providing reinforcement or incentives, can serve as a catalyst for children to exhibit specific patterns of behavior.

Another crucial aspect of this approach pertains to self-regulation. In the field of education, self-regulation refers to students' belief in their ability to control and manage their emotions, thoughts, and actions to achieve their goals (Santrock, 2018). Three processes are involved in this part: self-observation, self-judgment, and self-reaction (Schunk, 2011). Schunk contended that possessing this capability would enable pupils to achieve more personal agency. While the degree of self-regulation may differ, kids possess definite strategies and methods for achieving their goals. According to Winne

(2015, cited in Santrock 2018), self-regulated pupils possess certain traits such as goal planning, emotional management, fine-tuning, and evaluation.

2.1.4. Learning Theory in Indonesia's Education Context

Indonesia has undergone many theories of teaching and learning throughout the years. Behaviorism was the theory used prior to the latest transformation to constructivism (Anggraena et al., 2022). The constructivism learning theory as above explained refers to a philosophical approach that promotes the logical and conceptual development of students (Driscoll, 2000). The fundamental idea is the significance of interactions with the surrounding environment in student education. This theory posits that learners is capable of generating knowledge and constructing meaning through their personal experiences. According to Anggraena et al. (2022), the concept of constructivism has been being administered in the current education curriculum. In general, the administration of this learning theory in the education system is due that it has two primary elements, accommodation, and assimilation (Bada & Olusegun, 2015). Assimilation refers to the incorporation of new experiences into preexisting ones. This process results in the individual developing new viewpoints, reevaluating earlier misunderstandings, evaluating the importance of different elements, and ultimately altering their perceptions. Meanwhile, the accommodation entails modifying the surroundings and encountering novel situations to align with the current cognitive capabilities.

Individuals in this learning theory construct a distinct understanding of how the world operates through their cognitive processes (Caine & Caine, 1991, as cited in Bada & Olusegun, 2015). Accordingly, the core tenet of the constructivist theory is that learning is an interactive process in which learners actively engage. While knowledge

transfer can be compelled, comprehension cannot be coerced as it necessitates an inward acquisition of knowledge. Constructivism necessitates that teachers assume the role of facilitators, with the primary objective of assisting students in becoming active participants in learning. This involves enabling students to establish significant connections between their existing knowledge, newly acquired knowledge, and the various processes involved in the learning experience.

In the context of assessment and evaluation, I synthesize that the three above-mentioned theories offer different lenses for teachers in designing learning process and assessment methods. If I synthesized behaviorism, I would argue that this theory emphasizes the evaluation of visible and measurable learning outcomes in which teachers tend to use objective tests such as multiple-choices or true-false questions to evaluate students' conceptual and factual understanding. In the theory of constructivism, teachers will focus on the process of how students' conceptual understanding is constructed. The assessment using this theory tends to emphasize conceptual understanding, problem-solving, and interconnected ideas. Meanwhile, in socio-cognitive theory, teachers view social roles in students' learning. If teachers design assessments for students, they not only measure individuals' understanding, but also measure individual collaboration and communication with others.

2.2. Theory of Assessment Literacy

Stiggins (1995) and Mertler (2004) defined assessment literacy as the knowledge acquisition concerning the underlying principles of efficient assessment techniques. Girgla et al. (2021) contended that assessment literacy has been a historical area of focus in educational quality since it becomes the central element of learning and teaching. According to Pastore (2023), the assessment literacy entails the understanding, abilities, and attitudes that teachers need to accurately evaluate and facilitate learnings of students.

Further, Michigan Assessment Consortium (2020, p.2) defines assessment literacy as “the set of beliefs, knowledge, and practices about assessment that lead a teacher, administrator, policymaker, or students and their families to use assessment to improve student learning and achievement.” It can be inferred that an expertise in assessments is pivotal in which decisions regarding education system, educators, and students themselves are made in accordance with data collected from students’ performances.

Popham (2009) also opined that a genuinely literate teacher will have both an ability to design more appropriate assessments and an expertise in the selection of a wide variety of assessment types. However, if the educators are illiterate, the students would experience the greatest detriment. In other words, literate teachers in assessments are aware of what they evaluate, the rationale behind their evaluations, the method of evaluation, and the potential pitfalls that may arise (Girgla et al., 2021). Thus, the assessment literacy involves acquiring knowledge on creating and utilizing tests to aid student learning, ensuring fair and reliable decision-making, and offering feedback for enhancement (Pastore, 2023; Baartman, et al., 2023; Kawashima, 2022). This notion has transitioned from a narrow emphasis on standardized testing to a broader and more nuanced comprehension that takes into account the dynamic nature of learning, the context in which it occurs, and the influence of teachers' identities and practices (Shields, 2023). Accordingly, assessment literacy is somehow teachers’ compulsory expertise since the assessment must be aligned with learning objectives.

Shifting to the more specific terms definitions, Pastore (2023) highlights the terms of Teacher Assessment Literacy (TAL) as the competency consisting of knowledge and skills which should be possessed by teachers to effectively implement assessments in practices. Moreover, Xu and Brown (2016) drove this concept as teacher assessment

literacy practice (TALP) with which they defined assessment literacy as the competency which must be possessed by teachers in practical teaching. Different from the two abovementioned, Asamoah et al. (2023) specify the terms into Classroom Assessment Literacy or CAL. This tends to involve teachers' approaches to assessment within specific classroom context consisting of test designs, scorings, types in either formative or summative assessments. Overall, the definitions emphasize the essence of teachers' competence in assessment practices, with TAL focusing on foundational knowledge and skills; meanwhile, TALP and CAL delve the practical application of and approaches to assessment within the classroom setting. According to these, the terms Assessment Literacy (AL) can cover the of the above-mentioned.

Several aspects impact a teacher's proficiency in assessment. The aspects that contribute to this phenomenon encompass personal factors, contextual factors, perspectives on social studies, and competencies (Yan et al., 2022; Bourke et al., 2018; Alkharusi et al., 2014; Girgla et al., 2021; Paramo, 2008). Personal factors, including teacher rank, years of teaching experience, and self-efficacy, have been significant influences on teacher formative evaluation processes (Yan et al., 2022). Contextual factors, as shown by Bourke et al. (2018), have a significant role in shaping teachers' formative assessment practice. This encompasses the extent of assistance offered by the educational institution, the total count of teachers within the school, and the particular grade level being instructed. The third factor pertains to the perspectives of the teachers (Alkharusi et al., 2014; Girgla et al., 2021). They contended that teachers' perspectives on social studies, gender, teaching experience, and educational degrees have a substantial influence on their perceived requirements for assessment procedures. Finally,

competency which is inherent in teachers is a significant factor in understanding the variations in their approaches to processing classroom assessments (Paramo, 2008).

2.3. Standards and Principles of Assessment Literacy

a. Standards of Teacher Assessment Literacy

There are seven standards in teacher assessment literacy (Deluca et al., 2016). These include having an ability in; (1) deciding assessment purpose (know how to appropriately opt assessment), (2) conducting assessment process (know how to create, administer, score, and interpret assessment and results), (3) communicating assessment results (know how to give feedback to learners and share results to stakeholders), (4) providing assessment fairness (be aware of learners' diversity), (5) showing assessment ethics (uphold the students' rights and privacy, and report accurately), (6) applying measurement theory (understand psychometric principles), assessment for learning (using formative assessment), and (7) assessment education and support for teachers (continue professional development in assessments).

Slightly different from Deluca et al. (2016), Webb (2002), Mertler (2004) and Akayuure (2021) elucidated assessment standards which teachers ought to possess proficiency in (1) the selection and (2) cultivation of assessment techniques that are suitable for informing educational judgments. Proficiency in (3) the administration, scoring, and interpretation of both externally generated and teacher-generated assessment methods is expected. In addition, it is imperative that teachers possess a high level of proficiency in (4) utilising assessment outcomes to inform their decision-making processes about individual students, instructional planning, curriculum development, and overall school enhancement. The remaining three criteria encompass the following aspects: (5) proficiency in formulating sound grading methodologies based on student assessments; (6) effective communication of assessment outcomes to students, parents,

and fellow educators; and (7) identification and acknowledgement of unethical, illegal, and otherwise unsuitable assessment practices and the utilisation of assessment data.

Metler (2004) also elucidated that standards-based assessment literacy goes beyond test instruments and result interpretation. Knowing how assessment outputs can evaluate students' standards development, improve instruction, and maximise student learning programs is also important. Understanding assessments will show how they inspire and demotivate students, instructors, administrators, and staff to perform successfully. Xu and Brown (2016) further elucidated that not only teachers know how assessments can excite and educate stakeholders, but students also understand standards' learning objectives, structure, relevance to performance targets, and learning priorities. In other words, students need to involve the students in assessment practices.

b. Principles

In terms of assessment principles, Anggraena et al. (2022) acknowledged that assessments either it is an assessment *as, for, and of Learnings* are should not be separated from learning. These assessments are one a unit or intertwined elements in Indonesia's education system. According to Price et al. (2012), a connectedness of learning and assessment along with assessment principles is the assessment literacy itself which teachers should possess. This might be due that fifty percent of teaching and learning process is about assessments (Giraldo et al., 2018). Teachers in consequence are highly expected to use learning concepts and assessment principles to effectively create and execute meaningful learning experiences, fostering creativity, critical thinking, and innovation in students. So, to ensure the smooth functioning of the learning and assessment process, the government has implemented learning and assessment principles (Anngraena et al., 2022).

In line with the employment of learning principles, Anggraena et al. (2022) asserted the educators are required to focus on the following factors. The first factor is the learning should intentionally be tailored to align with the specific stage of development, the present level of achievement, and the individual learning needs of students. It also considers the unique traits and various backgrounds of students. The goal is to make learning both meaningful and pleasurable. The second factor is the purpose of the design and implementation of learning is to foster the capacity for lifelong learning. The third learning principle is that the learning process supports the development of students' competencies and character holistically. Moreover, ensuring pertinent education entails designing instruction in accordance with the students' context, environment, and culture, while also collaborating with parents and the community. Last principle is that the learning should be oriented towards a sustainable future.

Furthermore, Anggraena et al. (2022) specifically emphasized a number of assessment principles that teachers should now consider. The following constitutes the assessment principles:

1. *Integration*: this means that both learning and assessment are integrated elements. It should facilitate the learning and provide holistic information, as feedback for educators, students, and parents/guardians so that it can guide them in determining further learning strategies.
2. *Learning Objective Accordance*: this means that assessments are created and implemented in alignment with the assessment goals, allowing for flexibility in choosing the method and timing of the assessment to effectively meet the learning objectives.

3. *Fairness, Validity and Reliability*: this means that the assessments should be developed with fairness, proportionality, validity, and reliability in order to accurately describe the progress of learning, make informed decisions about next stages, and serve as a foundation for constructing appropriate subsequent learning programs.
4. *Simple and Informative Report*: This means that the reports on learning progress and student successes are concise and enlightening, offering valuable insights into the attained character traits and skills, along with subsequent action plans.
5. *Reflection*: this means that the results of the assessment serve as reflective material for students, instructors, education professionals, and parents/guardians to enhance the quality of learning.

Teachers who adhere to the principles of effective assessment has the knowledge and skills required to fulfil benchmarks of excellence (Stiggins, 1991). The criteria assert that effective assessments 1) are derived from and fulfil explicit objectives; 2) are derived from and align with appropriate and well-defined goals of achievement; 3) employ suitable assessment methods that are consistent with the intended purpose and goals; 4) adequately sample student performance; and 5) account for all pertinent factors that may introduce bias or distortion.

2.5. Levels of Assessment Literacy

Teachers who possess a comprehensive comprehension of assessment are adept at designing, selecting, administering, assessing, and utilizing high-quality examinations in order to promote the academic advancement of students (Celebre-Baird & Reid, 2023; Meijer et al., 2023). Stiggins (1991) differentiated the assessment literacy into three levels; low (users), moderate (practically literate), or high (advanced literate). Meanwhile, Xu and Brown (2016) classified the level of teacher assessment literacy into four categories.

Stiggins (1991) elaborated that the initial level is designated for those who primarily utilise data that has been produced by external sources. The individuals encompassed within this category consist of students, parents, members of school boards, lawmakers, staff members of education departments, education journalists, and any other public or private officials who rely on achievement data to inform their decision-making processes. The individuals in question possess knowledge regarding four essential characteristics of high-quality data, namely a well-defined objective, a suitable sample, identifiable sources of potential bias, and actionable outcomes. Such newly, qualified teachers are required to possess only a rudimentary comprehension of various assessment methodologies, encompassing both formative and summative approaches.

Different from the first level, Stiggins (1991) asserted that the second level of the assessment literacy involves practical literacy, which includes data generation and use. The group includes teachers of all levels, principals, guidance counsellors, school psychologists, superintendents, teacher trainers, and administrator trainers. Professional educators must improve their assessment literacy. Users, like functional literacy learners, must understand their decisions, their effects on pupils, and the data needed to make them. However, people must understand the qualities of good data, the factors that can damage it, and how to avoid them. Acknowledging that success goals without comprehension are impossible is crucial. Biased performance samples produce useless data. Interference variables must be controlled for assessment validity. User misinterpretation of assessment conclusions can waste valuable assessment resources.

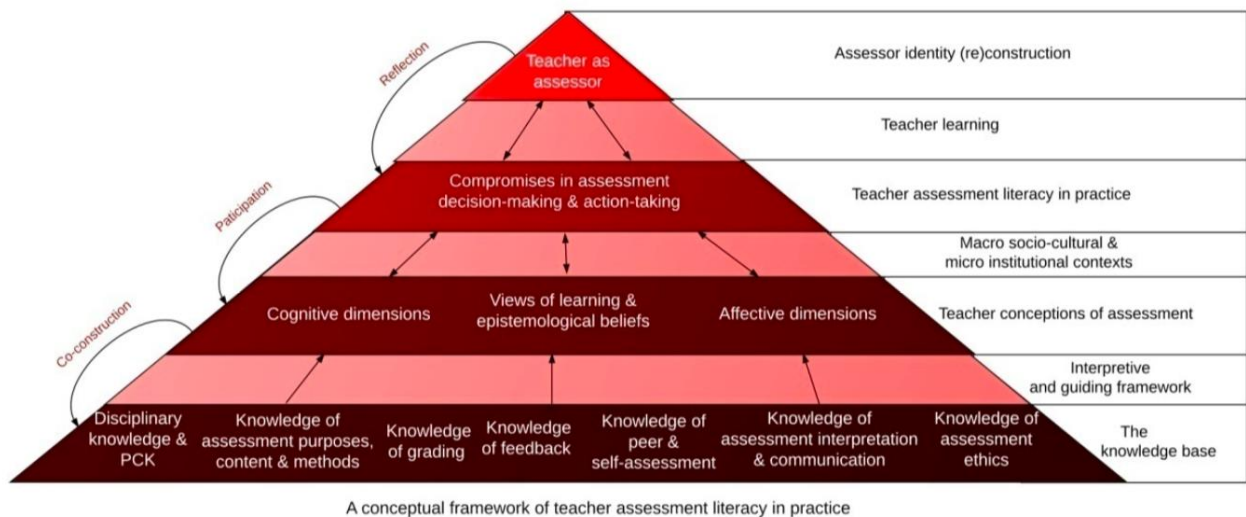
Stiggins (1991) posited that individuals who possess a high level of assessment literacy have the capacity to generate data that can be utilized by others, in addition to their proficiency in reliability and validity. These individuals are accountable for the

development of "high stakes" assessments, which are used to certify the minimum level of competence, select students for specialized programs, measure achievement on a large scale, and evaluate instructional programs. These assessments serve a variety of purposes. These individuals are professionals who specialize in measurement and evaluation and are required to exhibit a high level of proficiency in resolving assessment-related issues.

Different from Stiggins' concepts regarding the assessment literacy levels, Xu and Brown (2016) proposed conceptual framework of assessment literacy levels in practice.

Figure 1.

Adopted from Xu and Brown (2016, p.20)



As a teacher guide, Xu and Brown (2016) proposed a conceptual framework comprising four levels of interrelated practices. These include the knowledge base, assessment concept, assessment practice, and assessor identity construction. First, the foundation level includes a teacher's knowledge of their subject area (disciplinary knowledge), understanding of different grading and feedback purposes, ethical considerations in assessment, and how to teach their subject effectively (pedagogical

content knowledge). Teachers can develop and refine assessment practices with colleagues. These will lead teachers to understand imperative and guiding frameworks. The next level is the teacher's conception of the assessment itself. These include the teacher's cognitive and affective dimensions. A cognitive level pertains to teachers' epistemological comprehension of the subject matter, while affective dimensions concern teachers' emotive reactions to assessments. Understanding this level will guide teachers in comprehending socio-cultural and micro-institutional contexts.

The third level of teacher assessment literacy, which Xu and Brown propose, is the practice level. At this level, the teacher compromises the assessment decision-making and action-taking. In other words, teachers are highly expected to take learning actions based on the assessment results. Following this level, teachers are also expected to continue to level up their assessment skills, which then shifts their identity to become assessors. This level shows the transformative power of teachers in assessment literacy.

2.6. Procedure and Methods of Assigning Assessment

2.6.1. Assessment Procedures

In general classroom practices, Popham (2009) established a comprehensive framework for understanding assessment procedures by categorising different types of assessments based on numerous characteristics. The assessment is not limited to the traditional paper-and-pencil test, but it includes any kinds of formal and informal models. Popham identified two types of assessments' practices: summative and formative. The summative assessments can be referred to utilization of evidence-based assessment which aims to assess educational achievements upon the completion of a unit, course, or program. This can be in the form of a high-stake formal test with a variety of test forms, such as exams, essays, or projects. Since such assessment is restricted with time, limited feedback can be delivered to the students. On the contrary, the formative assessment is a

deliberate procedure in which teachers or students utilize evidence obtained from assessments to modify teacher current teaching practices (tailoring their instruction to meet the needs of all learners), and to strengthen students' learning motivation since it provides timely and targeted feedback.

In science classrooms, Bandede and Aluwatayo (2013) explicated procedures in science assessment can also be administered either formally or informally, considering to objectives of learnings and criteria of assessments. In formal settings, giving quizzes, homework, performance, or projects can be in the forms of response items in which students compose answers or select their response items. These can be delivered in forms of completion items, multiple-choice test, matching-items, or true-false items. These types of assessments require substantial validity and reliability. Meanwhile in informal settings, Bandede and Aluwatayo (2013) asserted that the assessments should also be able to obtain information which are valuable for students' learning. The informal delivery can be generated in forms of questionnaires, checklists, anecdotal records or peer-appraisal.

Related to this, Rustaman (2017) stated that there is a strong correlation between science instruction and science assessments. It means that assessment methods used in science learning should be aligned, one of which, with the inquiry-based nature of scientific research, promoting active learning and engagement in scientific practices. Abell and Siegel (2011, as cited in Joachim et al., 2020) introduced the STAL (Science Teacher Assessment Literacy) model, which consists of five components, to measure the level of assessment literacy among science teachers. The component comprises assessment values and perspectives on learning. This becomes an essential element of the model. The four sub-components consist of: (1) comprehension of the objective of

assessment, (2) identification of what needs to be assessed, (3) utilization of assessment methodologies, and (4) interpretation of assessment results and subsequent actions.

In the field of science, a teacher's understanding of "*what to assess*" mostly involves their conceptual knowledge and understanding of the experimental procedure, which students acquire as they develop their comprehension and encounter challenges (Joachim et al., 2020). Joachim et al. exemplify this concern in the field of Biology in which a common approach to discovering new information is through conducting experiments. This learning and assessment model is called an inquiry-based learning (Rustaman, 2017). In this context, assessments should be capable of measuring the students' proficiency in applying their scientific understanding, encompassing both factual knowledge of scientific concepts and comprehension of the scientific process. Hence, Xu and Brown (2016) argued that teachers should possess a thorough understanding of the STAL components to assess their students effectively and appropriately.

2.6.2. Methods of Assigning Assessment

Teachers should not rely on one single method of assessments. Often times a single way of measuring the students' performance is incapable of representing the students' overall performances. Stiggins (1995) and Santrok (2018) contended teachers who are knowledgeable enough about assessment literacy understand the appropriate utilization of distinct assessment methods within their classroom practices. Stiggins proposed four methods of assigning assessment to the students: selected response category, essay assessment, performance assessment, and personal communication. However, Santrok (2018) mentioned three methods of assessing students' learning,

including selected-response items, constructed-response items, and alternative assessment.

The first method which can be applied is through *Selected-Response Category or Items* (Stiggins,1995; Santrok, 2018). Here, students are required to read the exercise or questions, and choose the most suitable response from a limited set of options. The questions can be in forms of multiple-choice, true or false, quizzes, matching and fill-in-the-blank tests, and assignments. Another way is through *essay assessment* (Stiggins, 1995). Santrok (2018), however, classified these methods into two types: short-answer and essay items. Through this method, students are assigned a succinct writing assignment requiring them to generate original content. Following this, the text is evaluated in accordance with a pre-established set of grading criteria or standards. The grading criteria for essays typically revolve around the way in which the written response presents the given topic. The third method of the assessment which is believed less attractive is *performance assessment* (Stiggins, 1995); while Santrok (2018) classified this into authentic assessment. With this alternative, students are expected to show or develop a multidimensional product which can directly be evaluated, like portfolios. Not only is the issue of form addressed by the students, but the contents as well. The final assessment method offered by Stiggins (1995) is through what is named *personal communication*. Information regarding the achievements of the students is tracked by conducting a discussion, listening to students' responses to the given questions or initiating a talk with their parents and other colleague teachers who know better the students.

Furthermore, Girgla et al. (2021) placed a concept of adaptability and flexibility in assessment to meet students' diverse background. Hence, Girgla et al. proposed several

fundamental aspects which assure the assessment methods. The first aspect which needs to be considered is regarding the item content and format. These should be appropriately target-oriented. Another consideration in designing the assessment is based on the accessibility aspect. Also, all students must be well-informed and prepared for the assessment. It can also be associated with the mode of delivery, either it is through paper or online or through oral or written format. This aims to support students to readability to well perform during the assessment. Finally, the other essential aspect of the assessment delivery is reasonability. The assessments should not discriminate against students with a disability, for instance.

2.7. Assessment Approaches

2.7.1. Approach in Classroom Assessment

Since assessment is an integral part of learning activities, the assessment should be designed compatible with views of learning and motivation, be appropriate with learning targets and have high quality establishment (Santrock, 2018). In terms of its integration, Santrock (2018) asserted that teachers have responsibility to conduct assessment in three-time frames; namely pre-instruction assessment, during the instruction (formative assessment) and post instruction (summative assessment). In a pre-instruction assessment approach, teachers should look at prior students' grade either by giving the diagnostic tests, collecting their portfolios, or observing them for several days informally to know how well they perform in the class. If not conducting this approach, teachers might risk the class in either the instruction or the materials which are too advanced to their students.

Another approach is conducting assessment during instruction. This type of assessment is also called formative assessment or *assessment for learning* which functions as references not only for teachers to make decision regarding the next models

of learning activities if there are gaps between the actual level of student understanding and the target standard, but also for students to know their current proficiency and future improvement of the subject areas (Shepard et al., 2005; Santrock, 2018). This approach can also help teachers to detect which students need more individual attention. This approach is low stake evaluation which carries little even no point values. It can be run through group work or discussions, mini-quizzes, or self-assessment. Besides the formative assessment, teachers are also encouraged to conduct post-instruction assessment. This can be noted as a summative assessment, which is also familiar with a term “*assessment of learning*.” According to MacMillan (2014, cited in Santrock, 2018), this assessment approach gives information about to what extents the students comprehend the subject materials, how ready they are for the next unit level, what score they should deserve, and what types of feedback for student-parents, and what types of instructions would the teacher conduct.

2.7.2. Assessment Approach in Indonesia’s Context

According to the current national education curriculum, student’s learnings are evaluated using three approaches: *as, for, of* learnings. This is different from the previous assessment approach in which the assessment flows from *of, for, as* learnings (MoERCT, 2023). The assessment of the achievement of student learning outcomes in Indonesia’s schools should prioritize *assessment as learning* and *assessment for learning* when compared to using *assessment of learning*. The assessment *as learning* serves a purpose akin to assessment for learning, which is to serve as a formative assessment conducted throughout the learning process (Santrock, 2018). The assessment as learning engages students directly in the assessment activities. Participants were provided with the opportunity to acquire the skills to evaluate their own performance. Self-evaluation and

peer-to-peer assessment exemplified the concept of assessment as learning. Assessment as learning involves engaging students in the process of creating assessment methods, criteria, and rubrics/guidelines. This empowers students to have a clear understanding of what is required to achieve optimal learning outcomes.

Another approach in student's evaluation in Indonesia education curriculum is suggested through the assessment *for* learning. Assessment *for* learning is done during learning and used to improve current teaching and learning processes (MoECRT, 2023). Assessing student learning allows instructors to provide direct feedback, evaluate progress, and determine learning progress (Santrock, 2018). Teachers can increase student facilitation with assessment for learning. Formative assessments, such as assignments, presentations, projects, and quizzes, are examples of assessments for learning. Meanwhile, the assessment *of* learning is occurred after the learning (Santrock, 2018). Learning does not necessarily cease at the end of the year or when students graduate. After learning is complete, teachers assess learning to recognize attainment of learning outcomes. National exams, school/Madrasah exams, and summative assessments assess learning.

Because the assessment both a formative and summative is an integral activity to the learning process, the assessment plans must be incorporated into learning planning – the lesson plan (Anggraena et al., 2022). The summative assessment is conducted at the conclusion of the learning process to verify that the overarching learning objectives have been met. While the formative assessment serves the purpose of furnishing educators and students with information or feedback that can be used to enhance the learning process. The extent to which these two types of assessment are incorporated into a teaching

module or learning implementation plan is discretionary and contingent upon the learning objectives' scope (Anggraena et al., 2022).

Due to this rationale, educators or teachers ought to have the most comprehensive understanding of students' learning progress; therefore, teachers must possess the authority and autonomy to administer assessments that are tailored to individual students' requirements. This autonomy encompasses the ability to determine the criteria for accomplishing learning objectives, design the assessment, determine the implementation time, utilize assessment techniques and instruments, and process assessment results (Anggraena et al., 2022). Preferences concerning midterm assessments fall within this scope of discretion. Educators and educational institutions possess the jurisdiction to determine whether to administer such an evaluation.

2.8. Previous Study

Several studies, outside of Indonesia's context, have examined the teachers' understanding and knowledge of assessment literacy (Gunn & Gilmore, 2014; Cowie et al., 2014; Willis et al., 2013; Xu & Brown, 2016; DeLuca et al., 2018, 2019a, & 2020; Clark et al., 2022). Meanwhile, in terms of assessment literacy concepts, Gunn and Gilmore (2014) explored 37 early childhood teacher education students' assessment conceptions, by examining the extent to which these conceptions engage with current policy expectations, and by considering the extent to which they can shed light on student teachers' developing assessment capability. In addition, Clark et al. (2022) identified initial level of teachers' diagnostic assessment literacy, and they found that the teacher's conceptions of fundamental diagnostic assessment concepts, their comprehension of the assessment process and outcomes, and the impact of these outcomes on their instructional decision. While Cowie et al. (2014) and Willis et al.'s (2013) tried to investigate and detect what student teachers had learned about assessment in their university courses and

while on practicum. They revealed that the student-teachers' views of assessment had changed over the teaching experiences.

Shifting our attention to Indonesian context, some facets were examined in studies regarding the teacher assessment literacy. According to Bahtiar and Purnawan' (2020) finding that most teachers possess limited familiarity with the concept of Language Assessment Literacy (LAL), although demonstrating commendable levels of proficiency in LAL. Another study was undertaken by Aria et al. (2021) who discovered that most educators engaged in a variety of assessment activities, suggesting that they appeared to implement assessment for learning (AfL). Also, Nyudak et al. (2022) and Purwani et al. (2022) found in their study that the assessment literacy of the teachers in Bali is fair, with some influencing factors. Further, the latest research finding (Hutami & Putro, 2023) states that there are matches and discrepancies between perceived assessment literacy and the demonstrated assessment among the teachers in Banyumas regency.

Research on science teachers in Indonesia has dealt with various topics of science learnings. A study conducted by Faesal et al. (2019) has examined the experiences of Indonesian pre-service teachers in creating STEM-TPACK learning websites, focusing on the difficulties they face in communicating across different disciplines and integrating material. Another study is a comparative analysis of novice science teacher students from Indonesia and Thailand which was undertaken by Nuangchalerm and Islami (2018). They revealed no statistically significant disparity in their mean scores in science subjects, highlighting the commonalities in science education between the two nations. Furthermore, a survey study was conducted by Nida et al. (2020). The aim of this study to investigate the perspectives of Indonesian science teachers about the teaching of socio-scientific issues (SSIs). The study revealed that the teachers had different levels of

knowledge with SSI-based teaching methods and faced challenges in implementing them, despite their enthusiasm in incorporating SSIs into their lectures.

Iswatiningsih et al. (2022) revealed that training on designing e-assessment improves junior high school teachers' understanding of creating e-students' worksheet. This study focuses on how training improve teachers' competence in worksheet making. Unfortunately, however, related to science teacher's assessment literacy, there are very limited numbers of research studies in Indonesia explicating how high school science teachers' assessment literacy understanding along with their classroom assessment practices in accordance with the assessment principles stipulated in the curriculum. Thus, this research wants to delve into this particular concern.

CHAPTER III

RESEARCH METHODOLOGY

This chapter focuses on the methodological aspect of this research, particularly in investigating and analyzing teacher assessment literacy. In this chapter, I will describe the research methodology selected to examine this essential skill. I also provide a comprehensive explanation of the techniques used to gather and analyze data, explaining the rationale behind each strategy in terms of its ability to shed light on instructors' assessment practices and their comprehension of effective assessment procedures. Furthermore, this chapter will provide a plan for examining the complexities of teacher assessment literacy in the selected study environment.

3.1. Research Approach

The approach to research refers to the comprehensive set of processes employed in a study, encompassing general assumptions as well as specific methodologies for gathering, analyzing, and interpreting data (Creswell & Creswell, 2017). This research adopts a qualitative interpretivism paradigm to investigate the nature of research issues related to MA science teachers' assessment literacy and their classroom practices within the implementation of the current curriculum. Interpretivism in qualitative research seeks to comprehend how teachers in real-life experiences as individuals interpret assessment techniques, the circumstances that impact their choices, and the significance they attribute to assessment data (Cohen, et al., 2018). According to Creswell and Creswell (2017), a qualitative study is a form of investigation that delves into and offers profound understanding of real-world issues. To fully comprehend this study, it is necessary to elaborate on the elements of the qualitative approach used, including the research design and the specific methods or procedures followed throughout the investigation. Therefore,

by emphasizing individual perspectives, a detailed examination is conducted to understand how teacher views, backgrounds, and classroom environments influence their assessment literacy.

3.2. Research Design: Descriptive Case Study

To acquire comprehensive data and effectively capture the ever-changing nature of classroom assessment practices, I opted to utilise a specific qualitative research design which is a descriptive case study that more accurately reflects the present condition of the teachers. Qi (2009) asserted that the descriptive case study focuses on specific individuals or groups and highlights key events that are important to the case. Also, Cohen et al. (2018) contended that this research design is characterized by its precise, detailed, and analytical traits. Thus, applying this qualitative research design will deeply provide a nuanced understanding of teachers' perceptions and practice as well as how the assessment principles shape their assessment practices. Furthermore, Cohen et al. explains that the descriptive case study enables a researcher to deeply explore the personal experiences of an individual teacher/s or a small group. Conducting classroom observations, interviews, and document analysis can effectively capture their decision-making processes, opinions regarding assessment, and the influence on student learning.

3.3. Research Context

Madrasah Aliyah (MA) is an Indonesian secondary Islamic school, whether public or private, that is comparable to a Senior High School (SMA) in terms of educational level and duration in Indonesian education system, which is three years study at secondary level. Operated under the authority of the Ministry of Religion, this institution offers a curriculum that closely mirrors that of other high schools, but with a stronger focus on Islamic subjects such as Fiqh (Islamic law), Akidah (Islamic religion), and the Arabic

language (Kusaeri, 2018). MA schools are flourishing across the country. MoEC (n.d.) states that there is a total of 10,106 MA distributed over several provinces. According to Indonesian Central Statistics Agency/BPS (2013), there are 593 Madrasah Aliyah schools located in West Nusa Tenggara (NTB) with 10.750 teachers in total.

This investigation was conducted during late May and early June 2024 at a private MA situated in an urban area of West Lombok, NTB. This school became one of the sample schools in *Asesmen Nasional* program in 2022 in which the school principal, some teachers and students were respondents of the program. Interesting to note that the school has been operational for over five decades and is currently recognized with an A rating of accreditation level, indicating exceptional quality. Currently, there are more than 720 MA students enrolled in the school, comprising both male and female students. It is worth mentioning that this institution is of considerable size, accommodating boarding students from elementary to junior high and secondary high school levels, within a campus of around 10 hectares. The school's achievements exceed those of all other private schools in the region. In essence, the school's performance rankings exceed those of most other private schools, including public schools in the area. In addition, the school has sufficient resources to facilitate teachers' instruction and students' learning, including well equipped language and computer laboratories with essential materials and equipment.

In addition, the school has implemented a full-day schedule with six days of instruction per week. Morning classes from 7.00 am to 12.30 pm are designed to study general subjects such as math, social, and science subjects. In contrast, the afternoon class intends to study religious subjects such as Tafsir, Hadiths, and other traditional Islamic books. Classroom observations were carried out in year 10 classes, corresponding to students aged 14-15. Since there are 28 to 30 students in each parallel class within the

year group, and to preserve the voluntary aspect of participation, the teachers were granted the authority to choose which class to watch. In collaboration with the school principal, the teacher and principal agreed to choose year 10 classes for the observation.

3.4. Research Subjects

The subjects or participants of this study were selected from science teachers in one of MA in an urban area in West Lombok, NTB. These individuals have been formally registered with the Regency Ministry of Religion office and have successfully completed the teacher certification process. Using convenient sampling technique, three teachers were selected from a group of 15 science teachers of general Male and Female Madrasah Aliyah. The selection considered their fields of expertise and length of teaching experience as well as their willingness and availability to participate in. The selection of this sample method was driven by the goal of acquiring complete insights from individuals who possess the requisite knowledge, competence, and experience in the field, notwithstanding the possibility that their perspectives may not be representative or generalisable (Cohen et al., 2018). An exclusion of female teachers in female classes in this study was due to unwillingness to participate. Thus, the researcher cannot provide wide ranges of the sample genders for generalizability findings.

The participants

The initial participant is Mr. Budi (a pseudonym), a male Biology instructor with 11 years of teaching experience at the current educational institution. Before becoming a teacher at his current school, Mr. Budi had teaching experience at other schools following the completion of his bachelor's degree. He holds a bachelor's degree in biology education from a public institution in Mataram. Having worked as a teacher, he has participated in numerous teacher professionalism training sessions and is highly qualified to receive a

professional teaching diploma. Aside from his normal teaching duties, he also assumed the position of laboratory head and is currently being appointed as the coordinator of the school's annual scientific olympiad team, which is open to Madrasah Aliyah students.

Another participant is Mr. Rozi (the pseudonym), who has also been employed as a teacher at the present educational institution for a duration of 19 years. He graduated with a bachelor's degree in chemistry from the same university as the previous participant and achieved the status of a certified teacher. Furthermore, he actively engaged in multiple teacher training initiatives organized by both regional and central government entities, with a specific emphasis on pedagogy, evaluation, and educational administration. In addition to his primary function as a Chemistry teacher for classes 10 and 11, Mr. Rozi currently serves as the vice principal for student affairs. In this capacity, he is responsible for addressing and resolving student complaints. He many times needs to visit student-parents' home for reasons.

The third participant is Mr. Ujang (the pseudonym). Mr. Ujang teaches the subject of Physics in both the General and Special Classes of Male Madrasah Aliyah. He has dedicated his life to teaching in the present Madrasah Aliyah for a period of 17 years. Prior to commencing his teaching career, he obtained a bachelor's degree in physics from a local public institution in Mataram 18 years ago. Like the two prior participants, Mr. Ujang has also participated in numerous training courses focused on enhancing teacher development skills. Receive instruction in lesson planning and assessment through training provided by the regional office of the Ministry of Religion Affairs. Nevertheless, Mr. Ujang possesses a strong interest in computer science, which has endowed him with extensive knowledge in computer operating systems. Subsequently, he has also been

instructing students in computer proficiency and overseeing the maintenance of a reliable internet connection for all teachers and staff.

Table 1

Participant Demographic

No	Gender	Subject	Level of Education	Years of Experience	Status of Occupation	Professional Teacher Certification
1	Male	Biology	Undergraduate	11	Permanent	Granted
2	Male	Chemistry	Undergraduate	19	Permanent	Granted
3	Male	Physics	Undergraduate	17	Permanent	Granted

3.6. Data Collection Methods

In order to obtain comprehensive data, qualitative research data were collected from the research subjects through two times school observations followed by six-times classroom observations along with documents of lesson plannings, and a semi-structured interview.

a. *Classroom Observations*

In order to acquire a more profound comprehension of the issue, the three teachers in the field of chemistry, biology, and physics have been observed for three teaching sessions using an observation note. Prior to the observation, permission from the teachers have been granted. The observation note is adopted and or developed from Demirdogen and Korkut (2021) (see Appendix 2). During the observation, the researcher acted as the participant observer who structurally observed the events using structured and semi-structured observation forms. Using observation field notes, the researcher directly observed teachers conducting a variety of formative assessments during their teachings.

I realized that my presence as the participant observer could interfere with the classroom nuance; however, my main intention is to see teachers' activities on how they perform the assessment models and approaches. To get trustworthiness, I videotaped two of the classroom observations in which I could re-check my field notes during the analyses. So, I positioned the camera in the corner of the classroom in order to capture all learning activities throughout the teaching time. Additionally, the utilization of artefacts such as teachers' lesson plans for two meetings can serve as valuable resources and further insight for collecting qualitative data (Creswell et al., 2015). However, the opponents argue that qualitative research is lack of reliability, validity and generalizability (Sharma, 2010) since its aim not to generalize the finding, but to comprehend a particular issue.

b. *Semi-structured interview*

Another main research instrument which was employed in this study is a semi-structured interview that lasted forty-five minutes to an hour. An interview is believed as a powerful way of collecting data to get more description and explanation of social world of the individuals (Yeo et al., 2013). The interview can be driven in forms of structured, semi-structured, and unstructured forms. In structured interview, participants get the same questions with closed questions; however, in the semi-structured interview used in this study, the participants utilized pre-prepared questions but with some flexible and latitudes to ask further questions which the research considers essential (Bryman, 2008). In this study, three teachers (chemistry, biology, and physics) from one of private Madrasah Aliyah in West Lombok, NTB, have participated in semi-structured interviews. The interview were conducted in the first participants language - Indonesian, not in English in order to avoid misunderstanding and especially to ease them in understanding the

questions. The flow of the interview was unstructured to semi structured. The interview questions are derived and developed from the research done by Demirdogen and Korkut (2021) (see Appendix 1).

The interview questions address teachers' understanding and classroom assessment practice, both of which fall under the framework of assessment principles. In other words, it primarily delves into all research questions either in the levels of understanding and practices of assessments. Duration of the interviews varies per participant. The interviews with Mr. Budi (the first participant) lasted for 48 minutes, with Mr. Rosi lasted for 53 minutes and with Mr. Ujang lasted for 45 minutes. All the interviews were audio-recorded with permission from the participants to address trustworthiness and ethical aspects of the data collection procedure. The transcripts from the three interviews were de-briefed with the participants for clarification purposes. It is essential to note that the interview protocol and content validity of the interview has been priorly judged by two experts in the field.

After conducting the semi-structured interviews, I dedicated two to three timeslots to transcribing all the audio recordings. This allowed me to thoroughly review the interview process and ensure the correctness of the material obtained. Due to the verbatim transcription, the initial transcript of the audios using the application - YouTube was disorganized. As a result, I had to spend more time manually reviewing the audio from the early transcripts. I evaluate the aspects that were successful and identify areas that should be enhanced for additional validation. To enhance the readability and significance of the transcripts, I manually organized them by eliminating redundant words or phrases,

such as "eee... apa...ya ya or ok ok," and using short terms to provide a more comprehensive understanding of the responses.

Subsequently, I requested each of the participants to analyze and provide feedback on the transcripts before translating the transcripts into English. I have come to the realization that the specified techniques likely led to the distortion of data and diversion from its intended meaning. To prevent this problem, peer de-briefing is incorporated into the reference to ensure the reliability of data transformation and consistency of content. This includes obtaining feedback from participants and including a verbatim version. The following table is the example of the transcribing processes:

Table 2

Example of the Transcribing Processes

Version	Interview Transcripts
Original	Kalau pengalaman saya ya...eee.. untuk nilai yang di kelas lain ya saya ee..secara pribadi selalu menilai eee apa tidak hanya satu sisi begitu Jadi pembelajaran nilainya itu atau nilai-nilainya tidak hanya dalam bentuk eee apa ujian tertulis dalam bentuk eee ulangan harian tertulis Kita juga melakukan nilai-nilai yang lain gitu contoh nanti kita juga menilai sikap mereka saat melakukan proses pembelajaran tidak ada kegiatan-kegiatan praktikum yang kita lakukan kita nilainya juga nah aktivitas mereka kita nilai eee
Revised	Kalau pengalaman saya untuk penilaian di kelas sains, saya secara pribadi selalu menilai tidak hanya dari satu sisi. Setiap pembelajaran, penilaiannya tidak hanya dalam bentuk ujian tertulis atau dalam bentuk ulangan harian yang tertulis, tetapi juga dalam bentuk penilaian-penilaian yang lain seperti menilai sikap mereka saat melakukan proses pembelajaran, kegiatan-kegiatan praktikum yang kita lakukan, dan penilaian pada semua aktivitas-aktivitas mereka.
English Translation	Based on my experience in assessing students in science classes, I consistently evaluate from multiple perspectives rather than solely relying on one approach. Each lesson is assessed not only through written exams or daily tests, but also through other evaluations that include assessing students' attitudes during the learning process, their performance in practical activities, and the overall assessment of their activities.

3.7. Data Analysis Methods

Collected data from the interview are analyzed using thematic analysis with application of inductive coding procedures. Thomas (2003) and Braun and Clarke (2006) provide six stages for applying thematic analysis using inductive procedures in qualitative research. The initial phase is data familiarization, which seeks to gain a comprehensive understanding of the data derived from recorded voices of the participants. This can be accomplished repeatedly by thoroughly examining all interview transcripts, revising the transcripts and then translated them into English. The next step is the process of coding the data. During this stage, a Quirkos coding was performed by giving labels to text segments. After completing this stage, the theme searching process begins. This involves carefully examining the various coded patterns to detect any similarities or overlapping contents within the codes.

The fourth stage is reviewing themes. At this point, the discovered themes are subject to a critical evaluation to determine if they have been clearly defined, if they have adequately captured the subtle details of the data, or if any data should be discarded. Next, the fifth stage involves the process of clearly defining and assigning names to various themes. At this point, the researcher provides detailed explanations of each theme, clarifying its essential characteristics and its relationship to the data. Ultimately, the stage concludes by documenting the analytical findings. The procedure entails elucidating the research topic, expounding on the methodology, and thereafter conducting a comprehensive examination of the chosen themes supported by data.

Meanwhile, in analyzing the collected data from classroom observations, the researcher quantitatively identifies patterns in assessment delivery practices. The analyses were started from reviewing the observation notes, transcribing certain interactions

between teachers and students verbatim, and coding kinds of questions, responses, and feedback. Inductive procedures in coding using Quirkos application in the observation notes along with the document have also been applied. However, the thematic analysis may yield inadequate outcomes if erroneous data is included. This might occur from insufficient analysis or simple paraphrase and summarization of the data (Braun & Clarke, 2006). The processes of coding and theme using Quirkos can be seen in the appendix 4.

3.8. Ethical Issue of the Research

The researcher allocated a period of two weeks in the research environment, actively engaging with the research subjects to establish a sustained level of involvement. This study began by introducing the research with the necessary permits obtained from Universitas Islam International Indonesia. Additionally, this research has obtained approval from the school headmaster. Next, the researcher conducted both observations for three times (one observing the school along with small discussion with the researchers and twice for classroom observations) along with document analyses by physically being present at the research location and semi-structured interviews using Zoom application (recording the audio).

The participants will be provided with the findings of teachers' assessment literacy at both the perception and assessment practice levels, after conducting a thorough analysis of the data. Participants concurred with the findings; hence we ensured member checks. Furthermore, the active involvement of all teacher participants has been guaranteed by means of a written consent form. Participants are granted the autonomy to discontinue their involvement in the research at any time, and it is ensured that only researchers have the privilege of accessing the data. Additionally, the participants are notified that

pseudonyms are employed when documenting the study. By using these measures, ethical concerns in research, such as safeguarding participants from potential damage and ensuring anonymity, are addressed (Fraenkel & Wallen, 2006).

3.9. Limitation of the Research

Research using case study design has its limitations, despite having rich details of data in identifying science teachers' understanding on the assessment literacy concepts and practices together with their challenge and opportunities. The limitation in this research lies in a single instance of the findings which cannot be generalized. It is due that the samples only consist of male teachers in male-classes of the Islamic-based boarding school. The exclusion of female-teachers since having difficulty to get permission to conduct the research in the female-classes. Thus, findings of this study might (not) be applied to the other contexts. Apart from not applying triangulation in data analyses, my interpretation of the data can influence the results. It is because the case study often relies on subjective experiences, and there might be inaccurate information. Furthermore, data collection for conducting this study can be time-consuming and expensive.

CHAPTER IV

FINDING AND DISCUSSION

This chapter addresses and synthesizes findings from the four research questions. The research questions of the study are (1) to what extent do private Madrasah Aliyah (MA) science teachers understand their current assessment literacy? (2) what are the current specific classroom assessment practices utilised by MA science teachers which are align with the assessment approaches, principles, or procedures? (3) how do MA science teachers design and implement classroom assessments within the assessment components? and (4) what challenges and opportunities do MA science teachers face in implementing assessment practices? Using the information gathered from the classroom observation along with the teachers' documents - the lesson plannings, in addition to interviews that were generated in semi-structured. Data are presented and synthesized in accordance with the research questions obtained from the three instruments. Although it may exhibit duplication, the adaptation of this delivering technique is regarded as appropriate and manageable in order to guarantee the finding lucidity.

4.1. Findings

4.1.1. Teachers' Understanding on the Assessment Literacy

Finding from the inductive thematic data analyses with help using Quirkos identify that there are five themes which are used to elaborate the teachers' assessment literacy. The themes are teachers' understanding of definition, purpose, question types, levels of question difficulty, and approaches. These findings are primarily generated from the interview transcripts which are supported by the classroom observation notes and the

documents – lesson plannings. The themes are then deductively used to describe the teacher's understanding.

4.1.1.1. Teachers' Conception of Assessment Literacy

In terms of conception, the finding depicts that science teachers generally have a positive view of scientific literacy assessment. They understand that this assessment literacy is important to measure student learning progress and to provide useful daily feedback. Science teachers also realize that scientific literacy assessments can help them to improve the quality of learning in the classroom through the conduct "*assessment for learning*". They argue that the assessment literacy is not only about the teacher's ability to assess cognitive of individual students, but it includes to assess the whole learning activities which support the student's development. This encompasses not only daily assessments in several exam formats that will be quantitatively evaluated, but also additional factors such as efficacy, attitudes, and hands-on tasks like projects. The understanding of the assessment concept and purposes is regularly upgraded through attending regular training on teaching and assessment in their school. Annually, the school invites school supervisors from the regional education office to teach them about the updated education-teaching and assessment innovations.

Since they regularly attend the training from supervisors, the teachers are aware of an ideal assessment model or type/s and procedures which can be used to evaluate their student levels of knowledge. Also, they get an updated assessment practice together with the teaching contents in the current curriculum, the Merdeka curriculum – a new curriculum introduced from the central government. They said that the Merdeka curriculum with fewer chapters and supported by student-working sheets (LKPD) is more

practical where students are more inclined to project-based learning and assessment, compared to Curriculum 2013 (*henceforth: K-13*). They mentioned that the K-13 contains ten chapters (twice as many as in Merdeka curriculum) in which the students are primarily assessed based on their cognitive ability on theoretical concepts. With their conceptual ability gained from the trainings, the teachers can differentiate and create levels of question difficulty based on learning basic competency in the curriculum; however, not as good as the test questions made by the central government. As literary stated by Mr. Rozi,

"..... we at lower-level schools are not as skilled as those at the national level in using these stages, looking at basic competence (KD), then adjusting to the level desired by KD."

Source: interview transcript with Mr. Rozi

Discussions, both with fellow subject teachers and other subject teachers, are also an important part of continuing to upgrade their conceptual understanding of the assessment literacy. Mr. Budi narrated that

"within one month we routinely evaluate how to teach and how to appropriately assess students' performances, according to the experience of each teacher in accordance with the latest curriculum implementation."

Source: interview transcript with Mr. Budi

However, Mr Ujang narrated that;

"In the past, we have regular meeting discussing about teaching and assessment with other physics teacher in West Lombok, but for the last 3 years we have no meeting ..."

Source: interview transcript with Mr. Ujang

Their understanding regarding the assessment literacy concept was supported by the opportunity given to Mr. Budi and Mr. Ujang to be part of the team members for compiling Olympiad questions at the Madrasah Aliyah level at the sub-district level for the last few years.

4.1.1.2. Teachers' Preference in Assessment Types

Findings from interviews and documents show that the science teachers have various ways of evaluating their students' learning in both formative and summative assessments. Firstly, in the formative assessments, the teachers opt to conduct oral and written tests. Oral assessment is usually used for checking the student's prior knowledge in every beginning of the lesson and checking understanding during or after the material delivery. Mr. Ujang opines that oral tests in the classroom is indeed simple and direct technique to know the student's real knowledge. Another alternative is a written test in the form of quizzes, short-answer questions, or multiple choice. This is also considered effective, by Mr. Ujang, for both checking students' knowledge and understanding. However, it is different for the biology teacher, Mr. Budi, who contended that a portfolio or project is likely the most effective way to assess students understanding in Biology lesson. He stated that,

"In my view, it is inadequate for students to merely learn biology contents without attaining proficiency in practical application or project completion. By solely relying on rote memorization, students miss out on crucial aspects of scientific education (IQ.4-5)."

Source: Interview transcript with Mr. Budi

Nevertheless, if dealing with the summative assessments, the teachers narrated that they could adopt at least three out of five types of assessment questions which are based on the government guideline. Referring to the guideline and the results of the latest teacher meeting, the school obliges individual teachers to compile three types of questions which will be used as posttest in final semester questions. All teachers narrated that those questions are multiple-choice (simple and complex levels), short answer, and description questions. Apart from these genres of the test, the teachers are also familiar with the levels of questions they must design. They mentioned that the questions consist of basic,

intermediate, and advanced levels which are 20%, 30% and 40% respectively. Of the three levels, higher order thinking skill questions are loaded 40% of the total questions.

Teachers' Approach in Assessment

Based on the data analyses, it has been determined that all teachers utilize both *assessment for learning* and *assessment of learning* in their approach to assessment practice. The identification of these two themes is evident from the interview transcripts and is further substantiated by the classroom observations. During study hours, teachers assess students' learning by evaluating their knowledge and understanding, providing direct oral feedback, and facilitating peer-assessment. Engaging in practical activities, such as in the field of Biology, is also seen as a component of the assessment process for learning. Nevertheless, the diagnostic test, in the form of a brief written test, is exclusively administered by Mr. Rozi prior to the commencement of the semester. Mr. Rozi mentioned that he gives 2-3 questions as a diagnostic test. Meanwhile, an *assessment of learning* for students is conducted by all teachers at the conclusion of the semester. This involves administering written examinations to students and thereafter providing them with a study report that includes their cumulative scores. In the following table is the comparison of the teachers' view on the assessment literacy concepts, types of assessment preferences, and approach.

Table 3.

Summary of Teacher's Understanding on Assessment Literacy

Views of Assessment	Respondents		
	Mr Budi	Mr Ujang	Mr Rozi
Understanding of Assessment	“To me, effective assessment in Biology is assessing students through practice and project-based learning, not just theory.”	“In my opinion, assessment literacy has a broader meaning. We not only assess students individually, but also as a whole to understand the average ability of the class, identify students who stand out and need help, and paint a general picture of the class. So, it is the teachers' ability to write questions is good.”	In my opinion, Assessment Literacy is a way to assess students as a whole, both from a cognitive and affective perspective. This assessment is not just about numbers and test scores. The ideal assessment must be varied.
Assessment Preference	Both formative and summative, but he prefers formative assessment over summative assessment. Formative: lab practice, portfolio, Summative: Final semester tests (multiple-choice, short answer questions, and description)	Both formative and summative, but he prefers formative assessment over summative assessment. Formative: review video and short-answer questions. Summative: Final semester tests (multiple-choice, short answer questions, and description)	Diagnostic Test in early learning (2-3 short answer-questions) Formative during the study: oral presentation and written tests (short-answer questions, and multiple-choice) Summative as the final assessment (multiple-choice, short answer questions, and description)
Assessment Approach	Assessment for Learning Assessment of Learning	Assessment for Learning Assessment of Learning	Assessment for Learning Assessment as Learning Assessment of Learning

4.1.2. Teachers' Classroom Assessment Practice

The qualitative approach with the help using Quirkos was also carried out to code the data from the classroom observations notes. I used an inductive thematic approach to cater for emerging sub-categories or themes and found the following two main themes: *Knowledge Checking and Improvement, and Instruction and Monitoring towards Learning*.

2.1.2.1. Checking Knowledge and Understanding

The process of conducting analyses of the data from the classroom observation notes is somehow challenging since pedagogical and assessment practices have robust correlation. An inductive content analysis utilizing the Quirkos application has shown that teachers are frequently represented engaging in checking students' knowledge and understanding activities. These become the main core of the learning activities where teachers continuously conduct the assessment for learning. At the outset of each lesson, the teachers assess the students' understanding of the previous meeting by posing open-ended questions to specific students or the entire class. They may also provide information to the class or watch videos related to the prior subject matter. The students' responses to the knowledge checking questions are insufficient, sometimes consisting of brief oral answers.

There are multiple methods for assessing prior knowledge, which vary based on the teachers' preference. For example, Mr. Budi requests that students either present their assignments in front of the class or from their designated seats. The students have the option to expand upon what they remembered or refer to their written notes. Mr. Ujang, the Physics teacher, employs several methods to engage students in watching a video

related to the previous topic. He then prompts them to provide a concise summary or recap of their gathered material. On the contrary, Mr. Rozi focuses on providing a brief overview of the previous course and asking the students direct short questions to which they already know the answers. The following excerpt from Mr. Rozi's classroom teaching discourse presents an illustration of a knowledge assessment task.

Excerpt 1.

***Teacher:** Before we continue to discuss mole calculations, does anyone still remember the basic laws of chemistry? What are the basic laws of chemistry? Anybody knows?*

***Student:** Me, Sir., Law of Conservation of Mass, Law of Fixed Proportions, Law of Multiple Comparisons and Law of Volume Comparisons*

***Teacher:** Does anyone know what the law of Conservation of Mass says?*

***Student:** The law says that mass of substances before and after is the same*

Source: classroom observation notes of Mr. Rozi

The extract illustrates how the teacher commenced the learning process by initially verifying the students' prior knowledge through short-opened questioning and requesting the answers that the students already knew. This is also a component of providing students with a diverse array of information and answers that they have acquired, in which the subsequent materials are consistent with the previous subject matter.

Subsequently, students' knowledge is elicited and developed. In this section, the instructors also employ unique delivery methods. Mr. Ujang instructs students to watch a video in order to elicit the course material, while Mr. Budi instructs students to peruse the contents of the screen where the materials are presented. In his classroom, Mr. Rozi distributes the student's worksheets, which include details of the instruction and questions, to elicit the primary materials and facilitate comprehension of the subsequent material and queries. In the two observations, Mr. Rozi prefers to organize the students

into groups to help them work collaboratively to solve the assigned task and to provide support during the discussion. All instructors provide a concise explanation of the primary learning activities in advance, in addition to providing brief instruction. This eliciting and building knowledge share the overall learning delivery.

Furthermore, it is imperative that teachers ensure that students have a clear understanding of the material they have learned by offering clarification questions or allowing them to perform in front of the class. The second highest proposition in the observation notes is the sub-theme of "*checking the understanding*," which is quite often delivered by the teachers. The chemistry teacher highly encourages and involves students in presenting the results of their discussions to the class. If an error is made, the teacher requests feedback from other students. The sub-theme approach is a peer-assessment in which students evaluate the work of their peers and give constructive feedback. Also, the teachers at the same time provide constructive feedback with additional examples if the problem remains unresolved. Nevertheless, Mr. Ujang and Mr. Budi pose brief inquiries that necessitate the prompt and precise responses of one to three students. In addition to conducting direct comprehension assessments, the three teachers share the same objective of assigning assignments that will augment students' comprehension. The following is the excerpt on how the biology teacher checks the students' understanding.

Excerpt 2.

Teacher: *How do we know that river water is not polluted?*

Student: *Me, Sir., If there are many living creatures in it, Sir., such as fish because clean water contains a lot of oxygen which is needed by living creatures in the water.*

Teacher: *Yes, that's right, fish will breathe oxygen which is mixed with water.*

Source: classroom observation notes of Mr. Budi

From this excerpt, it is noticeable that the biology teacher wants to ensure the students get good understanding regarding the current subject he has just taught. This sub-

theme is also called assessment for learning with which teachers can utilize the results as a reference for the better following teaching. Thus, this part is somehow essential for the teacher to review his explanation whether or not the subject materials can be easily understood by the students.

Instruction and Monitoring

Even though instruction and monitoring are not directly involved in assessment practice, they do play a substantial role in the success of learning and assessment. Teachers are required to exhibit this pedagogical skill, which is a part of the assessment itself. Consequently, it is noteworthy that the instruction and monitoring sub-themes significantly contribute to the flow of learning. On the one hand, the learning outcomes are improved by the provision of explicit and organized instructions at each stage of the classroom learning process (Wang et al., 2014). Conversely, learners may experience confusion as a result of unclear and disorganized instruction. For example, the chemistry teacher who was observed explicitly instructs the students to form discussion groups with six students each. The teacher states:

[Please count 1-6] wait for all students to finish counting, then the teacher instructs “[to ease you understand the subject, please make your discussion groups based on your numbers, group 1 over there (while pointing at the position),]. The students immediately move to find their friends who have the same number.”

Source: classroom observation notes of Mr. Rozi

Besides giving good instruction, regular monitoring is an essential part in the learning since some students often fall asleep during classroom instruction. The teachers contend that it is very common in school-based Islamic boarding to discover students sleeping in the classroom. During the six-times field observation, it was found that at least

four to five learners were sleeping while the teachers were presenting the subject matters, particularly in the second to fourth teaching session. The teachers whom we have observed frequently wake up the sleeping students and request them to wash their faces. This shows that students are not ready to study. Thus, the issue of students' readiness becomes one of the assessment constraints in giving effective assessment. It is also in line with the findings from the interviews that approve such circumstances.

4.1.3. Factors in Designing and Implementing Assessments

Regarding the factors contributing to design and implementation of the assessment, I found two main themes from the interview transcripts which include learning approaches and assessment design and implementation.

Teachers' approaches in designing and implementing the types of assessments vary. Besides referring to the government guideline regarding the approach, each teacher during the teaching periods has his own considerations. According to Mr. Rozi, the assessment approach differs depending on whether it is early, middle, or final. The range of assessments is contingent upon the specific activity. Initially, my assessment methodology involves evaluating the condition of the students, including their cognitive capacity, which may be determined by several tests. At this point, we can proceed with developing the assessment and its various forms. In addition to that, I consider students' preferences. If individuals prefer practical activities, I tailor evaluations to align with this preference by incorporating practical tasks, such as light practical. Here, I can see and assess their performance during practice activities. Doing this lets me discern which students engage more in extracurricular pursuits. However, it is not possible for me to

evaluate practical activities such as this through written assessments. Nevertheless, observing the outcomes through student portfolios or activity reports may be possible.

Mr. Budi has a different approach in the assessment design and implementation. He often refers to general and specific learning objectives stipulated in the curriculum. He said that if we can achieve learning objectives with only practice, then we design and implement the assessments with practice. However, given the limitations in supporting facilities, adjustments to practice and assessment models also become important consideration.

"For example, when I was teaching material about biotechnology. The biotechnology practice there was superior to conventional Tempe biotechnology. But when I found out that the practical material was not available, I made it into practice for making simple organic fertilizer. To make organic fertilizer, we have the tools and materials. In essence, as long as the material both applies biotechnology....."

Source: Interview transcript with Mr. Budi

Meanwhile, if the goal of the learning is to master cognitive material, we design and implement the assessment with cognitive questions. Mr. Budi further mentions that.

"in the Merdeka curriculum students are essentially expected to produce works, such as a project in accordance with the objectives of the learning material. Even though the project results are not as good as expected, since it is a student project. The problem with good assessment results is number 3 or 4, the most important thing is that students perform better."

Source: Interview transcript with Mr. Budi

In contrast to the two responders mentioned before, Mr. Ujang's approach to planning and administering assessments considers the students' existing abilities, to some extent referring to curriculum. This is his primary reason. He noted that certain students may require extra specialized guidance prior to an assessment, despite having a high IQ.

However, they may still have deficiencies in terms of systematic understanding. So, his main consideration is on the student's capabilities.

4.1.4. Teachers' Constraints and Opportunities

Using the same data analysis procedure in analyzing the observation notes and interview transcripts, I found four main identified categories or themes. They are students and teachers' readiness, learning resources, involvement of stakeholders, and aspiration. These themes are synthesized and described in the following.

Teacher Constraints

The MA science teachers are not constrained by the latest curriculum revisions. Indeed, it provides them with concise chapters - fewer chapters and more pragmatic learning techniques. However, analysis of their interview responses reveals that all teachers encounter comparable limitations when it comes to the preparedness of their students and the available learning resources. Regarding the preparedness of the students, teachers have seen that a majority of the students exhibit a lack of motivation towards studying. Mr. Rozi, who also holds the position of vice-principal of student affairs, explained that he frequently contacts or visits some student-parents to inform them about their children's irregular attendance or frequent truancy. He also mentioned that there are a few students who show little desire in returning to school. According to Mr. Ujang, the students display greater enthusiasm when welcomed to play rather than when asked to study.

The findings from the six classroom observations validate the assertions made during the interviews. Within the classroom of the three science teachers, a significant number of students are frequently absent during school hours. Consider, for instance, Mr.

Budi's classes. During the initial observation, it was found that 18 out of 32 students were present in the class. In the subsequent visit to different class, just 9 out of 30 students were able to attend the class. Meanwhile, in Mr. Ujang's class, the situation was even worse when it was discovered in the second observation that only 2 out of the 28 students were in attendance, while the remaining 28 students stayed in the dormitories for reasons. In addition to absenteeism, teachers have reported that locating sleeping students during teaching sessions poses another obstacle to implementing assessments. During the classroom observation, when the teachers monitor the class, they frequently find the students are sleeping, then the teachers wake them up from sleep and instruct them to wash their faces. According to Mr. Ujang, students napping in class is a result of staying up late at night and engaging in physical activities. This condition has remained the same for a couple of years.

Besides the issue of students' readiness, the teachers admitted that their subject mastery are fundamental as they start to design the assessment questions. Mr. Rozi found himself experiencing such a difficult condition when trying to create high order thinking questions as he attended curriculum and assessment training. Further, an availability of supporting learning facilities becomes the other constraint to implement effective assessments. Occasionally, Mr. Budi encounters a difficult situation when he needs to design an assessment instrument in which the laboratory lacks the necessary tools and supplies for students to utilize during practical sessions. If he is required to obtain the essential equipment, a substantial amount of money is needed, and schools may lack the financial resources to pay for it. In addition, when asking students to buy outside, they are prohibited from leaving the dormitory outside of the authorized schedule, even if they need to ask someone to purchase goods for them. Therefore, one alternative he usually

does is to change the practice and assessment model, remaining the same in learning achievement targets

Teacher Opportunity

Despite the challenges, all teachers have passed the teacher professional program and deserved title as professional teachers. A regular training program is annually provided by the school where experts in the education field - the school supervisors from provincial Education Office are invited as the speakers. The program which is annually conducted aims at developing teachers' capacity in terms of teaching and assessment. The program is usually conducted before the start of the new school year. Mr. Rozi narrated that

“Usually, when preparing the questions, we always present the supervisors from the education office, For example, today, we focus on making questions in biology subjects. Here, we check everything from the beginning, from determining the KD, then this, and so on.”

Source: Interview transcript with Mr. Rozi

In other words, Mr. Rozi affirms that the activity aims not only to guide teachers in the preparation of making questions but also to keep improving and at least remind us that we have already learned those skills to avoid just making up in question making.

Teachers have another option to enhance their teaching skills through online training offered by the government via the *Si-Pintar* program. An application developed jointly by the Ministry of Education and the Ministry of Religion to serve as a platform for teachers to acquire knowledge and insights pertaining to education. According to Mr. Budi,

“the presenters of the online training on the Si-Pintar application are highly qualified education professionals with PhD and professor degrees, as well as extensive experience in the field of education.”

Source: Interview transcript with Mr. Budi

Furthermore, Mr. Budi explains that the teachers are also administered a post-test at the conclusion of the session to assess teachers' level of proficiency in the online learning contents. Teachers who successfully achieve a score of 60 will be awarded a certificate signifying their completion of the program, whilst those who do not meet the minimum passing grade will have the opportunity to retake the test. Actually, participation in this application is optional. According to Mr. Rozi, teachers who have the time can make use of this opportunity.

Apart from the opportunities the teachers can participate in, they have aspirations for assessment practices. Mr. Ujang expects the school or government to provide at least a virtual laboratory or assessment-based technology in which students can practice the subject of physics. Take, for instance, how to apply a computer line. He also hopes the assessment can be generated through test applications provided by Google. Using technology to assess students' performance is for time efficiency and effectiveness. In contrast to Mr. Ujang, Mr. Budi hopes there will be a special allocation to focus on question-making training because the training time provided by the school is deemed insufficient, especially in practices of question-making sessions at intermediate and difficult levels, such as higher-order thinking skill questions. Our time at Madrasah is not enough. He mentioned,

"Our time at this Madrasah is sometimes very short; we have just finished teaching and taken care of this and that. We haven't finished this yet; we still care for other things. So, in my opinion, we need more time."

Source: Interview transcript with Mr. Budi

4.2. Discussion

The subsequent explanations discuss the findings, which are elaborated in the previous section. It deals with the synthesis of the findings with the theoretical underpinnings.

4.2.1. Discussion on RQ1: Teachers' Understanding on the Assessment

As we already acknowledged, assessment literacy (AL) is understanding, abilities, attitudes which should be possessed by teachers to accurately assess students' performance (Stiggins, 1991; Xu & Brown, 2016). Regarding the teachers' assessment literacy, I synthesize that their conceptual knowledge needs improving. To some extent the teachers have generally met all seven standards of the AL as proposed by Deluca et al. (2016); however, they have not fully integrated the principles of the assessment based on the Indonesia's education curriculum which include *assessment as, of, and for learning* (Anggraena et al., 2022) in their classrooms. This underlines the need for a more holistic approach to assessment where different types of assessment are strategically applied to increase student learning. In addition, the variation in selection of assessment types stipulated in the curriculum preferred learning theories; constructivism, socio-cognitive, behaviorism. I opine that adoption of the specific learning theory either consciously or subconsciously reflects the ways of the instructions are given in the classroom teaching. The instruction is correlated to the assessment (Rustaman, 2017).

When I further synthesize the findings, I would argue that the science teachers' assessment literacy in West Lombok, to some extent, is relatively good if I refer to Stiggins' (1991) assessment framework. This considerable understanding on the assessment literacy is highly likely because the teachers have degrees in teacher education

which supported by long teaching experience, have attended school regular trainings in each early semester, and have participated in online training through *Si-Pintar* application. This is in accordance with the research findings conducted by Yan et al. (2022), Bourke et al. (2018), Alkharusi et al. (2014), Girgla et al. (2021), and Paramo (2008) regarding the factors which affect the teacher's expertise in assessment literacy. They argue that a teacher's educational degree, years of experience, self-efficacy, and school support have significant impacts towards teachers' understanding on the concepts of assessment literacy. Clark et al. (2022) also asserted that changing views of teachers regarding the assessment practice is affected by their teaching experience. Furthermore, it is also in line with the research finding of Istiwaningsih et al. (2022) who revealed that training on designing e-assessment (e-students' worksheet) improves junior high school teachers' understanding of the assessment literacy.

However, when dealing with assessment principles, I would argue that two teachers have not applied all the principles in designing the students' assessment, if referring to assessment principle or approach in Indonesia's education context (MoREC, 2023). This is depicted from the findings in the documents, the lesson planning they created and responses in the interview. Anggraena et al. (2022) has also acknowledged that the three principles (*assessment as, for, of learning*) should not be separated from learning. If the principles are separated, this means that these assessments are not one unit or intertwined elements in Indonesia's education system. Meanwhile, all teachers including science teachers are expected by the governments to apply the *assessments for, as, and of learnings* during their teachings; however, only one teacher actively participates students in the *assessment as learning* in form of peer-assessment. Indeed,

the teachers who adhere to the principles of effective assessment have the knowledge and skills required to fulfil benchmarks of excellence (Stiggins, 1991).

Regarding the types of assessments which teachers could apply in their classroom teaching, I view that it depends on the theory of learning and the assessment principles adopted in the educational curriculum. Two teachers highly likely consider constructivism and socio-cognitive theories in their teaching and assessment, meanwhile the other one tries adopting two theories of learning: behaviorism and constructivism - with over usage of behaviorism. It can be noticed from how the teachers design their lesson planning, assessment practices in their classroom as well as the types of final semester question types that are generated to their students. Constructivism explicates that learning as the process by which learners construct their own understanding while behaviorism views the learning is a process of creating connection between what they experience which are related reinforcements and punishment (Schunk, 2012). On the other hand, socio-cognitive theory asserts that learning is an activity to process information where both behavioral structures and environmental events are developed to form symbolic representation as an action reference (Bandura, 1986, cited in Schunk, 2011).

In my view, teachers who are assessment literate should conduct all the assessment principles and approaches (Anggeraina et al, 2022). They should conduct assessments in the three-time frames: namely pre-instruction assessment, during the instruction (formative assessment) and post instruction (summative assessment) (Santrock, 2018). Unfortunately, the finding shows that only one teacher applies assessments in pre, during and post instruction. Meanwhile the other two teachers apply

only during and post assessment approaches. If referring to Santrock (2018), the two teachers might risk the class in either the instruction or the materials which are too advanced to their students if not conducting the pre-instruction assessment or diagnostic tests. This pre-instruction assessment can also be done by looking at prior students' knowledge in their portfolios or observing them for several days.

Furthermore, in terms of understanding in designing the levels of questions, two teachers speculated that they could create three levels of test-questions including high order thinking (HOT) tests. On the one hand, I would argue that their levels of the assessment literacy of all teachers are at the second level (functional users) if referring to classification proposed by Stiggins (1991). Stiggins argues that at this level, the individuals possess knowledge regarding four essential characteristics of high-quality data, namely a well-defined objective, a suitable sample, identifiable sources of potential bias, and actionable outcome. I also argue that the teacher is likely capable of levelling up his level to advance literate if referring to what trainings that have attended annually in which they are supervised by the experts from the Education Office.

In the advance literate (third level), the teachers should be accountable for the development of "high stakes" assessments, which are used to certify the minimum level of competence, select students for specialized programs, measure achievement on a large scale, and evaluate instructional programs (Stiggins, 1991). However, the synthesis of the finding from two participants admitted that reaching the third level is somewhat challenging since they have no sufficient time for practice on question making as well as conceptual understanding on the major field is highly needed. This is line with what Xu

and Brown (2016) mentioned regarding the mastery of subjects becomes the foundation of being able to create quality questions, apart from being literate in assessment concepts.

4.2.2. Discussion on RQ2: Teachers' Classroom Assessment Practice

In the level of teachers' classroom assessment practice, only formative assessment can be obtained during the direct classroom observations. Checking knowledge of the students regarding the previous topic meeting becomes the first main activity done by the teachers which is followed by the current topic; however, I argue that the ways of the teachers in reviewing the previous knowledge may vary depending on the concepts of the learning the teachers hold. In socio-cognitivism, teachers use verbal or visual representations to enhance students' ability in recalling information (Schunk, 2011; Santrock, 2018). Two teachers engage the students' previous knowledge by showing them pictures or videos on the screen. In contrast, the other one teacher might prefer to the theory of constructivism in which the students are engaged in the process of actively constructing meaning by establishing connections between new material and their preexisting knowledge (James, 2008; Schunk, 2012).

I contend that this conception is also similarly adopted when checking students understanding of current course subjects. Take for instance, the Chemistry teacher groping the students and instructing them to discuss the materials and an issue. During the discussion the teacher gives direct feedback for each group, and after one group finishes the presentation in front of the class, other group students are also instructed to give feedback, a peer-assessment. According to Schunk (2011) and Santrock (2012), this model of learning is called inactive learning where students acquire knowledge through experiencing the consequences of their own actions. While the other two teachers refer to

notion that meaningful learning in complicated skills occurs when students can watch the materials along with given samples, and then do the given task (Santrock, 2018).

In my view, the way teachers assess student knowledge during classroom teaching reveals the underlying learning philosophies. In the socio-cognitivism, the teachers use visuals to enhance recall on the subjects (Schunck; Santrock, 2018), while constructivists encourage active meaning-making through discussions and peer-feedback (James, 2008; Schunk, 2012). This is further evident in how teachers check current understanding after teachers have given the course contents. The Chemistry teacher utilizes a “learning by doing” approach with group discussions and peer-assessment (Schunk, 2011; Santrock, 2018), while the others rely on visual and teacher-lead questioning, potentially reflecting a “watch-then-do” methods for complex skills (Santrock, 2018).

4.2.3. Discussion on RQ3: Designing and Implementing Assessments.

As far as I am concerned, the concept of learning on the current curriculum, the Merdeka Curriculum is inquiry-based learning in which students are expected to study and assess in accordance with this learning concept. To some extent, I would argue that all the teachers design the formative assessment based on the seven standards proposed by Mertler (2004) and Deluca et al. (2016), but they sometimes consider the constraint like students’ current knowledge which is not matched with the target of the learning. This is clearly stated by Mr Ujang who argues that *“sometimes we cannot apply for this test since the students’ knowledge is far behind the target”* (Source: interview response). Indeed, the principle of assessment says that that assessments are created and implemented in alignment with the assessment goals, allowing for flexibility in choosing

the method and timing of the assessment to effectively meet the learning objectives (Anggraena et al., 2022).

Moreover, the finding from Mr. Rozi also confirms that factors leading teachers in assigning assessment could be based on students' preference; while Mr. Budi argues heterogeneity of the students are reasons of creating the types of the assessment question. I would note that these views are in accordance with the fourth standard in assessment literacy in which teacher design the measurement types considering the assessment fairness that is being aware of the students' diversity (Mertler, 2004; Deluca, 2016). Otherwise, these assessments can demotivate students (Mertler, 2004).

I believe that capacity to modify assessment to suit field setting is also essential, as unforeseen circumstances may arise. So, the adaptability is the other consideration in designing the assessment measurements, particularly in practice. This is found in Mr. Budi issues when he would like to assess his students in the practice of biotechnology - making conventional Tempe, but finding tools and materials for the practice are not available in the laboratory. Alternatively, he alters the types of practice into simple organic fertilizers. This adaptability is in line with the second standard of teacher assessment proposed by Deluca et al. (2016) who asserted that conducting assessment process the teachers should know how to create, administer, score, and interpret assessment and results. Referring to Mr. Budi response that he prefers to use formative assessment over the summative. This is probably due that his learning theory constructivism where he views that something is missing from Biology learning if students have no practice.

However, in term of the summative assessment used in the final semester test, the teachers are ready enough with types of question they design. From the findings I am in convinced that consideration in selecting the question-types is based on the learning objectives as stated in the curriculum and assessment guideline from the central government, which is also in line with the assessment principle, that is learning objective accordance (Anggaraena et al., (2022). The guideline directs all teachers to create (at least three) types of questions from five proposed question-types. Apart from familiarizing students for national assessment, the inclusions of the questions are accordance with target oriented and concepts of assessment adaptability and flexibility to meet students' diverse backgrounds (Girgla et al., 2021). More importantly, the students must be well informed and prepared for such assessments to perform well during the final assessment.

4.2.4. Discussion on RQ4: Teachers' Challenges and Opportunities

Teachers' Challenges

It seems to be true having sufficient level of assessment literacy does not guarantee teachers can implement effective assessment practices. Even though all teachers are equipped with knowledge strengthened by the annual training in advance of early semester teaching, the finding revealed two main challenges: the students' readiness and source of learning. In my view, students are likely not ready and have low motivation with the learnings. These can be synthesized from high numbers of absenteeism in each observed classes together with the finding where some students are sleeping during the study hours.

Students with such conditions cannot be involved for summative assessment. Indeed, the involvement of students in assessment is essential since it is half of the

learning (Girgla et al., 2021). Also, Xu and Brown (2016) elucidated that not only teachers know how assessments can excite but also students need to involve the students in the assessment practices. If this continuous to happen, it could have significant impact on the learning outcomes, as we are noticed that students' learning success is determined by the students themselves and the teachers (Hattie, 2003). It means that both students and teachers shared responsibility in achieving the success of the learning objectives.

Regarding the issue of lack of tools and materials for practical assessments provided by the school, to some degree it could be a serious issue for teachers if they have no strong foundation of assessment literacy (Girgla, 2021). However, assessment literate teachers should have alternatives in advance if their teaching and assessment require specific needs. This has been shown by Mr. Budi who remains in the objective of assessment in Biology lesson he teaches (Rustaman, 2007) by strategically altered ways of assessing students, despite utilizing distinctive objects (Santrock, 2018).

Teachers' Opportunity

I see that teachers have great potential to develop their assessment literacy level. It seems that they can reach the high advanced level of assessment literate (Stiggins, 1991) or as an assessor identity level (Xu & Brown, 2016) if they are consistent in developing, attending, and participating in the provided assessment trainings. The annual school training given by educational experts from the regional education office has been conducted with additional focus on assessment literacy training for the last few years. Teachers are not only guided regarding conceptual understanding, but also practical skill in designing the question-types (low, medium, high level).

Besides the annual training, a similar training is also generated by the central government through an online platform - *Si-Pintar* Application. I opine that the focus of this online training is wider than that of the previous PPG training attended by the teachers. It is noted that the PPG curriculum does not place significant focus on the testing and evaluation abilities of teachers (Sulistiyowati et al., 2021). As narrated by the three teachers that there are lot of materials which can be used for teaching improvements provided in the application both in teaching and assessment. However, this online training is not mandatory now. Teachers themselves should have an initiative to develop their expertise in their field of study; otherwise, they cannot cope with the latest innovation and global demands, and worse be left behind in education.

CHAPTER V

CONCLUSION AND RECOMMENDATION

In this chapter, the researcher summarizes and reviews the study's findings and what contribution it offers to assessment knowledge and practice of the assessment which is related to the current curriculum implementation. Strengths and limitations which are pertinent to the research design and process as well as potential implication pedagogical and assessment practice are believed intertwined competence. Following these, the recommendations for further research.

5.1. Conclusion

5.1.1. Summary of the Findings

The result of this study has revealed that the MA science teachers have favourable concepts or understanding regarding assessment literacy. They acknowledge the essence of cognitive and holistic assessment in evaluating their students' learning activities. All teachers are obliged to attend annual training and professional development on teaching and the assessment from the school. They learn lot about methods and model of assessment practices driven by educational experts from the regional education office. They argue that the current curriculum – the Merdeka Curriculum is highly valued for its focus on project-based learning and assessment, compared to the Curriculum 2013. However, the teacher admitted their weaknesses to be able to design high quality assessments, which they think are still far behind the national level.

Subsequently, they consider collaborative actions amongst teachers are needed to develop assessment techniques. One of the actions is thorough the conduct of regular discussions with fellow teachers in the same field. This discussion allows teachers to obtain the latest innovation in education practices. Another constructive action which the

teachers can take pivotal roles is by being annual Olympiad test makers. They could sharpen their understanding and practical skills on assessment literacy which they have had learnt.

Regarding the types of assessment, the teachers have used both oral and written forms of assessment for formative purposes. The oral mode of assessment is used to give direct feedback, while the written format is designed to test the students' knowledge and comprehension of the current subject matters. The biology teacher highly endorses the project-based assessment in his assessment practices since he argues that it is more efficacious method. However, for the summative assessment, all teachers follow the government directives in which the teachers should employ multiple-choice, short answer, and descriptive types of questions in the final semester test. Moreover, the directives oblige the teachers to categorise the test questions in varying levels of difficulty, from low to higher order thinking skill questions.

In classroom assessment practices, the teachers are found to assess their student's previous knowledge through applying different ways of approaches. Direct oral questioning, assignment submission, video presentations or past project presentations are among the ways each teacher assesses students' prior knowledge. Only one teacher uses a variety of approaches, while the others do not. It is also noted that the teacher assesses students' comprehension through questioning, peer-assessment, and assignments. These approaches are called the assessment for learning. In addition, all the teachers were observed overseeing learners' levels of participation; of which is sleeping during the learning.

Shifting to factors influencing teachers in selecting the types of assessments' approaches. It is found that there are three factors, namely, government regulation, teachers' conceptual understanding on the assessment literacy, and students' attributes. Teachers agree to opt three out five types of questions proposed by the governments; however, for classroom assessment, the teacher has diverse preferences. The chemistry teacher strategically uses complex and multiple types of assessment to assess some aspects of learners' cognitive and affective abilities. In the beginning of the semester, he starts to assess students' current cognitive competency in which the results he uses as a guide for the formulation of later assessments, as well as finding out types of students' evaluation preferences. He opts to integrate practical assessments to accommodate students who prefer hands-on learning activities, besides incorporating portfolios and activity reports.

Mr. Budi's assessment design is closely in line with the curriculum objectives. He placed practical assessments as a higher priority whenever possible, but recognizes the restrictions imposed by a lack of resources. His versatility is seen in his capacity to modify assessment tasks to correspond with the materials at hand, as exemplified by his transition from biotechnology to organic fertilizer production. Mr. Budi also underscores the significance of student-centered learning, emphasizing the Merdeka's curriculum's emphasis on student projects and the necessity of prioritizing students' development over assessment results. Unlike rigorous adherence to the curriculum, Mr. Ujang's assessment planning is mostly guided by learners' previous knowledge. He acknowledges the necessity of providing personalized assistance to kids who may have difficulties, even when they possess excellent cognitive abilities. This method emphasizes the significance

of considering the varied needs and learning styles of students when creating assessments and evaluations.

This research finding also uncovered the limitations and possibilities encountered by the MA science teachers when implementing assessments. A major limitation faced by teachers was the level of students' learning which can be noticed from frequent absenteeism. Teachers encountered difficulties in captivating uninterested students and modifying instruction to suit different degrees of preparedness. In addition, teachers have raised concerns in creating assessment questions that require higher-order thinking skills. Apart from the above-mentioned, the presence of inadequate learning resources presented another obstacle, such as insufficient laboratory equipment which impeded practical assessments, while financial limitations prohibited the acquisition of essential supplies.

Despite facing difficulties, the science teachers had opportunities for professional development. The first is through the annual training programs led by education experts, with a focus on enhancing teaching and assessment skills. Another teacher professional support is through *Si-Pintar* online platform, supported by the central government. The teachers are provided with numerous course materials and practices which are not limited to only teaching and assessments. Teachers indicated their aspirations for enhanced assessment techniques. They promoted the idea of expanding access to technology, such as virtual laboratories and automated evaluation tools, to improve efficiency and efficacy. In addition, the teachers desired additional focused time for the development of high-quality assessment questions, namely those that target higher-order thinking skills.

5.1.2. Concluding Statement

This research has investigated the teachers' assessment literacy in one of the districts in West Nusa Tenggara, Indonesia. Findings have revealed the current teachers

understanding on the concept of assessment literacy, the assessment practice, factors in designing and implementing assessment, and the challenge and opportunity it offers. Looking at the teacher's conception of the assessment's literacy, though not complete senses, majority of the teachers understand what assessment literacy means. They all are aware of the components of good assessments for upgrading their pedagogical skills and enhancing the students' performances, starting from familiarizing themselves with the types, levels, procedures, and principles of the assessment literacy. The teachers' conceptions are regularly updated by the school, at least once a year. The government also has facilitated teachers for all levels of education to attend online training through an application called *Si-Pintar*. Since it is not compulsory, not all teachers allocate their time for upgrading their knowledge either in mainly assessment and or pedagogical competence.

In selecting the types and levels of assessment, either for formative or summative assessment, the teachers have several considerations. Not only do they need to pay attention to the general and specific objectives of the learning, which is stipulated in the Merdeka curriculum, but also think of the current students' knowledge. More importantly, the assessment classroom practices yield teachers' views of the theory of learning. Two of whom refer to constructivism and socio-cognitive theories, while the other one relies heavily on behaviorism with little adaptation to constructivism. The Merdeka curriculum approach is an enquiry-based learning (Aggaraena et al., 2023). This means that the teachers are expected to conduct the teaching based on the approach in which students are actively involved in learning and assessment, particularly in formative assessment.

During the teaching and assessment practices, teachers are faced with the reality that students' readiness and incomplete learning facilities. These two hindrances significantly affect their teachings. On the one hand, they want to improve their students learning through improving their capacity as teachers, however, low motivations depicted from the high percentage of students' absence along with sleeping during the school hours. Facilities to support the practices on certain course materials are unavailable. Not allowing students to go outside of the school in unscheduled time as well as lack of school fundings become the other challenges for teachers in appropriately apply the assessments. Fortunately, with some alternatives the assessment practices sometimes can be actualized.

5.2. Recommendation

Findings of this research have implication for teachers in practical and for research in theoretical. Since assessment literacy is a part of pedagogical competence which teachers need to possess, teacher education and other program for professional development courses are expected to allocate the opportunity for either pre-service, novice and in-service teachers to develop their expertise in the assessment knowledge. These can be the developments in standards, principles, approach, or methods of assessment literacy. Not only on this theoretical development but ensuring their assessment practices in line with all related teachings such curriculum and current students levels become pivotal for better learning performance results. In addition, this research is one of the efforts to provide additional sources of information majoring in science teacher assessment literacy in the context of current teaching practice design by Kemdikbud. In addition, conceptual knowledge in the field is highly like affects their assessment practices. Hence, both prospective teachers and in-service need to upgrade

their conceptual knowledge in their field and knowledge and practices of assessment and evaluation for the sake of better education attainment.

A qualitative research design is believe an appropriate way to more accurately reflects the present condition of the teachers since it focuses on specific individuals or groups and highlights key events that are important to the case (Qi, 2009) and is characterized by its precise, detailed, and analytical traits to deeply explore the personal experiences of an individual teacher/s or a small group (Cohen et al., 2018). However, the opponents argue that qualitative research is lack of reliability, validity, and generalizability (Sharma, 2010) since its aim not to generalize the finding, but to comprehend a particular issue. So, for further research, it is highly suggested to add more collected data instruments, such as evaluated students' worksheets and final test-questions, and then use triangulation data analyses in the research design which could increase the trustworthiness of the findings.

REFERENCES

- Alkharusi, H., Aldhafri, S., Alnabhani, H., & Alkalbani, M. (2014). Classroom Assessment: Teacher Practices, Student Perceptions, and Academic Self-efficacy Beliefs. *Social Behavior and Personality: an international journal*, 42(5), 835-855.
- Aldini, U. (2020). Pengelompokan Provinsi di Indonesia Berdasarkan Indikator Mutu Pendidikan Sekolah Menengah Pertama Tahun 2016–2018 Menggunakan Model Based Clustering (Doctoral dissertation, Universities PGRI Adibuana Surabaya).
- Anggraena, Y., Ginanto, D., Felicia, N., Andiarti, A., Herutami Indriyanti, H., Alhapip, L., & Setiyowati, D. (2022). Panduan Pembelajaran dan Asesmen Kurikulum 2013: Pendidikan Anak Usia Dini, Pendidikan Dasar, dan Menengah.
- Asiyah, S., Wiyono, B. B., Hidayah, N., & Supriyanto, A. (2021). The Effect of Professional Development, Innovative Work and Work Commitment on Quality of Teacher Learning in Elementary Schools of Indonesia. *Eurasian Journal of Educational Research*, (95), 227-246.
- Akhtar, Z., Hussain, S., & Ahmad, N. (2021). Assessment Literacy of Prospective Teachers in Distance Mode of Education: A Case Study of Allama Iqbal Open University, Islamabad. *Journal of Education and Educational Development*, 8(1), 218-234.
- Aria, D., Sukyadi, D., & Kurniawan, E. (2021). Teacher Assessment Literacy: Indonesian EFL secondary Teacher's Self-Perceived on Classroom-based Assessment Practice. *English Review: Journal of English Education*, 10(1), 15-26.
- Babaii, E., & Asadnia, F. (2019). A Long Walk to Language Assessment Literacy: EFL Teachers' Reflection on Language Assessment Research and Practice. *Reflective Practice*, 20(6), 745-760. DOI:<https://doi.org/10.1080/14623943.2019.1688779>
- Bahtiar, I., & Purnawarman, P. (2020, December). Investigating English teachers' comprehension in Language Assessment Literacy (LAL). In *1st International Conference on Information Technology and Education (ICITE 2020)* (pp. 303-310). Atlantis Press.
- Baker, B. A., & Riches, C. (2018). The Development of EFL Examinations in Haiti: Collaboration and Language Assessment Literacy Development. *Language Testing*, 35(4), 557-581. DOI: <https://doi.org/10.1177/026553221771>

- Baird, M. J. C. (2021). Exploring Teachers' Perceptions of Assessment Literacy in a High Performing School District in New Jersey (Doctoral dissertation, Seton Hall University).
- Bandele, S. O., & Oluwatayo, J. A. (2013). Assessing Assessment Literacy of Science Teachers in Public Secondary Schools in Ekiti State. *Journal of Education and Practice*, 4(28), 56-63.
- Bøhn, H., & Tsagari, D. (2021). Teacher Educators' Conceptions of Language Assessment Literacy in Norway. *Journal of Language Teaching and Research*, 12(2), 222-233. DOI: <http://dx.doi.org/10.17507/jltr.1202.02>
- Bourke, R., O'Neill, J., & Loveridge, J. (2018). What Starts to Happen to Assessment When Teachers Learn about Their Children's Informal Learning?. *The Australian Educational Researcher*, 45, 33-50.
- Budiyono, B. (2022). Assessment Literacy in PPG Batch 2020. *Magister Scientiae*, 49(2), 139-145.
- Barnes, N., Fives, H., & Dacey, C. M. (2014). Teachers' Beliefs about Assessment. *International handbook of research on teachers' beliefs* (pp. 284-300). Routledge.
- Celebre-Baird, M., & Reid, D. B. (2023). Exploring Teachers' Perceptions of Assessment Literacy in a High-performing School District in New Jersey. *Teachers and Teaching*, 1-17.
- Christoforidou, M., & Kyriakides, L. (2021). Developing Teacher Assessment Skills: The Impact of the Dynamic Approach to Teacher Professional Development. *Studies in Educational Evaluation*, 70, 101051.
- Clark, A. K., Nash, B., & Karvonen, M. (2022). Teacher Assessment Literacy: Implications for Diagnostic Assessment Systems. *Applied Measurement in Education*, 35(1), 17-32.
- Coombs, A., DeLuca, C., LaPointe-McEwan, D., & Chalas, A. (2018). Changing Approaches to Classroom Assessment: An Empirical Study across Teacher Career Stages. *Teaching and Teacher Education*, 71, 134-144.
- Coombs, A., Rickey, N., DeLuca, C., & Liu, S. (2021). Chinese Teachers' Approaches to Classroom Assessment. *Educational Research for Policy and Practice*, 1-18.
- Cohen, L., Manion, L., & Morrison, K. (2018). Research Methods in Education (eight edition). *Abingdon, Oxon*, 532-533.

- Cowie, B., Cooper, B., & Ussher, B. (2014). Developing an Identity as a Teacher-Assessor: Three student teacher case studies. *Assessment Matters*, 7(6), 64-89.
- Daud, A., Novitri, N., & Hardian, M. (2020). Evaluation of the Teacher Professional Education Program (PPG): English Students' Perspective. *AL-ISHLAH: Jurnal Pendidikan*, 12(2), 185-200.
- DeLuca, C., LaPointe-McEwan, D., & Luhanga, U. (2016). Approaches to Classroom Assessment Inventory: A New Instrument to Support Teacher Assessment Literacy. *Educational Assessment*, 21(4), 248-266.
- DeLuca, C., Lapointe-Mcewan, D., & Luhanga, U. (2016). Teacher Assessment Literacy: A Review of International Standards and Measures. *Educational Assessment, Evaluation and Accountability*, 28, 251-272.
- DeLuca, C., Willis, J., Cowie, B., Harrison, C., Coombs, A., Gibson, A., & Trask, S. (2019, November). Policies, Programs, and Practices: Exploring the Complex Dynamics of Assessment Education in Teacher Education across Four Countries. In *Frontiers in Education* (Vol. 4, p. 132). Frontiers Media SA.
- Deneen, C. C., & Brown, G. T. (2016). The Impact of Conceptions of Assessment on Assessment Literacy in A Teacher Education Program. *Cogent Education*, 3(1), 1225380.
- Dikbud Provinsi Nusa Tenggara Barat*. (n.d).
<https://dikbud.ntbprov.go.id/index.php/Beritaaa/Detail/438>
- Edwards, F. (2019). A Rubric to Track the Development of Secondary Pre-Service and Novice Teachers' Summative Assessment Literacy. In *Developing Teachers' Assessment Capacity* (pp. 85-107). Routledge.
- Fahrnunisa, F. (2021, October). Ini Daftar 100 SMA Terbaik Berdasarkan Hasil UN 2019. LIPUTAN 6. Retrieved from:
<https://www.liputan6.com/news/read/3993360/ini-daftar-100-sma-terbaik-berdasarkan-hasil-un-2019?page=9>
- Firoozi, T., Razavipour, K., & Ahmadi, A. (2019). The Language Assessment Literacy Needs of Iranian EFL Teachers with a Focus on Reformed Assessment Policies. *Language Testing in Asia*, 9(1), 1-14.
- Gareis, C. R., & Grant, L. W. (2015). Assessment Literacy for Teacher Candidates: A Focused Approach. *Teacher Educators' Journal*, 2015, 4-21.

- Giraldo, F., & Murcia, D. (2018). Language Assessment Literacy for Pre-Service Teachers: Course Expectations from Different Stakeholders. *GiST Education and Learning Research Journal*, 16, 56-77. Retrieved from: <https://files.eric.ed.gov/fulltext/EJ1184916.pdf>
- Giraldo, F. (2019). Language Assessment Practices and Beliefs: Implications for Language Assessment Literacy. *How*, 26(1), 35-61. <https://doi.org/10.19183/how.26.1.481>
- Girgla, A., Good, L., Krstic, S., McGinley, B., Richardson, S., Sneider-Gregory, S. and Star, J. (2021). *Developing a Teachers' Assessment Literacy and Design Competence Framework*. Australian Council for Educational Research. <https://ibo.org/globalassets/publications/ib-research/assessment-literacy-final-report-en.pdf>
- Gotch, C. M., & McLean, C. (2019). Teacher Outcomes from A Statewide Initiative to Build Assessment Literacy. *Studies in Educational Evaluation*, 62, 30-36.
- Gunn, A. C., & Gilmore, A. (2014). Early Childhood Initial Teacher Education Students' Learning about Assessment. *Assessment Matters*, 7(2).
- Hailaya, W., Alagumalai, S., & Ben, F. (2014). Examining the Utility of Assessment Literacy Inventory and Its Portability to Education Systems in the Asia Pacific Region. *Australian Journal of Education*, 58(3), 297-317.
- Hardiyanto, S. (2022, August 26). *Daftar Terbaru top 1.000 Sekolah Berdasarkan nilai utbk 2022, Cek di top-1000-sekolah.ltmpt.ac.id*. KOMPAS.com. <https://www.kompas.com/tren/read/2022/08/26/183000865/daftar-terbaru-top-1.000-sekolah-berdasarkan-nilai-utbk-2022-cek-di-top>
- Hattie, J. A., & Brown, G. T. (2010). Assessment and Evaluation. In *Educational Psychology: Concepts, Research, and Challenges* (pp. 116-131). Routledge.
- HUSSEN, A. A., & TESHOME, T. Z. (2016). Doğu Etiyopya'daki Orta Dereceli Okullarda Öğretmenlerin Öğrenci Öğrenmesine, Mesleklerine ve Topluluğa Yönelik Bağlılıkları. *Journal of Teacher Education and Educators*, 5(3), 289-314.
- Ingarianti, T. M., Suhariadi, F., Fajrianti, F., & Kristiana, I. F. (2022). The Effect of Antecedents of Teachers' Subjective Career Success. *International Journal of Environmental Research and Public Health*, 19(17), 11121.

- Ipah, Budi, Minarti., Atip, Nurwahyunani., Dewi, Purichasari. (2023). Assessment Literacy Profile of Science Teachers in SMAN 1 Gubug. *International Journal of Research in Education*. DOI: 10.26877/ijre.v3i1.13558
- James, M. (2006). Assessment, Teaching and Theories of Learning. *Assessment and learning*, 47, 60.
- Janatifar, M., & Marandi, S. S. (2018). Iranian EFL Teachers' Language Assessment Literacy (LAL) under an Assessing Lens. *Applied Research on English Language*, 7(3), 361-382. DOI: [10.22108/ARE.2018.108818.1221](https://doi.org/10.22108/ARE.2018.108818.1221)
- Jumlah Data Satuan Pendidikan (Dikmen) Per Provinsi. (n.d.). Kemristek Dikti. Retrieved May 11, 2024, from <https://referensi.data.kemdikbud.go.id/pendidikan/dikmen>
- Indonesia, B. P. S. (n.d.). *Jumlah Sekolah, Guru, dan Murid Madrasah Aliyah (MA) di Bawah Kementerian Agama Menurut Provinsi, 2022/2023 - Tabel Statistik*. Badan Pusat Statistik Indonesia. <https://www.bps.go.id/id/statistics-table/3/VUUxWVltazBUbII1VG5veWNIbFliek5uYmtGSVp6MDkjMw==/jumlah-sekolah--guru--dan-murid-madrasah-aliyah--ma--di-bawah-kementerian-agama-menurut-provinsi--2021.html?year=2022>
- Kaden, U., & Patterson, P. P. (2014). Changing Assessment Practices of Teaching Candidates and Variables that Facilitate that Change. *Action in Teacher Education*, 36(5-6), 406-420.
- Kremmel, B., Eberharter, K., Holzknacht, F., & Konrad, E. (2018). Fostering Language Assessment Literacy through Teacher Involvement in High-Stakes Test Development. *Teacher Involvement in High-stakes Language Testing*, 173-194. DOI: https://doi.org/10.1007/978-3-319-77177-9_10
- Kusaeri, K. (2018). The portrait of Madrasah Aliyah in Indonesia: A Critical Evaluation of the Mathematics Score in the National Examination. In *2nd International Conference on Statistics, Mathematics, Teaching, and Research* (pp. 1-7). IOP Publishing Ltd.
- Kusumawardhani, P. N. (2017). Does Teacher Certification Program Lead to Better Quality Teachers? Evidence from Indonesia. *Education Economics*, 25(6), 590-618.

- Latif, M. W., & Wasim, A. (2022). Teacher Beliefs, Personal Theories and Conceptions of Assessment Literacy—A Tertiary EFL Perspective. *Language Testing in Asia*, 12(1), 11.
- Levi, T., & Inbar-Lourie, O. (2020). Assessment Literacy or Language Assessment Literacy: Learning from the teachers. *Language Assessment Quarterly*, 17(2), 168-182. DOI: <https://doi.org/10.1080/15434303.2019.1692347>
- Menyikapi Asesmen Nasional bagi guru Dan Sekolah. (n.d.). Retrieved from <https://gtk.kemdikbud.go.id/read-news/menyikapi-asesmen-nasional-bagi-guru-dan-sekolah>
- Mertler, C. A. (2004). Secondary Teachers' Assessment Literacy: Does Classroom Experience Make a Difference?. *American secondary education*, 49-64.
- Michigan Assessment Consortium. (2020). Assessment Literacy Standards: A National Imperative. Assessment Literacy Standards.
- Ningsih, A. W., Shara, A. M., Andriani, D., & Kisno, K. (2021). Reading Duration and Financial Literacy in Vocational High School Pematangsiantar. *Jurnal Education and Development*, 9(1), 577-583. <https://doi.org/10.37081/ed.v9i1.2288>
- Nyudak, I. B. G. (2022). Assessment Literacy of EFL Teachers in Badung, Bali, Indonesia: Conception and practices. *Assessment*, 9(1).
- Oecd. (2023, December 5). *Equity in education in Pisa 2022*. OECD iLibrary. https://www.oecd-ilibrary.org/education/pisa-2022-results-volume-i_03c74bdd-en
- Ölmezer-Öztürk, E., & AYDIN, B. (2019). Investigating Language Assessment Knowledge of EFL Teachers. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 34(3), 602-620. doi: 10.16986/HUJE.2018043465
- Popham, W. J. (2009). Assessment Literacy for Teachers: Faddish or Fundamental?. *Theory into practice*, 48(1), 4-11.
- Pusat Penilaian Pendidikan. (2019). *Laporan Hasil Ujian Nasional / Kementerian Pendidikan dan Kebudayaan*. Retrieved May 14, 2024, from https://hasilun.pusmenjar.kemdikbud.go.id/#2019!sma!daya_serap!23&99&999!a&04&T&T&1&unbk!1!&
- Riskan, Qadar., Zeni, Haryanto., Lambang, Subagiyo., M., A., Junus., Muliati, Syam. (2022). Indonesian Science Teachers' Ability to Design Scientific Literacy Test.

- International Journal of STEM Education for Sustainability, 2(1):133-139. doi: 10.53889/ijses.v2i1.52
- Qi, S. H. E. N. (2009). Case Study in Contemporary Educational Research. *Cross-cultural Communication*, 5(4).
- Rahayu Saputra, E., Abdul Hamied, F., & Suherdi, D. (2020). The Development of Beliefs and Practices of Language Assessment Literacy: Does A Professional Learning Community Help?. *Journal of Education for Teaching*, 46(3), 414-416. <https://doi.org/10.1080/02607476.2020.1761250>
- Rachmatullah, A., Roshayanti, F., & Ha, M. (2017). Validating the Assessment for Measuring Indonesian Secondary School Students Performance in Ecology. Paper presented at the Journal of Physics: Conference Series.
- Rapor Pendidikan Indonesia - kemdikbud. (n.d).
https://merdekabelajar.kemdikbud.go.id/upload/file/201_1649322280.pdf
- Razavipour, K., & Rezagah, K. (2018). Language Assessment in the New English Curriculum in Iran: Managerial, Institutional, and Professional Barriers. *Language Testing in Asia*, 8, 1-18. <https://doi.org/10.1186/s40468-018-0061-8>
- Rogler, D. (2014). Assessment Literacy: Building a Base for Better Teaching and Learning. In *English Teaching Forum* (Vol. 52, No. 3, pp. 2-13). US Department of State. Bureau of Educational and Cultural Affairs, Office of English Language Programs, SA-5, 2200 C Street NW 4th Floor, Washington, DC 20037.
- Ron. (2022, November 16). *Rapor Pendidikan Terbit Akhir Tahun, Potret Kinerja Sekolah Diharap Meningkat*. Suara NTB. Retrieved April 15, 2024, from <https://www.suarantb.com/2022/11/16/rapor-pendidikan-terbit-akhir-tahun-potret-kinerja-sekolah-diharap-meningkat/ir> Tahun, Potret Kinerja Sekolah Diharap Meningkat
- Rustaman, N. Y. (2017, September). Assessment in Science Education. In *Journal of Physics: Conference Series* (Vol. 895, No. 1, p. 012141). IOP Publishing.
- Schunk, D. H. (2012). Learning theories an educational perspective. Pearson Education, Inc.
- Shapovalovy, Y. A. & Evans, C. M. (2022). Research synthesis: Developing the Educator Assessment Literacy Professional Learning Screener. Dover, NH: The National Center for the Improvement of Educational Assessment, Inc.

- Serpil, Ö. Z., & Derin, A. T. A. Y. (2017). Turkish EFL Instructors' in-class Language Assessment Literacy: Perceptions and Practices. *ELT Research Journal*, 6(1), 25-44. DOI: <http://dergipark.ulakbim.gov.tr/eltrj/>
- Shah Ahmadi, M. R., & Ketabi, S. (2019). Features of Language Assessment Literacy in Iranian English Language Teachers' Perceptions and Practices. *Teaching English as a Second Language Quarterly (Formerly Journal of Teaching Language Skills)*, 38(1), 191-223. DOI: 10.22099/jtls.2020.34843.2739
- Sharma, S. (2010, July). Qualitative Methods in Statistics Education Research: Methodological Problems and Possible Solutions. In *Proceedings of the Eighth International Conference on Teaching Statistics* (Vol. 8).
- Shepard, L. A. (2000). The Role of Assessment in a Learning Culture. *Educational researcher*, 29(7), 4-14.
- Stiggins, R. J. (1991). Assessment Literacy. *Phi Delta Kappan*, 72(7), 534-539.
- Sudianto, S., & Kisno, K. (2021). Potret Kesiapan Guru Sekolah Dasar dan Manajemen Sekolah dalam Menghadapi Asesmen Nasional. *Jurnal Akuntabilitas Manajemen Pendidikan*, 9(1), 85-97. doi:<https://doi.org/10.21831/jamp.v9i1.39260>
- Sultana, N. (2019). Language Assessment Literacy: An Uncharted Area for the English Language Teachers in Bangladesh. *Language Testing in Asia*, 9(1), 1. <https://doi.org/10.1186/s40468-019-0077-8>
- Technology-driven education reform in Indonesia. (n.d.). <https://repositori.kemdikbud.go.id/30538/1/Indonesias-K-12-Education-Quality-Improvement-English-05122023.pdf>
- Thomas, D. R. (2003). A General Inductive Approach for Qualitative Data Analysis.
- Umami, I. (2018). Moderating Influence of Curriculum, Pedagogy, and Assessment Practices on Learning Outcomes in Indonesian Secondary Education. *Journal of Social Studies Education Research*, 9(1), 60-75.
- Vazquez-Lopez, V., & Huerta-Manzanilla, E. L. (2021). Factors Related with Underperformance in Reading Proficiency, the Case of the Program for International Student Assessment 2018. *European Journal of Investigation in Health, Psychology and Education*, 11(3), 813-828

- Wang, H., Sun, W., Zhou, Y., Li, T., & Zhou, P. (2022). Teachers' Assessment Literacy Improves Teaching Efficacy: A View from Conservation of Resources Theory. *Frontiers in Psychology, 13*, 1007830.
- Willis, J., Adie, L., & Klenowski, V. (2013). Conceptualizing Teachers' Assessment Literacies in an Era of Curriculum and Assessment Reform. *The Australian Educational Researcher, 40*, 241-256.
- Xu, Y., & Brown, G. T. (2016). Teacher Assessment Literacy in Practice: A Reconceptualization. *Teaching and Teacher Education, 58*, 149-162.
- Yan, X., & Fan, J. (2021). "Am I qualified to Be a Language Tester?": Understanding the Development of Language Assessment Literacy across Three Stakeholder Groups. *Language Testing, 38*(2), 219-246. DOI: <https://doi.org/10.1177/0265532220929>
- Yan, Z., & Pastore, S. (2022). Are Teachers Literate in Formative Assessment? The Development and Validation of the Teacher Formative Assessment Literacy Scale. *Studies in Educational Evaluation, 74*, 101183.
- Yan, Z., Chiu, M. M., & Cheng, E. C. K. (2022). Predicting Teachers' Formative Assessment Practices: Teacher Personal and Contextual Factors. *Teaching and Teacher Education, 114*, 103718.
- Ye, W. (2023). Urban Public-school Teachers' Assessment Literacy in China's Zhengzhou city. *Education 3-13, 51*(5), 836-849.
- Zhang, Z., & Burry-Stock, J. A. (2003). Classroom Assessment Practices and Teachers' Self-perceived Assessment Skills. *Applied measurement in education, 16*(4). <https://radarlombok.co.id/hasil-asesmen-nasional-sekolah-di-ntb-cukup-buruk.html>
- Wang, J. S., Pascarella, E. T., Nelson Laird, T. F., & Ribera, A. K. (2015). How Clear and Organized Classroom Instruction and Deep Approaches to Learning Affect Growth in Critical Thinking and Need for Cognition. *Studies in Higher Education, 40*(10), 1786-1807.
- Webb, N. (2002). Assessment Literacy in a Standard-Based Education Setting.
- Wolf, D., Bixby, J., Glenn III, J., & Gardner, H. (1991). Chapter 2: To Use Their Minds Well: Investigating New Forms of Student Assessment. *Review of Research in Education, 17*(1), 31-74.

APPENDICES

Appendix 1

Classroom Observation Sheet

Adopted from a research undertaken by Demirdogen and Korkut (2021).

Teacher (pseudonym) :

Date :

Lesson Topic :

Class Period Observed :

Student Demographics :

Time	Facts/Activities	Comments:
At the beginning of class		.
During class		..
At the end of class		.

Appendix 2

Semi-Structure Interview Questions

Adopted from a research undertaken by Demirdogen and Korkut (2021)

No	Original Source	Adopted/Developed	Indonesian Contextual Translation
1	What is your personal opinion about assessment?	What is your personal opinion about assessment?	Bisakah Bapak/Ibu guru menceritakan pengalaman Anda dalam merancang dan menerapkan penilaian di kelas sains yang diajar?
2	Do you think that assessment is useful for teaching?	Can you describe your understanding of the Merdeka Belajar Curriculum's approach to assessment?	Biasanya jenis-jenis penilaian seperti apa yang sedang Bapak/Ibu guru gunakan untuk menilai siswa/siswi di kelas sains, dan apa alasannya bpk/Ibu guru menggunakan hal tersebut?
3	What do you know about assessment? a. Where did you learn about assessment? (e.g., teacher education, professional development, colleagues, curriculum, and own interest)	<ul style="list-style-type: none"> How does this approach differ from your previous experiences with assessment, and What are your initial thoughts on the strengths and weaknesses of the assessment approach? 	Bagaimana cara Bapak/Ibu guru berbagi pengalamannya tentang literasi penilaian, misalnya kepada sesama guru atau calon-calon guru muda
4	What are your assessment purposes?	What are your assessment purposes? How do you think assessment can be used to promote the development of the core competencies emphasized by the curriculum (e.g., critical thinking, problem-solving, collaboration)?	Bagaimana caranya Bpk/Ibu Guru menentukan pendekatan penilaian yang paling tepat untuk tujuan pembelajaran, seperti sains?

5	What are various ways that you use during an assessment?	What are various ways that you use during an assessment?	Sejauh pengalaman Bapak/Ibu guru mengajar dan memberikan penilaian kepada siswa/i, praktik penilaian yang seperti apa yang efektif jika Bapak/Ibu guru ingin mengukur pemahaman pembelajaran siswa dalam sains? (seperti apa contohnya)
6	What are the most frequent assessment strategies that you used during your teaching?	What specific types of assessments do you feel are most effective in supporting student learning?	Saat merancang penilaian, faktor apa saja yang bisa Bapak/Ibu guru jadikan pertimbangan untuk bisa memastikan penilaian tersebut sejalan dengan tujuan pembelajaran?
7	How do you select, plan and use those strategies?	Can you explain your typical assessment process (selecting, planning and using of specific types of assessment) for a unit or lesson plan?	Seperti apa tahapan yang bapak/ibu guru terapkan dalam mendisain dan mengimplementasikan penilaian di kelas sains?
8	How do you as a teacher interpret assessment results?	How do you as a teacher interpret assessment results?	Bagaimana Bapak/Ibu guru memeberikan feedback kepada siswa/i
9		What challenges have you encountered in implementing the assessment approach? How have you addressed them?	Sejauh ini, dalam menerapkan praktik penilaian yang efektif, apa tantangan tersebut yang pernah Bapak/Ibu guru hadapi? Baaik secara internal atau eksternal. Misalnya, perubahan kurikulum atau standar penilaian
10		Are there any resources or professional development opportunities you would find helpful to gain a deeper understanding of assessment practices?	Sebagai seorang guru yang sudah berpengalaman, peluang apa yang Bapak/Ibu guru lihat untuk terus bisa meningkatkan praktik penilaian dalam mata pelajaran sains di Madrasah Aliyah?

Indonesian Language:

Introduction:

1. Terima kasih, sudah meluangkan waktu untuk berpartisipasi dalam wawancara ini.
2. Tujuan dari wawancara ini adalah untuk memahami perspektif Bapak/Ibu guru tentang praktik penilaian di kelas sains yang Anda ampu. Menjadi atensi utama kami bahwa semua tanggapan Anda sepenuhnya akan dijaga kerahasiaannya, dan tanya digunakan untuk tujuan riset saja.
3. Bolehkan bapak/Ibu guru menceritakan pengalaman secara singkat menjadi guru?

Questions:

1. Bisakah Bapak/Ibu guru menceritakan pengalaman Anda dalam merancang dan menerapkan penilaian di kelas sains yang diajar? (RQ1)
2. Bagaimana cara Bapak/Ibu guru berbagi pengalamannya tentang literasi penilaian, misalnya kepada sesama guru atau calon-calon guru muda? (RQ1)
3. Biasanya jenis-jenis penilaian seperti apa yang sedang Bapak/Ibu guru gunakan untuk menilai siswa/siswi di kelas sains, dan apa alasannya bpk/Ibu guru menggunakan hal tersebut? (RQ2)
4. Bagaimana caranya Bpk/Ibu Guru menentukan pendekatan penilaian yang paling tepat untuk tujuan pembelajaran, seperti sains? (RQ2)
5. Sejauh pengalaman Bapak/Ibu guru mengajar dan memberikan penilaian kepada siswa/i, praktik penilaian yang seperti apa yang efektif jika Bapak/Ibu guru ingin mengukur pemahaman pembelajaran siswa dalam sains? (seperti apa contohnya?) (RQ2)
6. Saat merancang penilaian, faktor apa saja yang bisa Bapak/Ibu guru jadikan pertimbangan untuk bisa memastikan penilaian tersebut sejalan dengan tujuan pembelajaran? (RQ3)
7. Seperti apa tahapan yang bpk/ibu guru terapkan dalam mendisain dan mengimplementasikan penilaian di kelas sains? (RQ3)
8. Bagaimana Bapak/Ibu guru memberikan feedback kepada siswa/i? (RQ3)
9. Sejauh ini, dalam menerapkan praktik penilaian yang efektif, apa tantangan terbesar yang pernah Bapak/Ibu guru hadapi? Baaik secara internal atau eksternal. Misalnya, perubahan kurikulum atau standar penilaian (RQ4)

10. Sebagai seorang guru yang sudah berpengalaman, peluang apa yang Bapak/Ibu guru lihat untuk terus bisa meningkatkan praktik penilaian dalam mata pelajaran sains di Madrasah Aliyah? (RQ4)

Appendix 3

Interview Transcripts

Q/A	Statements
Q	Bismillahirrahmanirrahim Assalamualaikum warahmatullahi wabarakatuh
A	Walaikumsalam warahmatullahi wabarakatuh
Q	<p>Pertama-tama, Saya ucapkan terima kasih atas kesediaan Bapak/Ibu Guru berpartisipasi dalam sesi wawancara ini. Tujuan utama wawancara ini adalah untuk menggali perspektif Anda mengenai praktik penilaian dalam mata pelajaran sains di kelas. Perlu kami garis bawahi bahwa seluruh tanggapan Anda akan dijaga kerahasiaannya. Data yang diperoleh dari wawancara ini akan digunakan sebagai pelengkap data untuk tugas ahir.</p> <p>Bisakah Bapak/Ibu Guru menceritakan secara singkat pengalaman mengajar di Madrasah Aliah yang Anda ampu saat ini</p>
A	<p>Pengalaman mengajar di setiap lokasi memang tidak sama karena karakteristik siswa yang berbeda-beda. Hal ini berlaku tidak hanya antar sekolah, tetapi juga antar kelas dalam satu sekolah. Oleh karena itu, dalam proses pembelajaran, saya tidak dapat menggunakan satu teknik atau metode secara seragam. Di MA Putra, misalnya, saya dapat menggunakan metode tertentu untuk mengajar di kelas A, namun belum tentu metode tersebut cocok untuk kelas lain. Variasi metode ini dikarenakan proses belajar mengajar yang bervariasi pula, mengikuti karakteristik siswa yang berbeda-beda. Sebagai guru, kita harus beradaptasi dan menyesuaikan metode mengajar dengan kebutuhan masing-masing kelas.</p>
Q	Bisakah Antum menceritakan pengalaman dalam merancang dan menerapkan penilaian di kelas sains?
A	<p>Berdasarkan pengalaman saya dalam menilai siswa di kelas sains, saya selalu menerapkan penilaian yang komprehensif dan multidimensi. Penilaian tidak hanya berdasarkan ujian tertulis atau ulangan harian, tetapi juga meliputi berbagai aspek lain, seperti sikap siswa dalam proses belajar mengajar, partisipasi mereka dalam kegiatan praktikum, dan penilaian keseluruhan aktivitas mereka.</p>
Q	Jika ada guru baru atau guru PPL misalkan, gimana cara Antum menjelaskan tentang literasi penilaian sains khususnya di mata pelajaran Kimia itu sendiri atau gimana cara antum membagi pengalamannya tentang cara penilaian dalam kelas?

Q/A	Statements
A	<p>Awalnya, ketika mahasiswa PPL datang ke sekolah, kami mengadakan sesi diskusi untuk membahas target yang ingin mereka capai selama PPL di Madrasah. Dalam sesi ini, kami menjelaskan karakteristik siswa di setiap kelas. Setelah memahami karakteristik tersebut, barulah kami membahas secara umum tentang proses pembelajaran dan penilaian di masing-masing kelas. Sebelum memperkenalkan model pembelajaran di Madrasah, kami menyampaikan tahapan-tahapan diskusi ini. Kami juga bersikap terbuka untuk menerima ilmu pengetahuan baru yang dibawa oleh mahasiswa PPL dari kampus.</p>
Q	Seperti apa pemahaman Antum tentang makna <i>*Literasi Penilaian*</i> ?
A	<p>Literasi Penilaian menurut saya adalah cara untuk menilai siswa secara menyeluruh, baik dari segi kognitif maupun afektif. Penilaian ini bukan hanya tentang nilai angka dan nilai ujian semata. Penilaian yang ideal haruslah variatif, karena jika kita hanya berpatokan pada nilai angka atau hasil ujian saja, hal tersebut mungkin tidak dapat menggambarkan keseluruhan kemampuan anak-anak.</p>
Q	Kemudian, biasanya jenis-jenis penilaian seperti apa yang Antum gunakan untuk menilai siswa di kelas sains dan apa alasan penggunaan dari jenis-jenis penilaian yang Antum terapkan?
A	<p>Dalam proses pembelajaran kimia, saya menerapkan beberapa metode penilaian. Pertama, untuk penilaian harian, saya umumnya menggunakan tes tertulis. Selain itu, saya juga mengadakan tes awal kemampuan di awal semester, meskipun tidak selalu dilakukan. Tes ini terdiri dari 1-2 pertanyaan tertulis untuk mengukur pemahaman awal siswa tentang materi kimia. Hasil tes ini dapat dilihat dalam waktu singkat dan membantu saya memahami kemampuan dasar mereka.</p> <p>Tes awal ini juga berfungsi untuk mengetahui sejauh mana siswa telah menguasai materi sebelumnya. Untuk ulangan harian selanjutnya, saya kembali menggunakan tes tertulis sebagai metode penilaian.</p>
Q	Bagaimana cara Antum menentukan pendekatan penilaian yang paling tepat diterapkan di kelas?

Q/A	Statements
A	<p>Penilaian saya lakukan berdasarkan observasi terhadap kondisi dan kemampuan kognitif anak-anak. Tes awal membantu saya memahami kemampuan dasar mereka, dan dari situ saya dapat merancang jenis-jenis penilaian yang sesuai. Untuk siswa yang lebih menyukai kegiatan praktik, penilaian dilakukan melalui praktik itu sendiri. Saya memberikan praktik ringan yang memungkinkan saya untuk mengamati dan memberikan nilai secara langsung. Penilaian berbasis hitungan mungkin tidak ideal dalam situasi praktik. Melalui praktik, saya dapat melihat siswa mana yang lebih aktif dan menilai berdasarkan aktivitas mereka. Pendekatan penilaian ini bervariasi, dilakukan di awal, pertengahan, dan akhir pembelajaran. Jenis penilaiannya pun beragam, tergantung pada kegiatan yang dilakukan. Untuk kegiatan praktik seperti tadi, penilaian tidak dilakukan secara tertulis, melainkan melalui portofolio atau laporan kegiatan siswa. Pendekatan ini disesuaikan dengan jenis kegiatannya. Sebagai contoh, dalam pelajaran psikometrik (hitungan) di pertemuan sebelumnya, penilaian dilakukan melalui perhitungan dan tes.</p>
Q	<p>Sejauh pengalaman Antum mengajar dan memberikan penilaian kepada siswa/i, praktik penilaian yang seperti apa yang efektif jika Antum ingin mengukur pemahaman pembelajaran siswa dalam sains? Dan seperti apa contohnya?</p>
A	<p>Dalam mata pelajaran kimia, saya lebih sering menggunakan tes tertulis untuk menilai pemahaman siswa terhadap materi yang melibatkan perhitungan matematis. Untuk materi yang lebih mudah dipahami, saya terkadang menggunakan portofolio atau metode penilaian lainnya. Namun, tidak ada satu metode penilaian yang paling efektif. Keefektifan penilaian tergantung pada konteks materi yang diajarkan.</p>
Q	<p>Kalau yang paling sering Antum gunakan?</p>
A	<p>Berdasarkan pengalaman saya, bentuk penilaian yang paling sering digunakan di sekolah adalah tes tulis. Hal ini dikarenakan adanya kegiatan rutin seperti ulangan harian 1, 2, dan 3 yang mewajibkan siswa untuk mengerjakan soal-soal tertulis.</p>
Q	<p>Jenisnya seperti apa ditailnya, Ust?</p>
A	<p>Biasanya dalam bentuk soal isian, soal uraian.</p>
Q	<p>Untuk soal uraian, uraiannya singkat atau panjang atau bagaimana?</p>
A	<p>Sebelumnya, soal didominasi oleh uraian, tanpa pembedaan panjang atau pendek. Setahun terakhir, kami telah menerapkan tiga model soal baru yang sesuai dengan standar isi yang diperbarui. Sejak kemarin, kami telah mulai menggunakan format soal yang terdiri dari jawaban singkat, uraian, dan pilihan ganda.</p>

Q/A	Statements
Q	Apakah itu berdasarkan guideline atau arahan dari pusat atau memang sekolah memiliki inisiatif sendiri untuk mengembangkan jenis-jenis penilaian?
A	Model penilaian ini kami terapkan setelah berdiskusi dengan pengawas dan merujuk pada pedoman dari pusat. Kami menemukan bahwa setiap penilaian harus mencakup minimal tiga jenis soal, seperti pilihan ganda, benar salah, dan jenis soal lainnya. Setelah membaca pedoman pusat, para guru menjadi lebih terbuka dan kami sepakat untuk menerapkan model penilaian ini pada setiap kegiatan penilaian, meskipun hanya menggunakan tiga jenis soal yang disebutkan sebelumnya.
Q	Oke, tujuannya apa kira-kira meragamkan dari jenis-jenis penilaian?
A	Keragaman dalam soal bertujuan untuk membantu siswa, terutama dalam menjawab pertanyaan singkat. Siswa tidak perlu memberikan jawaban yang panjang, melainkan kalimat singkat yang padat dan jelas. Jawaban untuk pertanyaan singkat ini umumnya dapat diperoleh dari materi yang telah mereka pelajari selama proses belajar. Hal ini dapat membantu siswa dalam meningkatkan nilai mereka. Sementara, soal uraian lengkap umumnya memiliki persentase sekitar 40%. Meskipun tidak dapat dipastikan bahwa hanya 20% siswa yang mampu menjawabnya dengan sempurna, mereka yang mampu menyelesaikannya dengan baik kemungkinan besar akan mendapatkan nilai yang sesuai dengan standar.
Q	Pertanyaan berikutnya, saat merancang jenis-jenis penilalaian yang tadi, faktor-faktor apa saja yang menjadi pertimbangan Antum ataupun sekolah untuk memastikan bahwa penilaian itu sejalan dengan tujuan pembelajaran?
A	Pertama-tama, pertimbangan utama kami didasarkan pada aturan yang ditetapkan oleh Pusat. Hal ini merupakan keharusan yang harus kami patuhi. Pemilihan model-model soal yang digunakan saat ini memudahkan analisis soal secara terperinci. Misalnya, soal-soal dengan format ABC mudah dianalisis. Hal ini memungkinkan kami untuk dengan mudah mengidentifikasi kekurangan materi yang dialami oleh anak-anak. Kemudahan analisis ini merupakan salah satu pertimbangan utama kami. Dengan cara ini, kami dapat mengevaluasi efektivitas penerapan program yang telah kami lakukan. Pertimbangan ini sejalan dengan hasil pertemuan rapat dengan Bapak Kepala Madrasah.
Q	Apakah itu sama untuk kelas 1, kelas 2 dan Kelas 3?
A	Hasil kesepakatan bersama dalam rapat tersebut diberlakukan di semua jenjang pendidikan.

Q/A	Statements
Q	Kalau untuk, misalnya, membiasakan siswa ke jenis-jenis soal, taruhlah seperti soal yang sekarang ini <i>Asesmen Nasional</i> , Apakah sekolah menyiapkan soal yang mungkin menyerupai jenisnya atau levelnya?
A	Benar, jenis soal yang menyerupai soal Asesmen Nasional memang ada. Kami menggunakan jenis soal tersebut karena Kemendikbud telah menerapkan berbagai tipe soal, seperti yang telah disebutkan sebelumnya. Namun, tingkat kesulitannya belum mencapai level yang sama dengan soal yang dibuat oleh Kemendikbud. Paling tidak, kami berusaha untuk menyerupai dan menyesuaikan soal dengan Kompetensi Dasar (KD). Kemungkinan besar, soal kami tidak akan persis sama dengan soal Asesmen Nasional, dan tingkat kesulitannya mungkin lebih rendah.
Q	Seperti apa tahapan yang Antum terapkan dalam mendisain dan mengimplementasikan penilaian di kelas sains?
A	Dalam menyusun tes tertulis, umumnya penilaian difokuskan pada analisis Kompetensi Dasar (KD). Pertama, KD diidentifikasi dan dicocokkan dengan Kriteria Ketuntasan Kurikulum (KKU) yang ditetapkan. Selanjutnya, level kognitif yang ingin diukur dipertimbangkan, dan soal dirancang sesuai dengan level tersebut. Contohnya, jika levelnya analisis, soal analisis pun disusun. Namun, perlu diakui bahwa keterbatasan sumber daya di sekolah-sekolah level bawah menghambat penerapan proses yang komprehensif ini secara ideal, seperti yang dilakukan di level nasional.
Q	Jika siswa-siswa diikuti Olimpiade, bagaimana persiapan dari para asatis untuk mempersiapkan mereka, apakah membuat soal sendiri atau memang sudah ada bank soal?
A	Di sekolah kami, untuk persiapan Olimpiade, biasanya diadakan kelas khusus atau les khusus. Jumlah peserta yang mengikuti program ini bervariasi. Setelah mengikuti pelatihan, para peserta akan mendapatkan materi dan latihan menggunakan soal-soal tingkat nasional. Soal-soal ini dibahas bersama dalam pertemuan dan dikerjakan bersama sebagai latihan. Seperti yang saya sebutkan sebelumnya, untuk lomba OSN tingkat sekolah, materinya disesuaikan dengan standar nasional. Oleh karena itu, perlu disiapkan juga soal-soal latihan selevel nasional.
Q	Yang sifatnya penilaian kelas baik formatif atau sumatif itu sendiri gimana cara Antum memberikan feedback ke siswa, apakah secara langsung, secara oral atau tertulis atau bagaimana? Dan bagaimana dengan orang tua siswa?

Q/A	Statements
A	<p>Hasil ujian atau rapor, biasanya dalam bentuk cetak, dibagikan kepada siswa. Selanjutnya, wali murid diminta untuk menandatangani hasil rapor tersebut. Mengingat pertemuan dengan wali murid terkadang membutuhkan waktu lama, kepala asrama dapat mewakili siswa dalam menyerahkan rapor sebagai bentuk laporan kepada kepala asrama mereka. Ketika wali murid datang ke asrama, kepala asrama dapat menyampaikan hasil rapor tersebut kepada mereka.</p>
Q	<p>Kalau untuk tugas kelompok ataupun tugas rumah ataupun kegiatan kelas lainnya, bagaimana cara Antum memberikan feedback ke para siswa</p>
A	<p>Dalam memberikan umpan balik kepada siswa, tugas-tugas tersebut dikembalikan kepada mereka setelah diberi nilai. Hal ini bertujuan agar siswa dapat melihat hasil kerjanya dan kemudian berdiskusi bersama. Dalam diskusi tersebut, guru dan siswa bersama-sama menganalisis kekurangan dalam hasil kerja dan mencari solusi untuk memperbaikinya. Pendekatan ini menciptakan proses belajar timbal balik yang bermanfaat bagi pengajar dan siswa.</p>
Q	<p>Apa manfaat yang Antum dapatkan dengan melihat nilai siswa-siswa yang ada?</p>
A	<p>Sebagai imbalannya, saya dapat menilai kemampuan anak-anak pada poin-poin tertentu. Hal ini memungkinkan saya untuk memetakan pencapaian mereka, misalnya dengan mengidentifikasi anak-anak yang menunjukkan kelemahan dalam materi tertentu pada tahun ini. Informasi ini kemudian dapat digunakan untuk mengevaluasi proses belajar-mengajar pada tahap selanjutnya.</p>
Q	<p>Pertanyaan berikutnya, sejauh ini dalam menerapkan praktik penilaian yang efektif Apa tantangan terbesar yang Antum hadapi atau alami, baik secara internal maupun eksternal? misalnya, apakah perubahan kurikulum itu berpengaruh pada cara Antum memberikan penilaian?</p>
A	<p>Salah satu tantangan terbesar kami adalah kurangnya respons dari anak-anak terhadap hasil penilaian, meskipun mereka telah diberikan arahan. Hal ini menyebabkan anak-anak terkadang tidak termotivasi untuk mengubah atau memperbaiki kekurangan mereka. Saya menduga hal ini terjadi karena kurangnya semangat dari anak-anak untuk berkembang. Terkadang, mereka bahkan mengabaikan hasil ujian dan tidak menunjukkan respons terhadapnya.</p>
Q	<p>Berapa persen kira-kira kondisi siswa seperti itu?</p>

Q/A	Statements
A	Tingkat responsnya sangat rendah. Dari 30 orang di kelas, hanya sekitar 5 sampai 10 orang yang merespon, dan itupun saya tidak yakin apakah mereka benar-benar memahami materi.
Q	Terus kalau dari aturan-aturan yang terbaru, seperti kurikulum atau sejenisnya, apakah itu merupakan tantangan?
A	Perubahan kurikulum, justru mempermudah proses penilaian. Saya amati, baik dalam kurikulum terbaru maupun lama, perubahannya tidak terlalu signifikan. Mungkin hanya revisi kecil. Model pembelajarannya pun tetap scientific learning. Bagi saya, perubahan ini menguntungkan karena memberikan keleluasaan dalam mengatur kelas dan lain sebagainya. Di sisi lain, beberapa guru mungkin merasa dirugikan karena kehilangan beberapa jam pelajaran. Contohnya, beberapa mata pelajaran mengalami pengurangan jam pelajaran. Jadi, terdapat pro dan kontra terkait perubahan kurikulum ini. Ada yang merasa diuntungkan, ada pula yang dirugikan.
Q	Sebagai seorang guru yang sudah berpengalaman, peluang apa yang Antum lihat untuk terus bisa meningkatkan praktik penilaian dalam mata pelajaran sains di Madrasah Aliyah?
A	Terdapat banyak sekali kegiatan pengembangan diri yang tersedia bagi para guru, namun sifatnya masih sukarela, belum ada yang diwajibkan. Oleh karena itu, tingkat partisipasi guru dalam kegiatan pengembangan diri ini bervariasi. Bagi guru yang ingin serius meningkatkan diri, mereka akan mengikuti semua kegiatan yang tersedia. Pada dasarnya, semua fasilitas untuk pengembangan diri guru telah disiapkan melalui berbagai kegiatan pelatihan dan sejenisnya. Saya pribadi bersyukur selalu mengikuti kegiatan-kegiatan tersebut.
Q	Seperti apa Ustaz, kegiatan-kegiatan untuk meningkatkan kemampuan para guru dalam mendesain soal-soal ataupun penilaian itu sendiri?
A	<p>Pada pelatihan kemarin, kita membahas mengenai pembuatan soal, sebuah kegiatan rutin yang diadakan di awal semester. Kami mempelajari bagaimana menulis soal yang sesuai dengan Kompetensi Dasar (KD) dan level kognitif yang diinginkan. Bapak pengawas bertindak sebagai narasumber dalam pelatihan ini. Selain itu, setelah proses pembuatan soal, kami juga melakukan verifikasi soal untuk memastikan kualitasnya.</p> <p>Saya juga berkesempatan mengikuti beberapa pelatihan penulisan soal higher order thinking (HOT) melalui aplikasi si-pintar. Alhamdulillah, dalam dua tahun terakhir, saya sudah mengikuti tiga kali pelatihan ini.</p>

Q/A	Statements
Q	Oke berarti untuk soal higher order thinking tadi itu seberapa persen bisa kita terapkan atau para asatis terapkan dalam menilai kemampuan siswa?
A	Tentu saja, penyusunan soal akan mengacu pada silabus yang telah ditetapkan. Persentase soal pada setiap tingkatan, seperti low, medium, dan higher order thinking skill, akan ditentukan oleh para guru.
Q	Tadi soal-soalnya itu kira-kira persentasenya berapa?
A	Seperti yang disepakati di Madrasah kemarin, meskipun saya tidak ingat persisnya, sekitar 25% soal penilaian kita akan menggunakan format Higher Order Thinking Skills (HOTS). Penting untuk diingat bahwa tidak semua soal dapat menggunakan format HOTS. Jika tidak salah, sekitar 25% soal HOTS akan kita manfaatkan.
Q	Kelas tigas (12) misalkan, apakah ada persiapan untuk menghadapi ujian masuk perguruan tinggi di mana itu membutuhkan pembiasaan siswa pada soal high order thinking?
A	Fokus utama kami sejauh ini adalah mempersiapkan siswa untuk ujian asesmen madrasah dan nasional. Persiapan untuk ujian perguruan tinggi belum dapat kami laksanakan secara menyeluruh. Pada tingkat sekolah, fokus kami memang lebih terarah pada persiapan ujian asesmen nasional. Untuk persiapan perguruan tinggi, kami umumnya hanya membantu siswa dalam hal administrasi, pencarian sekolah, dan pendaftaran. Pemberian bimbingan belajar khusus untuk tes masuk perguruan tinggi belum tersedia saat ini.
Q	Menurut Antum nih, soal untuk masuk perguruan tinggi itu lebih tinggi levelnya dari asesmen nasional atau memang sama?
A	Dalam pandangan saya, tidak ada perbedaan yang signifikan karena saya mendasarkan penilaian ini pada analisis pengetahuan minimum yang wajib dikuasai oleh siswa.
Q	Selain dari yang tadi, penting tidak penguatan dalam membuat soal ataupun jenis-jenis penilaian kita sebagai guru dalam hal pemahaman konsep materi?

Q/A	Statements
A	<p>Pemahaman konsep sangatlah penting, bahkan bagi kita yang telah mengikuti pelatihan pun terkadang mengalami kesulitan dalam penerapannya. Hal ini dapat menyebabkan kehilangan fokus dan lupa materi yang telah dipelajari. Oleh karena itu, penting bagi sekolah untuk mengadakan kegiatan pembimbingan penyusunan soal di awal semester. Kegiatan ini dapat membantu meningkatkan pemahaman dan mengingatkan kita tentang materi yang telah dipelajari. Tanpa kegiatan pembimbingan, pembuatan soal terkadang hanya dilakukan secara asal-asalan.</p>
Q	<p>Oke, ada tidak kegiatan-kegiatan selain dari tadi aplikasi pintar itu pembimbingan pembuatan soal ini atau model secara langsung yang Antum bisa ikuti?</p>
A	<p>Pada awal setiap kegiatan ujian, kami biasanya menghadirkan Bapak Pengawas untuk membantu penyusunan soal. Namun, karena keterbatasan waktu, tidak selalu semua guru dapat hadir. Pada situasi seperti ini, kami biasanya hanya menghadirkan beberapa guru sebagai sampel. Sebagai contoh, untuk mata pelajaran biologi hari ini, kami memeriksa semua soal secara menyeluruh dari awal, mulai dari menentukan kode dasar (KD) hingga tahapan selanjutnya. Kami melakukan analisis mendalam pada mata pelajaran ini. Namun, karena keterbatasan waktu, tidak semua guru dapat menyelesaikan tahap analisis soal dan merumuskan perbaikan untuk kedepannya. Biasanya, hanya satu atau dua guru saja yang dapat menyelesaikan tugas ini.</p>
Q	<p>Dengan kata lain, kegiatan untuk meningkatkan kapasitas guru dalam literasi penilaian memang dapat dikatakan rutin. Saya rasa cukup sampai di sini untuk sesi wawancara terkait literasi penilaian, khususnya di kelas sains. Terima kasih atas waktunya Antum. Jazakalloh.</p>

Appendix 4

Data Coding Analysis



Quirkos Report

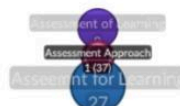
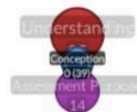
This report was generated by Mahurni on Jun 28, 2024 12:30:53 AM for the following file:
C:/Users/user/Downloads/Quirkos Interview.qrk

Sources Summary

Title	Author	Date and Time	Length	Quotes #
Revisi Transkrip Ujang - Translated	Mahurni	Jun 24, 2024 6:32:07 AM	17078	68
Interview with Mr Rozi	Mahurni	Jun 24, 2024 6:32:07 AM	18935	70
Revisi Transkrip Budi (Translated)	Mahurni	Jun 24, 2024 6:32:07 AM	23265	70
Observasi Mr Budi	Mahurni	Jun 27, 2024 10:22:38 AM	10116	21
Observation Mr Ujang	Mahurni	Jun 27, 2024 10:22:38 AM	5077	10
Observation mr Rozi	Mahurni	Jun 27, 2024 10:22:38 AM	8612	18
RPP PAK BUDI	Mahurni	Jun 27, 2024 5:23:21 PM	10023	11
RPP PAK UJANG	Mahurni	Jun 27, 2024 5:23:21 PM	11596	16
MODUL AJAR 4 SMA PAK ROZI	Mahurni	Jun 27, 2024 5:23:21 PM	17323	20

Main Canvas Views

Quirks Canvas - Primary



Appendix 5

Documents of Classroom Observation and Interviews



