CHAPTER I
INTRODUCTION

This chapter discusses the background of the study, research questions, research purpose, research significance, research scope, conceptual framework, and previous studies. Each section is arranged to provide a coherent foundation for understanding the direction of this research, beginning with the rationale for investigating rubric use and professional judgement in speaking assessment. The subsequent sections then clarify the research problem, objectives, and theoretical grounding that shape the overall study, before outlining the boundaries of the research and the conceptual model that connects its key constructs. Taken together, these sections establish the conceptual and contextual basis upon which the research design and methodology presented in the following chapters are developed.

A. Research Background

Speaking assessment has long been recognized as one of the most complex areas in English as a Foreign Language (EFL) assessment. Scholars such as Brown and Abeywickrama (2020) and McNamara and Roever (2020) argue that speaking is fundamentally different from receptive skills because it requires immediate language production under real-time conditions. Their studies demonstrate that students' spoken performance is influenced not only by linguistic competence, but also by anxiety, task demands, confidence, and interactional pressure. As a result, students may produce oral responses that do not fully represent their actual language ability, making speaking assessment particularly challenging for EFL teachers.

To address this complexity, many researchers have examined the use of rubrics in speaking assessment. Brookhart (2020) explains that rubrics are designed to provide clear performance criteria and to promote fairness and transparency in assessment. Similarly, Harris and Brown (2021), based on classroom-based studies, found that rubrics help teachers structure their assessment decisions and communicate expectations to students. In the Indonesian EFL context, Putra and Nugroho (2022) report that rubric-based speaking assessment is widely

implemented in senior high schools to support standardized evaluation aligned with curriculum demands.

However, research has also identified limitations in rubric-based assessment. Harris and Brown (2021) found that student speaking performances often do not align neatly with rubric descriptors, particularly when students demonstrate strong communicative meaning but weaker grammatical accuracy, or vice versa. Fulcher (2022) further argues that because speaking is dynamic and context-dependent, no rubric can fully capture all aspects of oral performance. These findings indicate that relying solely on rubrics may oversimplify the assessment of speaking skills.

Consequently, Wyatt-Smith, Klenowski, and Gunn (2020) have emphasized the importance of teachers' professional judgement in speaking assessment. They describe professional judgement as teachers' informed decision-making based on experience, pedagogical knowledge, and classroom context. Their research shows that teachers often consider factors such as student effort, progress, confidence, and task difficulty when evaluating speaking performance. In EFL classrooms, particularly in contexts with diverse student proficiency levels, professional judgement is frequently used to complement rubric criteria in order to maintain fairness and contextual relevance.

This issue also became relevant to the researcher during the Teaching Practicum Program (PPL) at SMAN 21 Bandung. During the program, a student asked how the teacher assessed their speaking ability, particularly because some students who were less fluent in speaking received similar scores to those who demonstrated more fluent speaking performance. This situation raised the researcher's interest in understanding how teachers applied assessment criteria and made evaluative decisions when assessing students with different levels of speaking performance. It further encouraged the researcher to explore how speaking rubrics and professional judgement are used in actual classroom assessment.

Although studies on rubrics and professional judgement are well documented, most of them examine these elements separately. Research on rubrics tends to focus on their design, reliability, and standardization (Brookhart, 2020; Harris & Brown, 2021), while studies on professional judgement focus on teacher expertise and

assessment decision-making (Wyatt-Smith et al., 2020; Fulcher, 2022). Very few studies investigate how EFL teachers actually experience the integration of rubrics and professional judgement in real speaking assessment practices, especially through teachers' personal narratives in Indonesian senior high school contexts.

This study addresses the gap by exploring two EFL teachers' experiences in using rubrics and professional judgement in speaking assessment at SMAN 21 Bandung through Narrative Inquiry. It examines how teachers interpret, negotiate, and reflect on assessment decisions in actual classroom contexts.

B. Research Questions

To focus the investigation, the research addresses the following questions:

1. How do EFL teachers at SMAN 21 Bandung use speaking rubrics as part of the assessment process to inform the evaluation of students' speaking performance?
2. How do the teachers describe the role of professional judgement in interpreting assessment results and making evaluative decisions about students' speaking skills?
3. What challenges and reflections do the teachers share about combining rubric and professional judgement in speaking assessment?

C. Research Purposes

Based on the research questions, the research aims to achieve the following purposes:

1. To explore how EFL teachers at SMAN 21 Bandung use speaking rubrics to assess and evaluate students' speaking performance.
2. To examine how teachers use professional judgement when interpreting assessment results and making evaluative decisions about students' speaking skills.
3. To identify the challenges experienced by teachers and their reflections on combining speaking rubrics with professional judgement in speaking assessment.

UNIVERSITAS ISLAM NEGERI
SUNAN GUNUNG DJATI
BANDUNG

D. Research Significance

The research is expected to provide both theoretical and practical significance. These contributions are intended to benefit not only academic discussions on language assessment but also the everyday realities faced by teachers when scoring students' speaking performance in the classroom. By addressing both dimensions, this study seeks to bridge the gap between conceptual understanding of assessment and the practical decision-making that teachers undertake on a daily basis. The following sub-sections elaborate on each type of significance in more detail.

1. Theoretical Significance

Theoretically, this research is expected to contribute to the growing body of knowledge on speaking assessment by offering insight into how rubric-based evaluation and teacher professional judgement are combined in real classroom practices. Previous studies on speaking assessment have predominantly examined rubrics as standardized assessment tools focusing on criteria design, reliability, and scoring consistency (Brookhart, 2020; Harris & Brown, 2021), while other studies have emphasized professional judgement as a form of teacher expertise in assessment decision-making (Wyatt-Smith et al., 2020; Fulcher, 2022). However, relatively limited research explores how these two elements function together through teachers' lived experiences in classroom-based speaking assessment.

By examining two teachers' narratives, this study sheds light on the interpretive nature of rubric use, illustrating that assessment is not merely a technical process of matching student performance to descriptors, but a pedagogical decision-making practice influenced by experience, contextual understanding, and reflective thought. Additionally, this research may strengthen the theoretical discussion on Narrative Inquiry as a methodological approach in language assessment research, emphasizing the value of teachers' stories as a legitimate source of knowledge that can reveal complexities in assessment practices which are often not fully captured through quantitative or experimental approaches.

UNIVERSITAS ISLAM NEGERI
BANDUNG

2. Practical Significance

Practically, this research provides benefits for several educational stakeholders. For EFL teachers, it offers reflective insights that can support more balanced development of speaking assessment, encouraging them to view rubrics not as rigid scoring instruments, but as flexible guides that require thoughtful interpretation. The findings can motivate teachers to trust their professional judgement and use it responsibly to ensure fairness and contextual accuracy when evaluating students' oral performance.

For researchers, the study may serve as a reference for future investigations related to language assessment, teacher judgement, and Narrative Inquiry. It can inspire further research that values teachers' voices and considers the lived realities of assessment practices, rather than focusing solely on measurement principles. For schools and educational policymakers, this study may provide considerations for developing assessment policies that support teacher autonomy while maintaining fairness and accountability. By acknowledging the coexistence of rubric and judgement, schools can design assessment systems that provide clear standards without undermining professional expertise. Ultimately, the study encourages a more holistic and humane approach to speaking assessment in senior high schools.

E. Research Scope

This research focuses on the narrative experiences of two English teachers at SMAN 21 Bandung who conduct speaking assessment using rubric-based scoring and professional judgement. The study does not attempt to represent the practices of all English teachers, schools, or educational institutions, but rather to explore deeply how the two teachers negotiate the use of rubric and personal evaluative decisions in their respective classroom contexts. By focusing on two participants, the research allows for both shared and distinct experiences in speaking assessment to emerge, enabling limited comparison while maintaining narrative depth. This approach is consistent with the principles of Narrative Inquiry, which emphasizes personal and context-bound experiences as meaningful sources of knowledge, prioritizing depth of understanding over broad generalization.

UNIVERSITAS ISLAM NEGERI
BANDUNG

Data in this study were collected primarily through semi-structured interviews, supported by assessment documents, particularly speaking assessment rubrics. The research did not predetermine the type of speaking task being assessed, as Narrative Inquiry values authentic circumstances that naturally emerge in real educational settings. Therefore, the speaking assessment tasks were identified during data collection, allowing the teachers to describe their experiences without restriction or manipulation by the researcher. This flexible approach allowed the narratives to reflect the teachers' reported practices rather than artificially imposed research variables.

F. Conceptual Framework

This research is grounded in Brown and Abeywickrama's (2020) perspective on speaking assessment, which provides the main theoretical foundation for understanding teachers' assessment practices. Speaking assessment is a complex process that involves both systematic assessment tools and teachers' professional judgement. In EFL classrooms, speaking assessment is not limited to scoring students' performance, but also involves interpreting assessment evidence to make evaluative decisions about students' speaking ability. According to Brown and Abeywickrama (2020), assessment refers to the process of collecting information about students' language performance, while evaluation involves making decisions or judgments based on that information. This distinction provides the foundation for understanding how rubrics and professional judgement function together in speaking assessment.

Within this framework, speaking rubrics are positioned as formal assessment tools that provide structure and criteria for evaluating students' oral performance. Rubrics typically consist of clearly defined criteria such as fluency, pronunciation, grammar accuracy, vocabulary use, and communicative effectiveness. According to Brookhart (2020), rubrics support assessment by clarifying performance expectations and promoting consistency, but they do not eliminate the need for interpretation. Therefore, rubric use represents the assessment component that generates evidence of students' speaking performance.

However, assessment evidence alone does not automatically lead to evaluation. Teachers must interpret rubric results in relation to classroom realities, task demands, and students' individual conditions. This interpretive process is referred to as professional judgement. According to Wyatt-Smith, Klenowski, and Gunn (2020), professional judgement is a central aspect of assessment practice, as it allows teachers to make informed decisions based on experience, pedagogical knowledge, and contextual understanding. In speaking assessment, professional judgement enables teachers to determine how rubric scores should be understood and weighted when making evaluative decisions.

The framework also emphasizes the role of contextual interpretation in connecting rubric-based assessment and evaluation. Teachers do not apply rubrics mechanically; instead, they interpret assessment criteria by considering factors such as students' effort, improvement over time, task difficulty, communicative intent, and classroom interaction. Harris and Brown (2021) argue that teachers' assessment practices are shaped by classroom context and professional reasoning, highlighting that judgement is embedded in real teaching situations rather than isolated scoring procedures.

Based on this perspective, the conceptual framework of this research illustrates the relationship between speaking assessment, rubric structure and criteria, professional judgement, and evaluation of students' speaking performance. Speaking assessment is viewed as primarily formative, as it supports feedback and learning development, while also contributing to evaluative decisions about students' speaking skills. Through Narrative Inquiry, this research explores how two EFL teachers experience and negotiate these components in their assessment practices, revealing how assessment and evaluation are interconnected through teachers' narratives.

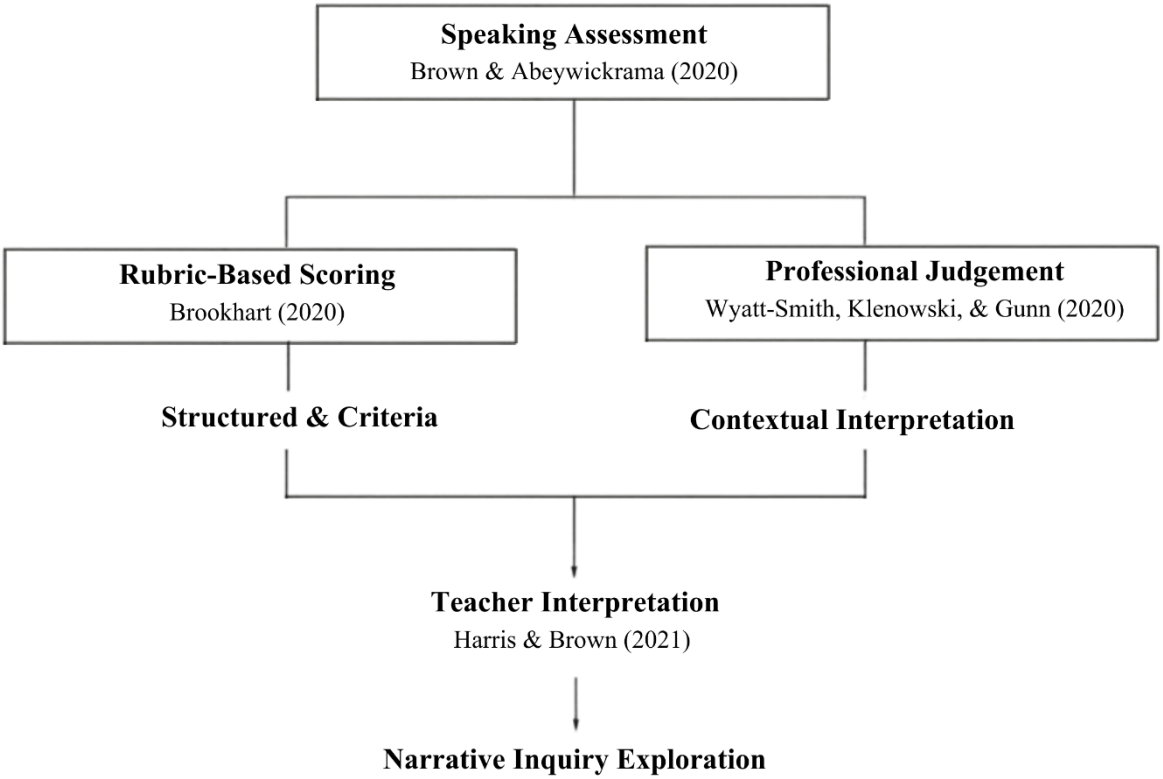


Figure 1. 1 Conceptual Framework

G. Previous Studies

Recent studies emphasize the importance of rubrics and teacher judgement in classroom assessment. Harris and Brown (2021) found that teachers use rubrics not as fixed scoring tools but as flexible guidelines interpreted based on classroom realities. Brookhart (2020) similarly stated that rubrics can support fairness only when teachers apply them thoughtfully through reflective judgement rather than mechanically. Wyatt-Smith et al. (2020) further demonstrated that professional judgement is an essential part of assessment practice and becomes reliable when grounded in experience, shared standards, and reflective thinking. Although these studies examine the relationship between rubrics and professional judgement, they primarily focus on how rubrics operate in general classroom evaluation and do not deeply explore teachers' lived experiences in specific contexts such as speaking assessment.

Research related more directly to speaking assessment has also highlighted the complexity of scoring oral performance. Brown and Abeywickrama (2020) pointed

out that the real-time nature of speaking makes performance difficult to capture using strictly standardized scoring tools, because spoken language is spontaneous, context-dependent, and influenced by individual interactional styles. Similarly, Alward et al. (2021), in a study involving secondary students' oral communication assessment, revealed that rubrics often fail to represent nuanced performance aspects such as effort, confidence, or communicative strategies. They emphasized that teachers frequently make scoring adjustments based on classroom observations that are not stated explicitly in the rubric. These findings show that speaking assessment requires subjective interpretation even when rubrics are present.

In the Indonesian context, a study by Putra and Nugroho (2022) reported that English teachers in senior high schools frequently negotiate between rubric criteria and their professional judgement due to curriculum demands, large class sizes, and varied student proficiency levels. While rubrics provide structured scoring categories, Indonesian teachers often revise scores to reflect student improvement, interaction strategies, task difficulty, and effort—elements they feel should be rewarded. However, these studies still focus on teachers' practices from a survey or systemic perspective and do not examine the internal reasoning process or reflective experiences that shape such decisions.

Narrative Inquiry-based investigations have recently gained attention in language education research because they allow teachers' voices to be heard. Caine et al. (2020) highlighted that Narrative Inquiry is especially useful for studying assessment practices because it captures personal interpretations, tacit knowledge, and emotions that cannot be accessed through questionnaires or experiments. Yet, its application in language assessment remains underrepresented internationally, and is even more limited in Indonesia, where teacher assessment research still leans toward quantitative and rubric-reliability analysis. These gaps indicate a need for narrative research that centers on teachers' reflections, beliefs, and decision-making processes when scoring students' speaking performance.

Therefore, this study fills the gap by focusing on two English teachers' personal narratives regarding the combined use of rubrics and professional judgement in speaking assessment at SMAN 21 Bandung. Rather than examining scoring validity

or comparing assessment tools, this research explores how the teachers interpret rubric descriptors, negotiate subjective and objective assessment considerations, and reflect on fairness and contextual accuracy in real classroom practice. By adopting Narrative Inquiry, the study offers an interpretive perspective that contributes to current assessment literature, particularly in the Indonesian senior high school context. The involvement of two participants also allows for a richer and more nuanced understanding of how different teachers may share common challenges while also navigating unique classroom experiences in their assessment practices.

UNIVERSITAS ISLAM NEGERI
SUNAN GUNUNG DJATI
BANDUNG