CHAPTER V

CONCLUSIONS, IMPLICATIONS, AND RECOMMENDATIONS

After presenting the research results and providing an in-depth discussion of the previous chapter, this chapter summarizes the main research findings in structured conclusions. In addition, this chapter also provides implications of the theoretical and practical research results, which are expected to contribute to the development of knowledge and practice in related fields. Furthermore, the author outlines the various limitations encountered during the research process so that they can be considered when interpreting the results. Lastly, this chapter is finalized with recommendations for future researchers and interested parties to provide direction and suggestions for further improvements.

5.1. Conclusions

As this study explored the EFL secondary teachers' assessment literacy as reflected in the way they design classroom tests and examined the design of the teacher-made tests through their alignment with established test construction guidelines, which reflect their assessment practices, it can be concluded that teachers' knowledge and practices in constructing summative tests are not yet optimal, as indicated by the discrepancy between the claimed knowledge and practice and the real practice in the field. This mismatch, which occurs in the planning, designing, and trying out stages, is mainly reflected in the practice of preparing blueprints that were often completed after the tests have been made, making it merely a fulfilment of administrative demands, not an initial guideline in test preparation; in the mastery of the technical aspects of test designs which was not fully reflected in the test products they produced; and in the process of cross-verification by peers and trying out the tests with students, which, although admittedly carried out, were not really carried out thoroughly. These situations point to a substantial gap between the theoretical understanding of assessment literacy and its practical application in constructing classroom-based tests.

The analysis of the seven teacher-made tests reveals a multifarious use of objective and subjective test types, including single and complex multiple-choice questions, true-false, matching, short answer, and essay. Of these types, single-answer multiple-choice tests keep

Enggar Pangesti Wibowo, 2025
INVESTIGATING EFL SECONDARY TEACHERS' ASSESSMENT LITERACY: THEIR KNOWLEDGE AND PRACTICE IN TEST CONSTRUCTION
Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

being the most commonly used format. The content of the summative tests in this study primarily targets vocabulary, reading, grammar, and writing skills. Moreover, in constructing the tests, teachers utilize various sources, including textbooks, internet, modules, teacher handbooks, previous years tests, question banks, student work or assignments in class, and exercises from standardized tests such as TOEIC, TOEFL, and IELTS, with textbooks and internet became the most frequently utilized references. In addition, some teachers also refer to other sources such as magazines, novels, newspapers, non-school textbooks, and even AI-based platforms such as ChatGPT.

Further, those tests show a degree of diversity in quality representing the extent to which the tests fulfil the principles of good test writing, with two tests categorized as excellent, two good, two fair, and one poor. In each category, defects that vary in severity are found, from minor errors to major issues. Tests in the excellent and good categories still contain minor errors that do not significantly affect the overall quality of the tests. In contrast, those in the fair and poor categories contain more noteworthy errors. The errors include unclear instructions, writing mechanics errors such as grammar, spelling, and punctuation, originality issues, correct answers that can be more than one for objective test types, ambiguity, redundancy, invalid items, poorly structured test formats, as well as minimal HOTS questions dominated by LOTS questions at recognition and recall level. Due to these errors, revisions are needed to improve the quality of the tests.

Teachers in this study reveal some of the difficulties they experience in the test construction process. They are aligning the level of difficulty of the items to match the cognitive level of students, creating HOTS questions, lack of standardized references for test construction, creating accurate blueprints, developing effective multiple-choice distractors, designing good quality stimuli (especially for HOTS questions), writing literacy-based tests, evaluating the validity and reliability of the tests they created, lack of access to PISA-type question banks as a reference source, designing contextual, innovative and up-to-date question items.

Based on the situations described, the discrepancy between teachers' knowledge and their actual practices in constructing summative test items cannot be attributed solely to a lack

Enggar Pangesti Wibowo, 2025
INVESTIGATING EFL SECONDARY TEACHERS' ASSESSMENT LITERACY: THEIR KNOWLEDGE AND PRACTICE IN TEST CONSTRUCTION
Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

of assessment-related knowledge. This gap is also shaped by policy constraints that generate significant time limitations. Thus, teachers do not have sufficient time to develop test blueprints, write high-quality items, pilot the questions, or conduct methodical reviews. Addressing these challenges requires strategic interventions that not only enhance teachers' assessment literacy but also reform policy frameworks and optimise time allocation, thereby enabling teachers to engage in all stages of assessment construction with the necessary thoroughness.

Practical trainings in test construction are needed to equip teachers with the ability to develop better quality tests and as a solution to the challenges they face. Moving forward, it is expected that teachers will not only be able to work individually but can also engage in collaborative practices with colleagues and actively participate in the construction of test instruments together within the scope of the school and the teacher community to develop assessment instruments that are more valid, reliable and in accordance with learning needs. In addition, more systematic and comprehensive teacher training education programs that address teachers' workload and teachers' competence should be conducted.

5.2. Implications

The implications of this study relate to how the research findings can be implemented or contribute to the development of theory and practice in assessment. The results of this study are expected to be constructive input for teachers, prospective teachers, and policy makers in designing strategies to improve teacher's assessment literacy.

Theoretically, the results of this study reinforce the findings in the previous literature review that there is a significant gap between teachers' theoretical knowledge of assessment literacy and its implementation in practice. Then it is clear that conceptual understanding alone is not enough to guarantee the quality of test construction, so the theory of assessment literacy needs to be developed by considering the context of its application in the field more realistically.

Furthermore, there are implications for strengthening assessment literacy theory in classroom or school-based assessment settings. This research shows that assessment literacy theory, which has been developed in the context of generic and systemic (macro) education, needs to be recontextualised to be more relevant to the real conditions of EFL teachers in

Enggar Pangesti Wibowo, 2025
INVESTIGATING EFL SECONDARY TEACHERS' ASSESSMENT LITERACY: THEIR KNOWLEDGE AND PRACTICE IN TEST CONSTRUCTION
Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

classroom-based assessment practices. Supposing that administrative pressures faced by teachers, such as limited time to develop quality test instruments and lack of access to adequate assessment training, require a theory that not only emphasises the ideal, but is also flexible and adaptive to existing limitations. The flexibility must still be accompanied by efforts to maintain the quality of the tests, for example through a minimum checking mechanism by colleagues or involving the exam committee and school leaders as validators. In addition, flexibility in the use of various sources of tests, such as textbooks and the internet, needs to be accompanied by a modification process so that the material used is relevant to the needs and abilities of students. Thus, it is not just copying as a whole without adjustment.

Moreover, the finding of a mismatch between teachers' claimed knowledge and practice and the teacher-made tests indicates that it is not enough to evaluate teachers' assessment literacy solely through questionnaires or interviews. Assessing the quality of real products such as blueprints and tests produced is important to build a new theoretical understanding of how assessment literacy can and should be assessed.

Practically, continuous hands-on training is needed, not only theoretical training, which focuses on the technical aspects of test construction, such as blueprints writing, HOTS question preparation, test validation, and the use of contextual and innovative stimuli. The following one is on the development of collaborative models in the local teacher association forum. Collaborative activities like peer review, joint test writing, and joint testing need to be revived. This collaboration can serve as both a shared learning environment and a quality control tool for the evaluation instruments created. Another crucial aspect concerns the provision of resources and access to standardized references. It is essential for the government and relevant stakeholders to ensure that teachers have easy access to those references.

Thereafter, schools and education supervisors should establish an academic supervision system that goes beyond administrative compliance by evaluating the substantive quality of teacher-constructed tests. This is accomplished by means of reviewing the blueprints alignment, analyzing cognitive levels, and assessing the extent to which competencies are addressed. In addition, assessment literacy should be an integral part of teachers' continuous professional

Enggar Pangesti Wibowo, 2025
INVESTIGATING EFL SECONDARY TEACHERS' ASSESSMENT LITERACY: THEIR KNOWLEDGE AND PRACTICE IN TEST CONSTRUCTION
Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

development programs through formal training and learning communities. Thus, teachers are able to situate theoretical understanding with field practice in designing assessments.

5.3. Recommendations

The researcher did not adequately probe into some of these issues. First, this study's coverage area is limited to two provinces, West Java and Central Java, and does not even include all cities within them. This reflects the limited generalisability of the findings to English language teachers across Indonesia. In addition, the study's focus only covers junior and senior high school levels, so it does not represent assessment practices at other education levels.

Next, this study did not involve teachers' direct observation of the test construction process despite using three data instruments: questionnaires, semi-structured interviews, and teacher-made tests. Therefore, some important aspects, such as the duration of preparation time, writing patterns, and teacher considerations in writing test items, were not explored in depth. The analysis of the teacher-made tests was done interpretatively based on the criteria from the experts, which means that the assessment of the tests contains an element of subjectivity and can be different when done by other assessors. This study also did not use a quantitative approach to measure the level of validity, reliability, or difficulty of the questions statistically, which, if done, would provide a more objective assessment and help researchers obtain clear and measurable information.

Finally, only teachers who claimed to have a high understanding of their knowledge and practice in test construction were further analyzed, while this study did not thoroughly reveal the possibility of teachers with low claims producing higher-quality tests.

Based on the abovementioned limitations, several recommendations are made for future research on similar topics. Future research is expected to broaden the area of exploration of teachers' assessment literacy in terms of knowledge and practice to other regions in Indonesia to obtain a more comprehensive picture of English language teachers' assessment literacy nationally. In addition, it is also important to expand the level of education studied, such as primary school and higher education, to provide a more extensive coverage of the findings.

Enggar Pangesti Wibowo, 2025
INVESTIGATING EFL SECONDARY TEACHERS' ASSESSMENT LITERACY: THEIR KNOWLEDGE AND PRACTICE IN TEST CONSTRUCTION
Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

To produce further exhaustive results, it is recommended that future research involve teachers' direct observations of the test construction process. Thus, researchers will be able to verify the procedures and steps taken by teachers in developing test instruments. Furthermore, quantitative analysis using software such as SPSS or ITEMAN can be carried out for certain test types to test the items' validity, reliability and difficulty level statistically.

Lastly, further research is also recommended to include participants who, on the questionnaire, show low levels of claims (e.g. those who answer principally 'disagree', 'strongly disagree', and also 'neutral') related to assessment knowledge and practices. Besides, respondents who choose a neutral answer need to be explored further to identify their tendency, whether they are more inclined to understand or not understand the principles of assessment. These participants could be interviewed, and their test products could be analysed to see if their low claims aligned with the results of the tests they created or if they showed better results than expected.

Enggar Pangesti Wibowo, 2025
INVESTIGATING EFL SECONDARY TEACHERS' ASSESSMENT LITERACY: THEIR KNOWLEDGE AND PRACTICE IN TEST CONSTRUCTION
Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu