

DOI: <https://doi.org/10.20372/JSSHR.V2.i2.05>

ISSN: 2226-7522 (Print) and 2305-3372 (Online)

Journal of Social Science & Humanities Research

J. Soci. Sci. & Hum. Res., July – Dec., 2026, 2(2), 57-76

Journal Homepage: <https://journals.wgu.edu.et>

Original Research

Developing and Validating Four Resource Model-Based Critical Literacy Assessment Instruments for EFL Learners

 Ashenafi Shimeles^{1,*}, Dereje Tadesse², Abera Admasu³, & Alemayehu Getachew⁴

¹Ph.D Candidate at School of Foreign Language and Journalism, Haramaya University, Haramaya, Ethiopia.

²Associate Professor, School of Foreign Language and Journalism, Haramaya University, Haramaya, Ethiopia.

³Assistant Professor, School of Foreign Language and Journalism, Haramaya University, Haramaya, Ethiopia.

⁴Assistant Professor, School of Foreign Language and Journalism, Haramaya University, Haramaya, Ethiopia.

Abstract	Article Information
<i>This research was done using the Four Resource Model as a framework to validate tools for assessing literacy awareness in English learners. The study employed a mixed methodology, integrating qualitative elements with quantitative validation, and focused on fourth year English students at Oda Bultum University. Tools, including a questionnaire, observation checklist and test items, were developed for the purpose of assessing the level of literacy and tools' worthiness in reading classroom. A preliminary study of 16 participants demonstrated the reliability of the questionnaire and Cronbach's alpha of .71 indicated initial feasibility. The results underlined the importance of the Four Resource Model as a theoretical basis for the design of assessments and to promote literacy that goes beyond decoding to include socially-contextualized and critically engaged practice. From the teaching point of view, the study highlighted the need for context-sensitive tools that not only assess but also improve the ability of learners to analyze texts, thus facilitating the transformation of English as Foreign Language learning. More research into long-term methods may be needed to bridge this gap and to check the long-term consequences of theoretical and pedagogical views in the English as a Foreign Language context.</i>	Article History: Received: 05-04-2026 Revised : 26-06-2026 Accepted : 09-09-2026 Keywords: assessment, critical literacy, foreign language, literacy assessment, pedagogy
	*Corresponding Author: Ashenafi Shimelis E-mail: ashenafishimelis1@o bu.edu.et

Copyright@2026 JSSHR, Wollega University. All Rights Reserved.

INTRODUCTION

English has solidified its position as an international language in the rapidly changing field of international communication, serving as a crucial tool for social, professional, and

academic interactions. In this context, English as a Foreign Language (EFL) plays a vital role in equipping students with both linguistic proficiency and the skills necessary to engage

critically with texts and societal issues. While traditional language instruction often focuses on grammar, vocabulary, and reading comprehension, modern pedagogical approaches advocate for a broader emphasis that includes critical literacy. This approach encourages students to analyze, challenge, and question the social, political, and ideological constructs present in texts (Luke, 2018; Bobkina & Stefanova, 2016; Bacon, 2017).

There are specific challenges in assessing these intricate literacy skills. Traditional assessments often fail to recognize students' ability to critically analyze texts and understand their societal implications, instead focusing on decoding and comprehension. As a result, there is a growing demand for assessment tools that can effectively measure critical literacy skills, especially in EFL settings where linguistic and cultural disparities may impact interpretive and analytical capabilities (Papadopoulos and Griva, 2019; Silvhiany, Huzaifah, & Ismet, 2020).

Rationale for the present initial tools assessment

For a number of reasons, the development of valid and trustworthy critical literacy assessment instruments in English contexts is crucial. First, by providing teachers with useful information about students' levels of literacy engagement, these tools can guide focused instructional interventions (Silvhiany, Huzaifah, and Ismet, 2020). Second, by aligning assessment procedures with educational objectives, they can foster a classroom environment that promotes social awareness and literacy skills (Papadopoulos & Griva, 2017). Third, current assessments often fail to fully capture the complexity of critical literacy, highlighting the need for the creation of theoretically sound and contextually

relevant tools (Luke & Freebody, 1999). There is still a dearth of research on creating and assessing assessment instruments specifically for English students, despite the acknowledged significance of critical literacy. Most modern instruments either do not have a solid theoretical basis or are adapted from native-language settings rather than our national context.

The core ideas of Luke's and Freebody's Model of Critical Literacy include four key roles for readers (Appendix 1). The first is the Code Breaker, who deciphers words and sentences. Next is the Text Participant, who makes meaning from what they read. The third role is the Text User, who understands the purpose of reading. The last is the Text Analyst, who looks at the social, cultural, and political ideas in the text.

This alignment corresponds with the specific research objectives (Luke, 2013; Shin and Crookes, 2005). Consequently, the tools did not emphasize the inquiries necessary for effectively managing the study and alleviating students' uncertainties while completing the test and the closed-ended questionnaire.

The purpose of this initial study was to develop tools intended to gauge English language students' literacy abilities. The study sought to create authentic, comprehensive tools that reflect the intricate aspects of literacy, informed by the Four Resource Model and its theoretical framework

By offering assessment tools that can be used to improve literacy awareness and social awareness in English reading classrooms, the main goal was to advance both theoretical understanding and pedagogical practice. Although previous reviews have begun this work, none of them have focused on the full array of literacy awareness that can be addressed during the effects of literacy skills

for English language students at the university level and also utilized the Four Resource Model instruction training to rigorously examine reliability and validity. Therefore, the present initial tools assessment sought to address this gap in the literature by stating the following questions: How is assessment tools designed to measure literacy awareness among English learners within a theoretical framework? How valid and reliable are the developed assessment tools for evaluating literacy awareness in an English Reading context? And how do the assessment results inform pedagogical practices aimed at fostering critical literacy among English students?

This research is significant for educators, curriculum developers, and language education researchers. It improves scholarly understanding of integrating critical literacy assessment into language teaching and emphasizes the need for contextually relevant and multidimensional assessment approaches. The study offers a theoretically sound and pedagogically useful assessment framework to evaluate and enhance critical literacy in English language learners. Additionally, it provides insights for curriculum design, instructional approaches, and future research on critical literacy assessment, language proficiency, and social consciousness in various educational contexts.

The development of effective assessment tools in language learning requires a solid grounding in relevant theoretical frameworks that explain what literacy is, how people learn, and how they are taught (Gee, 2004; Luke and Freebody, 1999; Luke, 2000; Janks, 2013). This chapter provides a review of some important theoretical stances relevant to the present study and their pedagogical implications in English Language Reading

Contexts (Gee, 2004; Luke and Freebody, 1999; Luke, 2000; Janks, 2013).

The importance of understanding literacy as a complex, multidimensional concept has long been emphasized by theorists. Early literacy models primarily focused on decoding and understanding skills (Morrell, 2015), but modern views recognize the social, cultural, and political aspects inherent in literacy development (Luke, 2000). Theoretical frameworks have developed to guide pedagogical practices and evaluation strategies that promote critical engagement, as literacy increasingly incorporates social criticism and empowerment.

The transition from traditional to critical literacy paradigms emphasizes the importance of students' capacity to analyze texts, challenge underlying ideologies, and actively participate in social justice issues (Gadotti and Torres, 2009; Luke and Freebody, 1999). This change affects the design of assessments, necessitating the use of tools that gauge critical analytical abilities in addition to literacy awareness in the text.

Critical Literacy and the Four Resource Model

Defining Critical Literacy

The ability to examine, evaluate, and alter social realities through texts is a component of critical literacy, which goes beyond conventional ideas of reading comprehension (Luke and Freebody, 1990). It entails identifying and comprehending the power dynamics, ideologies, and biases present in texts that contribute to social change. Particularly in diverse and globalized contexts where students are exposed to texts with social and political undertones, critical literacy is seen as a crucial part of literacy education.

Studies by Shor and Freire (1987) and Luke (2012) have shown that critical literacy

training improves students' capacity to become engaged citizens who can challenge prevailing narratives and promote social change. Therefore, assessment instruments that align with the principles of critical literacy must measure students' ability to critically examine texts and connect textual analysis to broader social issues.

The Four Resource Model of Critical Literacy by Luke and Freebody

The Four Resource Model developed by Luke and Freebody (1999) offers a comprehensive framework for understanding and enhancing literacy skills in a social context. According to this model, proficient readers engage in four interconnected roles when interacting with texts: Code Breaker, which involves decoding language, understanding vocabulary, and interpreting grammatical structures. This role aligns with traditional literacy skills. Meaning Maker, which involves deriving meaning from texts through comprehension, inference, and contextual interpretation.

Text user, which involves applying comprehension of texts within specific social practices and recognizing how texts function in particular contexts. Text Analyst, which involves critically examining texts for underlying ideological messages, power dynamics, and social constructs. According to this framework, literacy is a dynamic, socially situated process that requires students to transition fluidly between these roles depending on their reading context and purpose.

Pedagogical Implications of the present initial tools assessment

The assessment results, based on the initial tools developed for this study, emphasize the pedagogical importance of Luke and Freebody's Four Resources Model in

promoting critical literacy among English language learners. With Cronbach's Alpha values ranging from .71 to .86, the instruments demonstrated acceptable to high reliability, providing reliable data that assessed students' performance in decoding, making meaning, using text, and analyzing text. Specific tasks within the assessment reflected these dimensions: vocabulary-focused items addressed the role of code breaker, comprehension and interpretation questions were aligned with the meaning maker, functional application tasks corresponded to the text user, and critical evaluation items targeted the text analyst. Refinement of items with low correlations further enhanced the internal consistency of the tools, ensuring they effectively measured these layered competencies. This evidence-based approach is in line with the work of Luke and Freebody (1999), Luke (2000), and Janks (2010), who stress the importance of thorough evaluation in ensuring learners acquire the skills needed to actively participate in societal discourse. The contextual evidence from the pilot study shows that validated assessment instruments not only assess students' literacy awareness and active participation but also assess teachers in developing pedagogical practices that help learners move beyond surface comprehension toward reflective, analytical, and socially aware engagement with texts.

Application in English as Foreign Language Contexts

Within the field of English as a Foreign Language (EFL), the Four Resources Model has been widely recognized as a framework that supports the concurrent development of linguistic and literacy skills. Luke and Freebody (1999), Luke (2000), and Janks (2010) emphasize that effective literacy instruction must engage learners as code

breakers, meaning makers, text users, and text analysts, moving beyond surface comprehension toward critical engagement with texts. Empirical studies further demonstrate that assessment tools grounded in this model promote deeper interaction with culturally and socially relevant texts, encouraging learners to decipher complex vocabulary, apply their understanding to social contexts, and critically evaluate ideological messages (Shin & Crookes, 2005). These findings align with communicative language teaching principles, which prioritize meaningful interaction and critical thinking as central to language learning.

In the pilot study conducted at Oda Bultum University, the incorporation of this model into assessment design provided evidence of its pedagogical value. Reliability and validity analyses confirmed that the tools could capture students' layered competencies across the four roles. These results reinforce the broader literature by illustrating how validated instruments informed by the Four Resources Model can serve as diagnostic tools for identifying learner strengths and weaknesses, while simultaneously guiding instructional practices that foster reflective, analytical, and socially aware literacy development in EFL contexts.

Empirical Studies and Program Implementations

Empirical evidence consistently demonstrates the effectiveness of the Four Resources Model in promoting critical literacy in English as Foreign Language (EFL) contexts. Instructional strategies based on this framework help learners analyze texts for social biases, interpret power dynamics, and generate critical responses (Luke, 2013). In EFL settings, tasks centered around the four roles have been shown to enhance social

awareness and engagement (Shin & Crookes, 2005). The reliability of assessment tools aligned with the model has also been confirmed, with Brown & Alford (2023) reporting strong internal consistency ($\alpha = 0.92$), validating their ability to measure comprehension and analytical skills.

Similar results were found by Saito (2022), who showed that refining assessment items increased reliability ($\alpha = 0.82$) and improved the accuracy of assessing critical literacy competencies. In higher education settings, Lee et al. (2025) noted significant enhancements in students' ability to analyze texts and engage in reflective discussions when critical literacy instruction was incorporated into classroom activities. Recent studies, such as Ma, Crosthwaite, Sun, & Zou (2024), have also suggested that exposure to text analysis from various critical perspectives led to noticeable improvements in learners' capacity to construct arguments supported by textual evidence. Overall, these findings underscore the educational value of the Four Resources Model, demonstrating that assessment and instruction grounded in its principles not only aid in language development but also nurture reflective, socially conscious, and critically engaged learners in EFL environments.

Challenges and Considerations

Despite providing a strong framework, the Four Resource Model presents challenges when used in assessment design. Subjective judgments are often required in assessing literacy analysis and social application, so it is important to have multiple evaluators and clear rubrics to ensure reliability. Additionally, contextualized assessment practices are crucial because learners' cultural backgrounds influence how they interpret information (Luke, 2012).

METHODOLOGY

The chapter describes the research design, participants, tools, processes, and data analysis techniques used in developing and assessing literacy assessment tools for English as Foreign Language learners. To ensure the findings' theoretical and pedagogical relevance, the methodological framework is based on both qualitative and quantitative approaches.

The researcher utilized a mixed-methods research design, incorporating quantitative techniques for validation and reliability testing and qualitative processes for instrument development. This approach allowed for a comprehensive assessment of the instruments, ensuring they were both practically and theoretically sound in

evaluating literacy skills among English as a Foreign Language students at Oda Bultum University (Creswell and Plano Clark, 2018; Dörnyei, 2007; Johnson and Onwuegbuzie, 2004). Phase 1 involved qualitative development of the assessment tools, including item creation based on theoretical frameworks and professional advice. Phase 2 included reliability testing, item analysis, and pilot testing as part of the quantitative validation process.

Participants**Selection Criteria**

The participants were fourth-year English as a Foreign Language students enrolled in the English department at Oda Bultum University (table 1).

Table 1*Demographics of Participants*

Variable	Description/Range	Notes on Representation
Sample size	11-19	Total number of participants
Gender	Male and female	Reflective of population distribution
Age Range	20-25	Typical for fourth-year undergraduates
Academic Level	Fourth year EFL students	Advanced course work completed
Program Status	Currently enrolled	Ensured curricular consistency

They were purposefully chosen based on their enrollment in the English as a Foreign Language program at the time of data collection, their availability, their registration with second-year students in the same program, and their voluntary consent. This ensured curriculum consistency, academic experience, linguistic proficiency, and ethical compliance. Nineteen students, both male and female, between the ages of twenty and twenty-five took part. Demographic data was gathered to

provide context for the results. The distribution of age and gender was considered typical of Oda Bultum University fourth-year English as Foreign Language participants.

The study group consists of 11-19 fourth-year English as a Foreign Language students, aged 20–25, who are currently enrolled (table 1). They represent both male and female learners. Their uniform curriculum and advanced academic progress, with the

majority of coursework completed, make them ideal participants for the study.

Instruments Development: theoretical foundation

A set of test items, an observation checklist, and a structured questionnaire were used as assessment tools in this study. Luke and Freebody's Four Resource Model of critical literacy (1999) provided the general theoretical framework, which was carefully matched with these tools. An extensive review of relevant literature and previously validated instruments informed item development to ensure content relevance and consistency with this framework. Various formats, including true/false statements, multiple-choice questions, short-answer, summary tasks, and textual analysis prompts, were incorporated into the examination, in addition to the 42 closed-ended items in the questionnaire.

The Four Resource Model describes four roles: Code Breaker, Meaning Maker, Text User, and Text Analyst. The assessment tools were specifically designed to operationalize these roles. This model has gained widespread recognition as a comprehensive framework for literacy instruction due to its incorporation of linguistic decoding, meaning construction, functional application, and critical analysis into a cohesive pedagogical approach (Luke, 2000; Freebody, 2007; Janks, 2013). Such a framework is particularly crucial in the context of English as a Foreign Language, as it allows students to move beyond simple comprehension and engage in more in-depth interpretive exercises, contextual applications, and critical analyses of the ideological stances present in texts. By firmly grounding the assessment tools in this theoretical framework, the design was made both theoretically sound and pedagogically applicable, providing

English as Foreign Language students with a reliable way to assess their literacy abilities.

Limitations of the Methodology

The study had several potential limitations, including a limited sample size, the specific context which might limit the generalizability of the findings, and the subjective nature of assessing literacy skills. Small sample sizes are often recognized as reducing statistical power and external validity (Creswell, 2018; Dörnyei, 2007). Additionally, conducting the study within a single institution could limit its applicability to wider English as a Foreign Language context, where literacy practices are influenced by sociocultural factors (Gee, 1996; Street, 2003).

The interpretive aspect of literacy also presents difficulties in defining constructs like ideological critique, which may introduce bias in the design and scoring of assessment items (Luke, 2000; Janks, 2010; Comber, 2015; Vasquez, 2017). To address these limitations, the study employed purposive sampling, used multiple instruments for triangulation, and aligned with Luke and Freebody's Four Resource Model, thereby improving the study's validity, credibility, and relevance for teaching (Freebody, 2007; Johnson & Onwuegbuzie, 2004; Lincoln and Guba, 1994; Creswell & Plano Clark, 2018).

FINDINGS AND DISCUSSIONS

Introduction

This chapter provides a detailed exploration of the study findings related to the design, validation, and practical application of assessment instruments in English as a Foreign Language (EFL) reading contexts. The analysis is conducted both descriptively and through the lens of Luke and Freebody's Four Resource Model, offering a comprehensive understanding of how these tools support

English learners. The insights gained contribute to improving assessment practices in an English language context.

FINDINGS

In a standard educational approach, Allan Luke and Peter Freebody's critical literacy framework emphasizes empowering learners to actively interrogate texts rather than passively consume them (Appendix 1). Their Four Resource Model portrays four essential roles that students should develop to become competent and critical readers. The first role, Code Breaker, focuses on decoding and understanding the language, symbols, and conventions used in texts. The second role, Text Participant, involves constructing meaning by making connections and interpreting content. The third role, Text User, emphasizes applying knowledge from texts to real-world situations and personal contexts. Lastly, the Text Analyst role encourages learners to critically evaluate and question the messages, motives, and biases embedded within texts (Appendix 1). This comprehensive framework aims to foster not only literacy skills but also critical thinking and awareness, enabling students to navigate and challenge the multitude of texts they encounter daily.

In their fourth year at Oda Bultum University, English students undertook a significant task by piloting new assessment tools aimed at evaluating literacy levels in an English reading context. Throughout this process, they actively provided detailed feedback on the clarity, relevance, and

practicality of these instruments. Their academic background and coursework equipped them with the necessary skills to critically analyze and suggest improvements, which in turn helped refine the assessment tools. This repetitive refinement ensured that the instruments were not only theoretically sound but also pedagogically effective in real English classroom settings. As a result, the validity of the assessments was enhanced, highlighting the crucial role of learner perspectives. This experience emphasized the importance of involving students in assessment design to better align with their needs and promote more meaningful literacy skills learning. To quantify the initial assessment tools' outcomes for the main study, the researcher analyzed the results using Cronbach's Alpha (table 2).

Table 2

Reliability and Validity of the Closed-Ended Questionnaire Items

Reliability Statistics	
Cronbach's N of Alpha Items	
.71	42

The reliability and validity findings of the closed-ended questionnaire have important theoretical and pedagogical implications within Luke and Freebody's Critical Literacy framework (table 2). A Cronbach's Alpha of 0.71 across 42 items indicates acceptable internal consistency, suggesting the instrument reliably measures students' critical engagement with texts in English as a Foreign Language

reading classroom. This supports the tool's validity within the literacy paradigm, aligning with Luke and Freebody's four resources model (code-breaking, meaning-making, text-making, and text-solving). The coherence among items demonstrates that the instrument effectively captures these interconnected dimensions, providing a credible basis for analyzing students' critical literacy level and reflection.

Testing with fourth-year students before second-year students exemplifies reflective practice. Refining questions for clarity enhances the instrument's accuracy in

assessing the development of literacy skills in English as a Foreign Language reading classroom, reducing misinterpretation and fostering deeper engagement.

The findings suggest room for improvement such as reviewing items for redundancy to achieve higher reliability (above 0.80). This would improve sensitivity to refinement differences in students' literacy level, supporting more targeted pedagogical strategies. Overall, ongoing validation and refinement are essential to effectively foster literacy awareness and literacy in English as Foreign Language contexts.

Table 3

Distribution of Descriptive Statistics Pilot Test Score

Descriptive Statistics					
	N	Min	Max	Mean	Std. Deviation
Pilot Test Score	16	36	59	45.8	5.83
Valid N (list wise)	16				

The descriptive statistics of the pilot test scores showed that all 16 valid cases were included in the analysis (table 3). Scores ranged from a minimum of 36 to a maximum of 59, with a mean of 45.8 and a standard deviation of 5.83. These findings suggest moderate variability in student performance, with most

scores clustering around the mean but still reflecting individual differences across the sample. The inclusion of all valid cases enhances the reliability of the reported measures and establishes a solid foundation for subsequent analyses in line with the study's objectives.

Table 4

Frequency Distribution of Pilot Test Score

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 36	1	4.0	6.3	6.3
40	2	8.0	12.5	18.8

Table 4 Continues

41	2	8.0	12.5	31.3
44	2	8.0	12.5	43.8
45	1	4.0	6.3	50.0
46	2	8.0	12.5	62.5
47	1	4.0	6.3	68.8
49	1	4.0	6.3	75.0
50	1	4.0	6.3	81.3
52	1	4.0	6.3	87.5
53	1	4.0	6.3	93.8
59	1	4.0	6.3	100.0
Total	16	64.0	100.0	
Missing System	9	36.0		
Total	25	100.0		

The distribution shows that valid responses were distributed among various values, with the most frequent responses recorded at 40, 41, 44, and 46 (each contributing 12.5% of valid responses). Lower frequencies were seen at values like 36, 45, 47, and others, each accounting for 6.3%. In total,

16 valid responses (64.0%) were collected, while 9 cases (36.0%) were missing (table 4). This suggests that while most participants provided usable data, the presence of a significance number of missing responses should be taken into account when interpreting the results.

Table 5

Feasibility, Implementation, Clarity, and Understanding of observation items (Student Perspective: Observed Behaviors)

Main Theme	Items Included
Student Comprehension & Engagement	1. Vocabulary understanding
	2. Understanding instructions
	3. Navigating format
	4. Completing all items (skipped questions)
	5. Engagement while completing
	6. Signs of confusion/frustration
	7. Peer consultation due to difficulty
	8. Decoding text (Breaking the Code)
Assessment of Literacy Skills	9. Meaning Making
	10. Text Using (functional literacy)
	11. Text Analyzing (critical literacy)
Test Appropriateness & Coverage	12. Grade-level reflection
	13. Comprehensive assessment of literacy skills

Main Theme	Items Included
<i>Table 5 Continues</i>	
Administration & Logistics	14. Time sufficiency
	15. Physical format (font size, layout)
	16. Administration procedures clarity
Learning Environment	17. Classroom environment conducive to completing the test

The pilot study showed that the assessment tools effectively captured key literacy dimensions, including decoding, meaning-making, text use, and text analysis. Observational data indicated that students comprehended test instructions and formats, completing all items independently, confirming the instruments' clarity and accessibility (table 5). While the tools aligned with Luke and Freebody critical literacy model and the study's goals, the lack of explicit evidence on student engagement identified a need for further investigation. Practical factors like time constraints, clear formats, administration procedures, and classroom environments validated the feasibility of incorporating these assessments into practice. Overall, the results emphasize the importance of refining assessment design and instructional methods to better support diverse learners in English as Foreign Language settings.

Key Findings

Feasibility: The administration of the assessment tools showed practicality and efficiency, as most students completed the questionnaires and tests within the allocated time frame. From a pedagogical perspective, this indicates that the assessment design aligns well with classroom constraints and can be smoothly integrated into instructional routines. However, some participants had difficulty understanding certain items, suggesting a need

for further refinement of item clarity and instructional support (table 2-4). Theoretically, this underscores the importance of designing assessments that are accessible and inclusive, ensuring that all students can engage meaningfully with critical literacy tasks as emphasized in Luke and Freebody's framework.

Acceptability: The majority of students perceived the assessment instruments as relevant and appropriate within their educational context. Their positive response to the variety of test items, especially the innovative approaches to text analysis, reflects the pedagogical value of incorporating diverse, engaging tasks that foster critical engagement with texts. Nonetheless, some items particularly those requiring contextual interpretation induced confusion, pointing to the necessity for clearer, more straightforward instructions. From a theoretical standpoint, this underscores the significance of aligning assessment tasks with students' linguistic and cognitive levels to effectively measure critical literacy skills without causing unnecessary frustration or misunderstanding.

Potential Challenges: Several challenges emerged during the pilot, notably time constraints and response inconsistencies. These issues suggested that certain items lacked sufficient clarity or failed to fully align with the objectives of Critical Literacy Instruction as outlined by Luke and Freebody.

Pedagogically, this emphasizes the need for careful item design that considers cognitive load and ensures tasks are comprehensible and appropriately challenging. Theoretically, it calls for ongoing refinement of assessment tools to better capture the processes involved in critical literacy development, such as decoding, meaning-making, and text analysis, to support more accurate evaluation and targeted instruction.

The limitations identified in the pilot study have important theoretical and pedagogical implications within the framework of Critical Literacy Instruction by Luke and Freebody, particularly in the EFL reading context (Appendix 1).

From a theoretical perspective, the limited sample size of 16 participants constrained the generalizability of the findings, thereby limiting the extent to which conclusions about students' critical literacy development could be applied to the broader student population at Oda Bultum University. This underscores the need for further research with larger, more diverse samples to develop a comprehensive understanding of how critical literacy concepts are perceived and enacted across different demographic groups. Additionally, the homogeneity of the sample might have restricted the exploration of how various socioeconomic, cultural, or linguistic backgrounds influence critical literacy engagement, which is essential for developing inclusive pedagogical strategies aligned with Luke and Freebody's model.

Pedagogically, the predominance of quantitative data and the reliance on closed-ended questionnaire items limited the depth of insights into students' experiences, perceptions, and challenges related to critical literacy instruction. This highlights the necessity for incorporating qualitative

approaches such as interviews or open-ended responses to enrich understanding of learners' perspectives and inform more refined instructional designs. The variability in responses, particularly across demographic groups suggested potential biases in assessment tools and indicated the importance of continually refining instruments to ensure cultural and contextual relevance, clarity, and appropriate difficulty levels.

Furthermore, issues related to item difficulty and time constraints underscore the importance of carefully designing assessment tools to accurately measure critical literacy skills without causing frustration or fatigue. Aligning assessment difficulty with learners' literacy levels could enhance the validity of evaluations and support targeted pedagogical interventions. The variability and interpretative challenges in data also emphasize the need for further testing and refinement of assessment instruments to ensure they effectively capture critical literacy development over time.

In summary, addressing these limitations through ongoing cycles of testing and refinement is essential to establish reliable, valid, and inclusive tools for evaluating critical literacy in the EFL classroom. Such efforts would support the development of pedagogical practices that are responsive to diverse learner needs, ultimately fostering more effective critical literacy instruction grounded in Luke and Freebody's theoretical framework.

Discussions

The tools construct validity and acceptable reliability ($\alpha = .71$) align with recognized benchmarks in literacy assessment, where values above .70 are considered adequate for early-stage instrument development (Polit and Beck, 2010; Al-Osail et al., 2026). Similar research has shown that item refinement improves reliability; Sato, Mutoh, and

Kleinsasser (2022), and Watkins (2018) reported enhanced consistency through targeted changes, supporting Messick's (1996) theory that validity is supported by reliability.

In terms of pedagogy, the tool's incorporation of Luke and Freebody's Four Resources Model aligns with research by Smith, Jones, Gilbert, & Wieman (2013) and Wang & Wang (2024), who emphasize that organized frameworks promote analytical, interpretative, and reflective abilities. Taken as a whole, these findings support more general evidence that critical literacy-based instruction in EFL contexts requires the use of verified, theoretically based tools.

Luke and Freebody's Four Resources Model positions literacy as a multifaceted practice that extends beyond decoding to interpretation, functional use, and critical analysis of texts. This aligns with critical literacy perspectives that emphasize literacy awareness and examination of social functions and power relations. Strong reliability ($\alpha = 0.87\text{--}0.92$) supports effective assessment of learner literacy awareness and text critique skills, according to validation studies (Taherdoost, 2016; Watkins, 2018). Research also shows that structured frameworks promote questioning, bias detection, and sociocultural awareness. High reliability ($\alpha = 0.71\text{--}0.86$) and the tool's potential to promote literacy awareness and text evaluation abilities were confirmed by the Oda Bultum University pilot study. However, poor items total correlations and improved diagnostic accuracy need attention.

In line with earlier research by Nijakowska, Tsagari, & Spanoudis (2020); Sato, Mutoh & Kleinsasser (2022), item analysis and the elimination of weak items improved reliability, supporting Messick's

(1995) claim that validity is based on reliability.

Furthermore, research on test designs reveals that Multiple Choice Questions (MCQs) encourage participation and literacy awareness, while short responses add depth but require explicit rubrics. True/false items can put students at risk for guesswork (Razavipour, & Raji, 2022; Smith, Jones, Gilbert, & Wieman, 2013; Banitalebi, Estaji, & Brown, 2025).

Collectively, this data underscores the importance of careful item selection, refinement, and utilizing different formats to enhance reliability, validity, and pedagogical efficacy in EFL literacy assessment.

While the use of critical literacy frameworks in textual analysis strengthens reasoning and evidence-based argumentation (Ma, Crosthwaite, Sun & Zou (2024); Smith, Jones, Gilbert & Wieman (2013)), engagement through summary writing has been demonstrated to improve students' capacity to extract important information, deepen understanding, and develop literacy skills (Kato, 2022). These findings are in harmony with piloting studies of literacy assessment tools, where diverse forms such as summaries, textual analyses, and structured exercises have consistently enhanced literacy awareness and learner engagement. For example, Kormos & Indrarathne (2025) and Sato & Kato (2016) showed that improving item types and varying formats improved validity and reliability, supporting Messick's (1995) assertion that reliable measurement techniques are essential for meaningful research tools evaluation.

Likewise, Watkins (2018) showed that research tools with good internal consistency and easily understood rubrics facilitated better comprehension and literacy awareness, reflecting Razavipour & Raji (2022) caution

that simplistic forms like true/false items risk guesswork and weaker outcomes. Together, these results highlight the significance of careful assessment design in EFL contexts by demonstrating how piloting tools with a variety of well-designed formats not only improves reliability and validity but also fosters the critical literacy skills of interpretation, analysis, and reflection necessary for academic success.

Limitations

This pilot study provided valuable insights; however, it was limited by several factors. Firstly, the sample size was small, with only sixteen participants, which restricted the statistical power and generalizability of the results. This highlights the need for a larger and more diverse participant group in future research. Secondly, the study was conducted within a single timeframe, which may have overlooked variations in student responses over a longer period. A more extended study could provide a more comprehensive understanding of the intervention's effects. Additionally, the subjectivity of participant responses introduced biases that require careful analysis. While multiple-choice and true/false items offer efficiency and objectivity, they may not fully capture literacy skills. Future studies could enhance scoring reliability through rater training and activity sessions.

Using 4th-year students for pilot testing has advantages but also limitations. Differences in academic maturity and experience could influence responses, affecting the generalizability of the results to the 2nd-year population. To address this, the main study included validity checks such as factor analysis to confirm the instrument's applicability to the target group. Furthermore, the small pilot sample size limited the depth of statistical validation. While larger pilot samples are

preferable, they may not always be feasible due to resource constraints. Lastly, contextual differences in the academic environment and student engagement levels may vary between sample populations, influencing responses. Future research should consider cross validating instruments across different student groups to enhance the tools' robustness.

Discussions with Implications for the Full Study Design

The results of the pilot study indicated a promising start in evaluating tools for Critical Literacy Instruction, while also highlighting aspects that require further examination before conducting a more extensive study. The reported acceptance of these instruments suggested their applicability within an EFL context; however, additional clarification on certain items would be essential to enhance comprehension and improve the accuracy of responses.

The insights derived from this pilot study would have several implications for the overall study design:

- 1) *Enhancement of Instruments:* Utilizing feedback from the pilot to improve the clarity of items, particularly those that have caused confusion, is essential to ensure that all participants fully understand the tasks at hand.
- 2) *Sample Expansion:* Increasing the sample size for the primary study to encompass a broader range of participant demographics is crucial for enhancing the reliability and validity of the findings, as well as for improving the representativeness of the collected data.
- 3) *Data Collection Methods:* Incorporating additional qualitative approaches, such as interviews or focus groups, alongside quantitative measures might provide richer insights into students' learning experiences and their development of literacy skills.

The findings from this pilot study provided valuable insights into the effectiveness of the developed questionnaire and various assessment items for measuring reliability and validity when assessing the effects of critical literacy instruction in an English as a Foreign Language (EFL) reading classroom (e.g., Smith & Jones, 2022). While the instruments demonstrated potential, further refinement and validation were necessary to establish their reliability and validity (e.g., Luke, 2000).

These preliminary results have significant implications for academicians aiming to integrate critical literacy approaches into English as a Foreign Language (EFL) reading curricula. This integration could enhance student engagement with texts, foster critical thinking skills, and contribute to the development of informed global citizens (Lopez-Núñez et al., 2024).

CONCLUSIONS

In the initial pilot study conducted with fourth-year English majors at Oda Bultum University, the assessment tools grounded in Luke and Freebody's Four Resource Model of Critical Literacy were found to demonstrate acceptable reliability and validity. Although there is room for further improvement in internal consistency, a Cronbach's Alpha score of approximately 0.71 was interpreted as indicating adequate internal reliability in measuring key literacy dimensions.

The tools were deemed to be generally comprehensible, relevant, and capable of effectively assessing critical literacy skills such as decoding, understanding, application, and analysis during the pilot testing process. In addition, a variety of question formats including true/false, multiple-choice, short answer, summarization, and text analysis were shown to be successful in capturing diverse

literacy skills in alignment with the theoretical framework.

Several limitations were noted: the small sample size of 16 participants was recognized to restrict the generalizability of the findings and limit the statistical power; certain items were observed to lack clarity, indicating that their wording should be revised for better comprehension; and the use of a homogeneous sample was acknowledged to potentially constrain the applicability of the instruments to larger and more diverse student populations. Furthermore, reliance solely on quantitative data was considered to possibly obscure insights into students' difficulties and experiences.

The importance of repetitive testing and continuous refinement of assessment instruments was emphasized as essential for enhancing their validity and reliability. Recommendations for future research included the use of larger and more diverse samples, the integration of more qualitative data to deepen understanding, and the adoption of longitudinal approaches to monitor changes over time.

In conclusion, despite demonstrating promising potential for measuring critical literacy in an EFL context, further testing and improvement are deemed necessary before the tools can be applied in larger-scale research or integrated into curricula. It is believed that such efforts would contribute to more effective instructional practices and promote more critical engagement with texts among learners by ensuring that assessment of critical literacy development is more accurate, reliable, and comprehensive.

List of Abbreviation

EFL: English as a Foreign Language

TEFL: Teaching English as a Foreign Language

CLI: Critical Literacy Instruction

Acknowledgement

We are happy to express our genuine appreciation to those who support me both physically and mentally to the completion of this work.

Author contributions

ASS led the initial study and manuscript drafted; DTB, AAE, and AGT provided continues consultant and contribution to refining and adjusting the manuscript for clarity and accuracy. The teacher-researchers reviewed and approved the final version prior to submission.

Competing interest

The teacher-researchers declare no competing interests.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During the preparation of this manuscript, the teacher-researchers utilized GPT, Grammarly, and Quillbot for language enhancement and grammar corrections, aimed at improving text clarity and flow, refining grammar, and enhancing style; additionally, all AI-generated content underwent thorough review and editing to assure accuracy, coherence, alignment with research objectives, and adherence to the four resource model.

Ethical approval

The research underwent an ethical review and received approval from the Graduate Council of Postgraduate Programs (CPGP) at Haramaya University on November 3, 2024. The initial author Ashenafi Shimeles Shumiye (Ph.D. Candidate, ID No: PHD061/14) received this approval as a component of his doctoral research in the English Language and Literature department. The approval process adhered to ethical standards established to guarantee the research's acceptance and

complied with ethical principles regarding human participation.

Informed consent

A hard copy of informed consent was shown (given) to the classroom before assessment and intervention; and also informed all literacy-related activities and discussions to be collected as part of their weekly performance in the reading course. The teacher-researchers indicated that the students' departure from the intervention might not affect their academic standing; and also ensured confidentiality and anonymity throughout the data collection process.

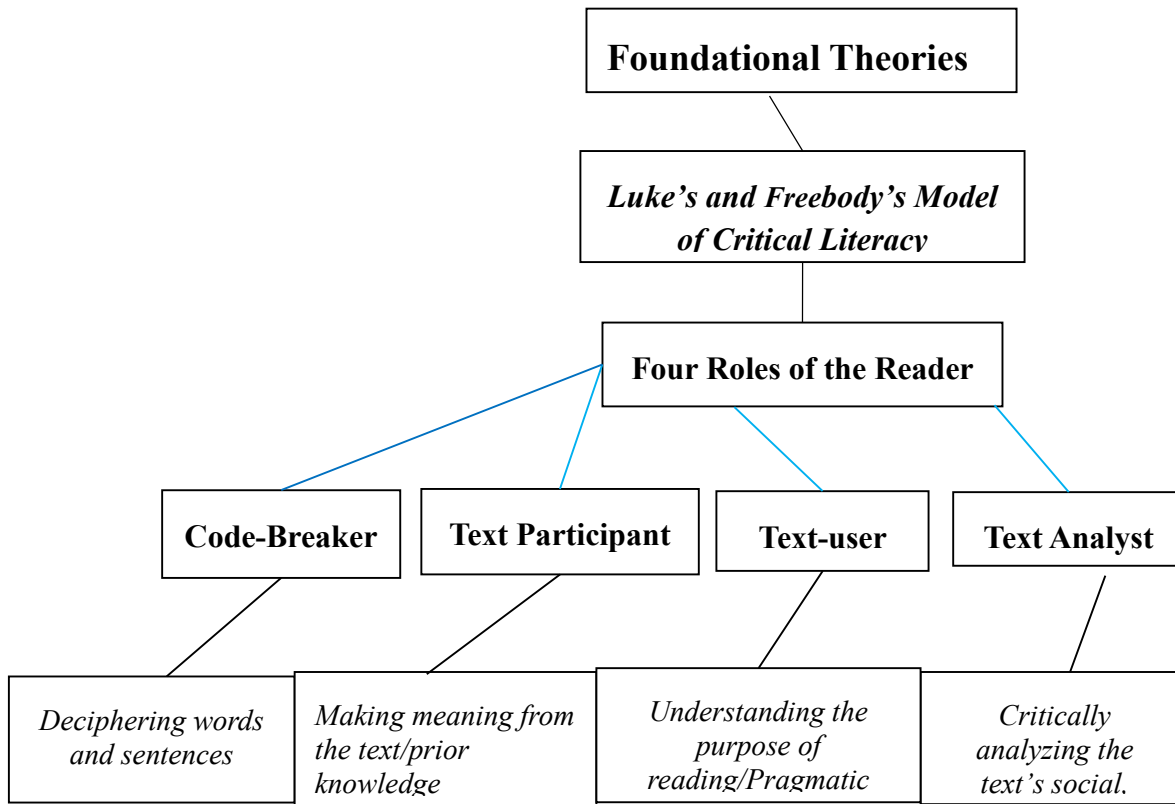
REFERENCES

- Luke, A. (2018). Critical literacy in Australia: A matter of context and standpoint: In *Critical literacy, schooling, and social justice* (pp. 168-188): Routledge.
- Bacon, C. K. (2017). Multilanguage, multipurpose: A literature review, synthesis, and framework for critical literacies in English language teaching. *Journal of Literacy Research*, 49(3), 424-453.
- Griva, E., & Papadopoulos, I. (2019): "MulDiLanguaging" for Fostering Intercultural Communicative Competence in Second/Foreign Language Education: *Language Teaching Research Quarterly*, 11, 54-67.
- Luke, A. (2013): Defining critical literacy. In *Moving critical literacies forward* (pp. 19-31): Routledge.
- Shin, H., & Crookes, G. (2005): Exploring the possibilities for EFL critical pedagogy in Korea: A two-part case study. *Critical inquiry in language studies: An international journal*, 2(2), 113-136.
- Gadotti, M., & Torres, C. A. (2009): Paulo Freire: education for development:

- Development and change, 40(6), 1255-1267.
- Shor, I., & Freire, P. (1987): What is the “dialogical method” of teaching: Journal of education, 169(3), 11-31?
- Janks, H. (2010): Language, power and pedagogies: Sociolinguistics and language education, 1(3), 40-61.
- Brown, C. W., & Alford, J. (2023): Critical Literacy in the Language Classroom: Possibilities for Intercultural Learning through Symbolic Competence. Intercultural Communication Education, 6(2), 36-52.
- Saito, S. (2022): Language Awareness and High School English Teaching in Japan: Taking a Translanguaging Lens to a 'Monolingual' Setting (Doctoral dissertation, Monash University).
- Lee, N. A. A., Tazijan, F. N., Adam, A. F. M., Ikhsanudin, I., & Bakar, R. A. (2025). Comparative analysis of digital literacy and 21st-century skills among university graduates in Malaysia and Indonesia: The role of collaboration, critical thinking, communication, and creativity. Journal of Nusantara Studies (JONUS), 10(1), 166-191.
- Ma, Q., Crosthwaite, P., Sun, D., & Zou, D. (2024): Exploring ChatGPT literacy in language education: A global perspective and comprehensive approach. Computers and education: Artificial intelligence, 7, 100278.
- Gee, J. P. (1999). Critical issues: Reading and the new literacy studies: Reframing the national academy of sciences report on reading: Journal of literacy research, 31(3), 355-374.
- Street, B. (2003). What’s ‘new’ in new literacy studies? Critical approaches to literacy in theory and practice: Current issues in comparative education, 5(2).
- Guba, E. G., & Lincoln, Y. S. (1994): Competing paradigms in qualitative research: Handbook of qualitative research, 2(163-194), 105.
- Polit, D. F., & Beck, C. T. (2010): Essentials of nursing research: Appraising evidence for nursing practice: Lippincott Williams & Wilkins.
- Al-Osail, A. F., Ibrahim, A. M., Al-Dosari, M. N., Ibrahim, A. R., Abdellatif, M. S., Tashtoush, M. A., ... & Nemt-Allah, M. A. (2026). Statistical Analysis of the Academic Procrastination Scale Demonstrates its Psychometric Properties Through a Rigorous Four-Factor Validation Study.
- Sato, K., Mutoh, A. N., & Kleinsasser, R. C. (2022): Longitudinal research on EFL teacher professional development in (Japanese) contexts: Collaborative action research projects: Language Teaching Research, 26(3), 477-503.
- Watkins, M. W. (2018). Exploratory factor analysis: A guide to best practice: Journal of black psychology, 44(3), 219-246.
- Messick, S. (1996): Validity and washback in language testing: Language testing, 13(3), 241-256.
- Smith, M. K., Jones, F. H., Gilbert, S. L., & Wieman, C. E. (2013): The Classroom Observation Protocol for Undergraduate STEM (COPUS): A new instrument to characterize university STEM classroom practices. CBE—Life Sciences Education, 12(4), 618-627.

- Wang, X., & Wang, S. (2024): Exploring Chinese EFL learners' engagement with large language models: A self-determination theory perspective: *Learning and Motivation*, 87, 102014.
- Taherdoost, H. (2016): Validity and reliability of the research instrument; how to test the validation of a questionnaire/survey in a research: *International journal of academic research in management (IJARM)*, 5.
- Nijakowska, J., Tsagari, D., & Spanoudis, G. (2020): Cross-country comparison of EFL teacher preparedness to include dyslexic learners: Validation of a questionnaire. *Studies in Second Language Learning and Teaching*, 10(4), 779-805.
- Razavipour, K., & Raji, B. (2022): Reliability of measuring constructs in applied linguistics research: A comparative study of domestic and international graduate theses: *Language Testing in Asia*, 12(1), 16.
- Santos, J. S., Andrade, W. L., Brunet, J., & Melo, M. R. A. (2020, October). A systematic literature review of methodology of learning evaluation based on item response theory in the context of programming teaching: In *2020 IEEE Frontiers in Education Conference (FIE)* (pp. 1-9) IEEE.
- Banitalebi, Z., Estaji, M., & Brown, G. T. (2025): Measuring teacher assessment literacy in digital environments: Development and validation of a scenario-based instrument. *Educational Technology & Society*, 28(2), 169-215.
- Kato, M. (2022): Examining the Dependability and Practicality of Analytic Rubric of Summary Writing Using Multivariate Generalizability Theory: Focusing on Japanese University Students with Lower-Intermediate Proficiency in English. *English Language Teaching*, 15(9), 82-94.
- Kormos, J., & Indrarathne, B. (2025): Specific learning differences in learning, teaching, and assessing additional languages. *Language Teaching*, 1-25.
- Razavipour, K., & Raji, B. (2022): Reliability of measuring constructs in applied linguistics research: A comparative study of domestic and international graduate theses: *Language Testing in Asia*, 12(1), 16.
- López-Nuñez, J. A., Alonso-García, S; Berral-Ortiz, B., and Victoria-Maldonado, J. J (2024): A systematic review of digital competence evaluation in higher education: *Education Sciences*, 14(11), 1181.
- Bobkina, J., & Stefanova, S. (2016). Literature and critical literacy pedagogy in the EFL classroom: Towards a model of teaching critical thinking skills: *Studies in second language learning and teaching*, 6(4), 677-696.
- Novianti, N., Thomas, A., & To, V. (2020): Addressing challenges in the practice of critical literacy in EFL classrooms: A new framework.
- Papadopoulos, I., & Griva, E. (2017): Promoting critical literacy in the EFL context: Implementing a project to young learners: *European Journal of Language and Literature*, 7(1), 107.
- Silvhiany, S., Huzaifah, S., & Ismet, I. (2021): Critical digital literacy: EFL students'

- ability to evaluate online sources: Indonesian Journal of EFL and Linguistics, 6(1), 249.
- Gee, J. P. (2004): Learning language as a matter of learning social languages within discourses. *Language learning and teacher education: A sociocultural approach*, 18, 13-31.
- Janks, H. (2013): Critical literacy in teaching and research¹. *Education inquiry*, 4(2)-225-242.
- Morrell, E. (2015). *Critical literacy and urban youth: Pedagogies of access, dissent, and liberation*: Routledge.
- Creswell, J. W., & Plano Clark, V. L. (2018): *Designing and conducting mixed methods research* (3rd ed.). Thousand Oaks, CA: Sage.
- Dörnyei, Z. (2007): *Research methods in applied linguistics*. Oxford: Oxford University Press.
- Johnson, R. B., & Onwuegbuzie, A. J. (2004): *Mixed methods research: A research paradigm whose time has come*. *Educational Researcher*, 33(7), 14–26.
- Comber, B. (2015): *Critical literacy and social justice*: Routledge.
- Vasquez, V. M. (2017): *Critical literacy across the K–6 curriculum*: Routledge.



Appendix 1

Theoretical Framework of Critical Literacy Instruction by Allen Luke and Peter Freebody