

EFL Teachers Challenges in Designing Assessment Material for Students' Listening Skills

Nurhayati Nurhayati¹, Puji Wahyu Setyawaty², Sahril Nur³

¹ Universitas Klabat, Manado, Indonesia (<https://orcid.org/0000-0002-8572-5267>)

² Universitas Negeri Makassar, Makassar, Indonesia

³ Universitas Negeri Makassar, Makassar, Indonesia

Corresponding Email: sahrilnurfbsunm@unm.ac.id

To cite this article:

Nurhayati, N., Setyawaty, P. W., & Nur, S. (2024). EFL Teachers Challenges in Designing Assessment Material for Students' Listening Skills. *ENGLISH FRANCA : Academic Journal of English Language and Education*, 8(2), 409–422. <https://doi.org/10.29240/ef.v8i2.12053>

Abstract. This study investigates the challenges EFL teachers face when designing listening assessment materials, with a focus on understanding these obstacles and identifying strategies to address them. The research aims to explore the alignment of assessment practices with student proficiency levels, resource accessibility, and the impact of time constraints on assessment quality. The study involves eight active English teachers from diverse educational institutions as participants. Data were collected using an online questionnaire designed to gather insights into teachers' perspectives and strategies, guided by the CIPP (Context, Input, Process, Product) evaluation model. This structured framework facilitated a comprehensive analysis of contextual needs, available resources, procedural approaches, and assessment outcomes. The data were analyzed descriptively using Google Forms, with key findings highlighting significant challenges in sourcing appropriate audio materials, ensuring inclusivity and validity, and managing time for creating bespoke assessments. Results revealed that while 50% of teachers successfully aligned assessments with student abilities, 25% faced difficulties in selecting suitable materials due to resource limitations. Teachers frequently relied on self-created or freely available online resources and employed pre-made tests due to time constraints. Despite these efforts, gaps in assessment literacy and resource access persisted. The study underscores the need for professional development programs to enhance language assessment literacy, improved access to diverse, high-quality audio materials, and systemic support to address these challenges. By addressing these issues, this research contributes to more effective listening assessment practices, fostering equitable and comprehensive EFL education.

Keywords: *EFL Teachers, Challenges, Assessment Materials, Listening Skills*

Introduction

The role of listening assessments in English as a Foreign Language (EFL) instruction has gained increasing attention in recent years due to their critical importance in both language acquisition and communication. Listening is a foundational skill in language learning, as it directly influences how students engage with spoken language in real-world settings. However, designing effective listening assessments that accurately measure students' listening comprehension is a complex task. Effective listening assessments require students to process various linguistic aspects, including intonation, contextual cues, stress patterns, and cultural nuances, which are often difficult to capture within standardized test formats (Bizunesh, Alemu, & Ashenafi, 2021). As such, EFL teachers face numerous challenges in creating assessments that accurately reflect the multifaceted nature of listening

Article info:

<http://journal.iaincurup.ac.id/index.php/english>

Received 24 April 2024; Received in revised form 10 July 2024; Accepted 24 August 2024, Available Online 24 November 2024

Published by Institut Agama Islam Negeri (IAIN) Curup on behalf of ENGLISH FRANCA: Academic Journal of English Language and Education. This is an open-access article under the CC BY-SA license Copyright (c) 2024 Author

comprehension. This article aims to explore these challenges in depth, providing insights into the difficulties faced by EFL teachers in designing listening assessments and offering potential solutions to enhance assessment practices in diverse educational contexts.

Listening is not merely about decoding spoken words but involves complex cognitive processes such as identifying and interpreting the meaning of words in context, recognizing various accents, understanding cultural references, and responding to emotional cues. Bizunesh et al. (2021) note that EFL teachers in secondary schools face significant difficulties in assessing students' listening skills due to the inherent complexity of this skill. Listening comprehension requires a sophisticated understanding of both language and the situational context in which the language is used. For example, comprehension of a conversation may depend not only on the students' ability to understand the vocabulary but also on recognizing the tone, rhythm, and the speaker's intent, which are often not easily conveyed through written assessments (Chie, 2023). Moreover, listening assessments need to account for the wide range of accents, speech rates, and linguistic variations encountered in real-world communication. The limitations of traditional assessment formats in capturing the full range of listening abilities are significant. Standardized tests, which often rely on multiple-choice questions or short-answer formats, may fail to accurately measure students' listening comprehension, as they do not account for the complexities of real-world listening tasks (Jeanne & Burns, 1988). Furthermore, teachers must balance the need for reliability and validity in assessments with the need to avoid bias, especially given the linguistic and cultural diversity present in many EFL classrooms. For instance, accents that students are less familiar with may unintentionally disadvantage them in listening tests, making the assessment process even more challenging (Podder, 2011).

The challenges associated with listening assessment extend beyond technical and pedagogical concerns to include systemic and contextual issues that further exacerbate these difficulties. Resource limitations, including the lack of high-quality audio materials, pose a significant barrier. Teachers often find it difficult to locate or create materials that reflect authentic speech, a critical element for fostering listening comprehension. Syahrin and Salih (2020) emphasize the importance of technological tools and authentic resources in designing reliable listening assessments. However, such tools and materials are often scarce, particularly in under-resourced educational settings, leaving teachers reliant on simplified or contrived listening tasks that fail to simulate real-world communication. Moreover, the COVID-19 pandemic highlighted the fragility of existing assessment practices and the lack of preparedness among educators to adapt to remote learning environments. Teachers faced unprecedented challenges in administering listening assessments through digital platforms, compounded by inadequate training and access to appropriate technological tools (Al-Jarf, 2021; Shofatunnisa et al., 2021). This disruption not only underscored the need for innovative approaches to assessment design but also revealed a critical gap in teacher training and support for online assessment modalities. As Ghanbari and Nowroozi (2021) argue, many teachers lack the expertise to design effective online assessments, particularly for complex skills such as listening comprehension. This gap in technological and pedagogical readiness remains a significant obstacle to the advancement of equitable and effective listening assessment practices.

Another pressing issue is the gap in language assessment literacy (LAL) among EFL teachers. Language assessment literacy refers to the knowledge and skills required to design, implement, and interpret assessments effectively (Fulcher, 2012). Without adequate training, teachers may struggle to develop assessments that are both valid and reliable. Jeanne and Burns (1988) argue that inconsistencies in how listening skills are defined and measured can lead to assessments that do not accurately capture students' abilities. These inconsistencies are often compounded by the pressure to prepare students for high-stakes testing, which frequently prioritizes standardized formats over comprehensive evaluations of listening comprehension. Consequently, many assessments fail to account for the nuanced and

context-dependent nature of listening skills, limiting their effectiveness in fostering genuine language proficiency (Rylander et al., 2023).

Cultural and linguistic diversity in EFL classrooms further complicates the task of designing listening assessments. Students bring a wide range of linguistic backgrounds and cultural experiences to the classroom, which can influence their ability to understand different accents, speech patterns, and cultural references. Podder (2011) highlights the potential for biases in listening assessments that favor certain linguistic or cultural groups over others. For example, students who are familiar with a particular accent may perform better than those who are not, even if their overall listening comprehension skills are comparable. This creates an equity issue that must be addressed to ensure that listening assessments are fair and inclusive. Vygotsky's sociocultural theory (1978) underscores the importance of designing assessments that reflect the diverse social and cultural contexts of learners. By incorporating culturally relevant materials and diverse linguistic inputs, teachers can create assessments that are more representative of real-world communication and more equitable for all students. Despite the growing body of literature on these challenges, significant research gaps remain. While previous studies have identified individual obstacles such as resource limitations or gaps in assessment literacy, few have examined how these factors interact to shape assessment practices. Additionally, there is a lack of research on the systemic and institutional barriers that hinder the development of effective listening assessments, including the role of curriculum design, policy support, and teacher collaboration. This study addresses these gaps by adopting a holistic approach that considers the interconnected nature of these challenges and their implications for EFL education.

The novelty of this research lies in its application of the CIPP (Context, Input, Process, Product) evaluation model as a framework for analyzing the complexities of listening assessment. Unlike descriptive studies that focus on isolated aspects of assessment design, this study uses the CIPP model to provide a structured and comprehensive evaluation of the contextual factors, resources, methods, and outcomes associated with listening assessments. By doing so, it offers actionable insights that can inform the development of more effective and inclusive assessment practices. The CIPP model's emphasis on aligning assessment practices with educational goals and available resources makes it particularly well-suited to addressing the systemic and contextual challenges identified in this study. Furthermore, this research contributes to the field by emphasizing the importance of equity and inclusivity in listening assessment. By highlighting the need for culturally responsive assessment practices, it provides a framework for addressing the biases and inequities that often characterize traditional assessment formats. This focus on inclusivity aligns with broader trends in EFL education that prioritize the development of global communication skills and intercultural competence. The study also underscores the potential of technology to transform listening assessment practices, particularly through the use of interactive and adaptive tools that can accommodate diverse learner needs and contexts. By integrating technological innovation with pedagogical best practices, this research sets the stage for a new era of listening assessment that is both effective and equitable. In summary, this study addresses critical gaps in the literature by providing a comprehensive analysis of the challenges associated with listening assessment in EFL contexts. Its use of the CIPP evaluation model, emphasis on equity and inclusivity, and focus on technological innovation represent significant contributions to the field. By offering practical recommendations for educators and policymakers, this research aims to enhance the design and implementation of listening assessments, ultimately supporting the development of more effective and equitable EFL education systems.

Theoretical Framework

To frame the challenges faced by EFL teachers in designing assessment materials for listening skills, it is essential to root the discussion in robust theoretical foundations. This section synthesizes relevant theories, models, and frameworks from the literature, emphasizing their alignment with the key issues identified in the study.

Importance of Listening in Language Learning

Listening is a fundamental skill in language acquisition, forming the foundation for communication competence. According to Krashen's (1985) Input Hypothesis, comprehensible input is vital for language acquisition, with listening serving as the primary means of receiving this input. Effective listening assessments, therefore, must measure students' ability to process and interpret spoken language in context. This encompasses intonation, cultural nuances, and contextual cues (Bizunesh et al., 2021). Moreover, Vandergrift and Goh (2012) argue that listening is not a passive activity but an active process involving cognitive, metacognitive, and socio-affective strategies. Assessments must account for these dimensions to accurately capture students' listening comprehension levels. Failure to do so could lead to an incomplete evaluation of their abilities and hinder effective instruction.

Resource Limitations and Technological Challenges

The availability and quality of resources significantly impact listening assessments. Nurhandayani et al. (2020) highlight the negative effects of poor-quality audio equipment and limited resources on listening comprehension outcomes. Similarly, Syahrin and Salih (2020) emphasize the role of technology in facilitating effective listening assessments, particularly in the wake of the COVID-19 pandemic. The theoretical lens of the Technological Pedagogical Content Knowledge (TPACK) framework (Mishra & Koehler, 2006) can be applied to understand how teachers integrate technology into listening assessments. TPACK highlights the interplay between technological, pedagogical, and content knowledge, suggesting that effective use of technology requires a balance of these elements. In contexts where resources are scarce, teachers may struggle to achieve this balance, impacting their ability to design meaningful assessments (Syahrin & Salih, 2020).

Curriculum and Contextual Adaptation

Curriculum changes and contextual factors often require teachers to adapt their practices, as the dynamic nature of curricula necessitates flexibility and responsiveness in assessment design. The CIPP (Context, Input, Process, Product) evaluation model (Stufflebeam, 1971) provides a comprehensive framework for assessing and adapting educational programs. This model offers insights into how teachers navigate challenges in designing listening assessments by aligning their efforts with curriculum goals, utilizing available resources, refining practices for reliability, and evaluating assessment effectiveness.

Language Assessment Literacy (LAL)

Language Assessment Literacy (LAL) refers to teachers' knowledge, skills, and principles in designing, implementing, and interpreting assessments (Fulcher, 2012). Ghanbari and Nowroozi (2021) note that gaps in LAL are a significant barrier for EFL teachers in creating effective listening assessments. This includes understanding principles such as reliability, validity, and fairness, as well as the ability to select or develop appropriate assessment instruments. Jeanne et al. (1988) argue that inconsistencies in how teachers define and measure listening skills lead to discrepancies in assessment outcomes. Geranpayeh and Taylor (2013) emphasize the importance of using validated tools and frameworks, such as Rasch analysis, to ensure the reliability and fairness of listening assessments.

Cultural and Linguistic Diversity

Classroom diversity introduces additional challenges in designing listening assessments. Teachers must consider cultural and linguistic differences to avoid biases that disadvantage certain student groups. Podder (2011) highlights the importance of inclusive assessment practices that address the needs of diverse learners. Sociocultural theory (Vygotsky, 1978) provides a useful perspective for understanding these challenges. It emphasizes the role of social interaction and cultural context in learning, suggesting that assessments should reflect the diverse backgrounds of students. For example, listening materials should include culturally relevant content to ensure that all students can engage meaningfully with the tasks.

Time Constraints and Practical Challenges

Time limitations often force teachers to rely on pre-made assessments, which may not align with their students' needs (Wallerstedt, 2011). The management framework (Macan, 1994) can be applied to explore how teachers prioritize tasks and manage their time effectively. Teachers who lack adequate time to design customized assessments may face challenges in ensuring the validity and reliability of their tools.

Validation of Listening Assessments

Validating listening assessments is critical for ensuring their effectiveness. Rasch analysis, as described by Rylander et al. (2023), provides a robust method for measuring item difficulty and ensuring consistency across test versions. This approach aligns with the principles of construct validity, which require that assessments accurately measure the intended skills. Engaging and interactive assessment methods, such as the use of mindful listening rubrics, can enhance the validity of listening assessments by capturing a broader range of listening abilities (Rylander et al., 2023). Jeanne et al. (1988) also emphasize the importance of using diverse assessment methods to provide a comprehensive evaluation of students' listening skills.

Addressing Listening Comprehension Challenges

Listening comprehension involves multiple processes, including decoding, linguistic knowledge, and higher-order cognitive skills. Chie (2023) highlights the role of cognitive assessments in identifying and addressing listening difficulties. Teachers must design assessments that capture these dimensions to provide a complete picture of students' listening abilities. The Interactive-Compensatory Model (Stanovich, 1980) suggests that listeners use compensatory strategies to overcome difficulties, such as relying on contextual cues when linguistic knowledge is insufficient. Assessments should account for these strategies to provide a more accurate measure of listening comprehension.

Implications for Teacher Training and Professional Development

Addressing the challenges identified in this study requires targeted professional development programs designed to enhance teachers' skills in designing and implementing listening assessments. Training programs should focus on developing Language Assessment Literacy (LAL), ensuring teachers understand essential principles of assessment design, such as reliability and validity, as emphasized by Ghanbari and Nowroozi (2021). Additionally, integrating technology effectively using the Technological Pedagogical Content Knowledge (TPACK) framework (Mishra & Koehler, 2006) can help teachers balance technological, pedagogical, and content knowledge. Furthermore, adapting assessments to diverse classroom contexts can be guided by the Context, Input, Process, Product (CIPP) evaluation model (Stufflebeam, 1971), which provides a structured approach to aligning assessments with educational goals and resources. These tailored programs aim to equip teachers with practical strategies and tools to address the complex challenges of designing effective listening assessments in varied educational settings. The theoretical framework for this study

integrates insights from language acquisition, assessment literacy, technological integration, and cultural diversity to provide a comprehensive understanding of the challenges faced by EFL teachers in designing listening assessments. By grounding the study in these theories, the findings can contribute to the development of practical solutions and inform policy and practice in EFL teaching. Future research should explore how these theoretical perspectives can be applied in different contexts to further enhance the effectiveness of listening assessments.

Material and Method

This quantitative descriptive study employs an ex post facto survey design to investigate the challenges faced by EFL teachers in designing listening assessment materials. Ex post facto research, meaning "after the fact," is particularly suitable for examining existing conditions without manipulating variables, allowing for the exploration of relationships in naturally occurring settings.

Selection of Subjects

Participants were purposefully selected to ensure representation of active English teachers with diverse experiences in designing listening assessments. The eight teachers chosen for this study were identified based on their roles in EFL instruction across various educational contexts. Criteria for selection included their current engagement in teaching English as a foreign language, prior experience in designing assessments, and willingness to participate in the study. This purposeful sampling ensures that the findings are grounded in the lived experiences of practitioners directly involved in listening assessment design.

Instrumentation

Data collection was facilitated through an online questionnaire developed to capture detailed insights into the challenges and strategies associated with listening assessments. The questionnaire was designed following the CIPP (Context, Input, Process, Product) evaluation model, providing a structured framework to examine various aspects of the assessment process. The instrument included 20 indicators distributed across the four CIPP components:

1. Context: Indicators focused on understanding the environment in which teachers operate, including institutional policies, curriculum requirements, and student demographics.
2. Input: This section evaluated the availability of resources, access to technology, and professional development opportunities.
3. Process: Indicators explored the methods and procedures used by teachers in designing and implementing listening assessments, such as time allocation and material selection.
4. Product: Questions assessed the outcomes and effectiveness of the designed materials, emphasizing alignment with student needs and overall validity.

Validity and Reliability

To ensure the validity and reliability of the questionnaire, a two-step process was employed. First, content validity was established through expert review. Two specialists in EFL education and assessment reviewed the instrument to confirm that the indicators were comprehensive and aligned with the study's objectives. Second, a pilot study was conducted with three EFL teachers who were not part of the main study. Their feedback was used to refine the language, clarity, and relevance of the questions. Internal consistency reliability was calculated using Cronbach's alpha, achieving a coefficient of 0.85, indicating high reliability.

Data Analysis

The data collected through the questionnaire were analyzed descriptively using Google Forms' analytical tools, which provided immediate summaries of responses, including

frequency distributions and percentages. For a more nuanced understanding, the data were exported to SPSS software for advanced statistical analysis. This allowed for the identification of patterns, correlations, and significant trends within the data. The analysis focused on the responses to the CIPP model indicators, providing a comprehensive view of the challenges and strategies employed by teachers. By combining descriptive and inferential statistical techniques, the analysis not only summarized the findings but also offered deeper insights into the systemic and procedural issues faced by teachers. This methodological rigor enhances the reliability and validity of the study, ensuring that the results are both actionable and reflective of the complexities of listening assessment in EFL contexts.

This methodological approach provides a robust framework for evaluating the challenges encountered by EFL teachers, offering valuable insights to inform targeted interventions and support mechanisms aimed at improving listening assessment practices.

Results and Discussion

Designing effective listening assessment materials is a multifaceted challenge faced by English as a Foreign Language (EFL) teachers. This study explores how teachers navigate these challenges across four key stages of material design: context, input, process, and product. Each stage reveals distinct obstacles and corresponding strategies that teachers employ to enhance the effectiveness of listening assessments.

The Context Stage

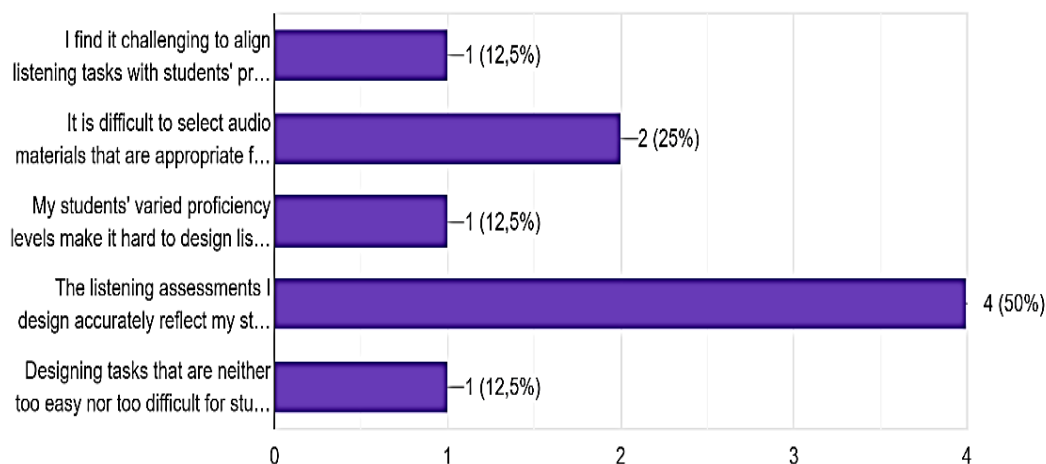


Figure 1

The perceived difficulty of aligning materials with proficiency

The context stage is concerned with the various needs such as the goals or program objectives that must be align the material with the listening proficiency level. The result shows 50% teachers said “the listening assessments that they designed accurately reflect their students’ abilities. But, 25% teachers faced difficult to select audio materials that are appropriate for their students’ listening abilities. Supporting the difficulty, Chie (2023) said integrating comprehensive evaluation approach with cognitive assessments in auditory tests is contributing listening difficulties and improve the identification of affected individuals. There are some teachers’ argument to solve the problem.

“There are a lot of listening materials but most of them are not free. So, I'll have to prepare more to start a listening practice because I have to make it by myself so it will be cheaper.”

"I have prepared many kinds of level listening sections so I am capable to teach them based on their level."

"Because before I try to use listening method, I observe some vocabularies of my students, and I pick easy topics that related with their ability by using listening method."

To sum up, the teachers solved this problem by prepare the material early before the listening class starts.

The Input Stage

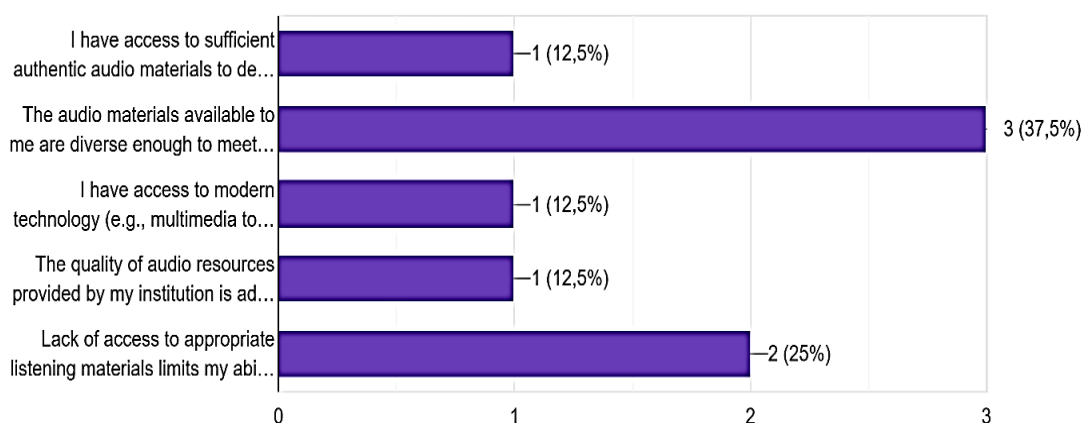


Figure 2

The challenges relating to availability of resources in designing listening assessment material

The input phase focuses on the availability of resources in designing listening assessment material. The result of this research showed 35% teachers said the audio materials available for them are diversity enough to their student varying needs. But, 25% of teachers meet challenge about lack of access to appropriate listening materials limits teachers' ability to design effective assessments. Nurhandayani, et.al (2020) said poor quality audio equipment negatively affected students' listening experiences. Podder (2011) mentioned teachers face challenges in creating listening assessments due to insufficient resources. There some statements from teachers how they solved the problem.

"Using modern technology to create listening assessment is simple, available and practice."

"I have collected a bunch of listening audios so I just need to play one of those audios as student's level."

To solve the problem, I used to find another resource from internet which is free and easy to access and the quality of the audio is better than the one that I got from the school.

Based on the statements above the teachers solve the problem by searching free resources in google which fix their students ability.

The Process Stage

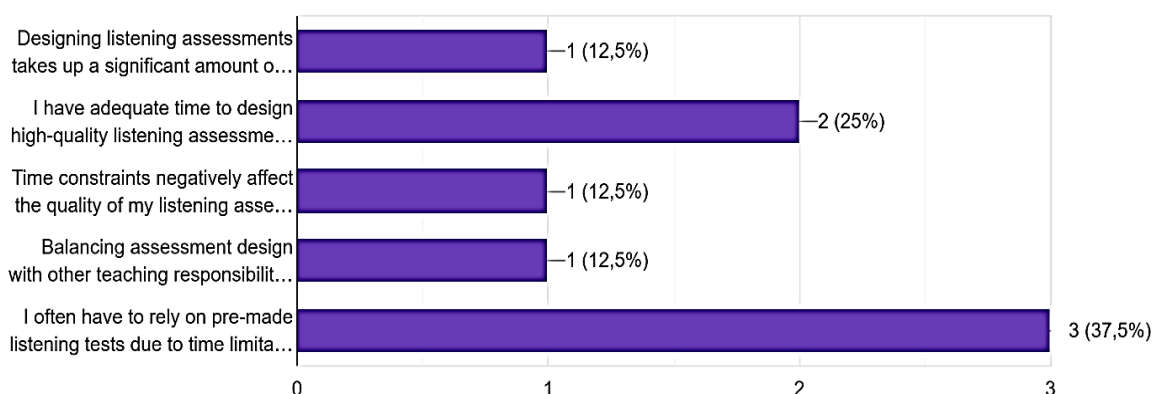


Figure 3

Designing listening relates to time constraints assessment

This stage presents teachers' attempting to implement their preparation material. There are 37,5% teachers, they often have to rely on pre-made listening tests due to time limitations. Wallerstedt (2011) said teachers face challenges in creating listening assessments due to the temporality and managing dimensions of variation. There are some arguments how teachers overcome the problem.

I always use stopwatch so I know when the time is too short or already run out.

In designing listening assessment, we can choose to select or develop, if we want to make it more suitable to the condition where we teach or to make it appropriate to our students, we may develop the listening assessment, but it's more consuming the time and cost highly.

Sometimes, due to time limitations, I often have to rely on pre-made listening test, so I try to pick easy topics to overcome this.

Summarizing the teachers' ways to solve the problem, they tried to control time for setting how long time they need to design the assessment. Then, they use test that creates before.

The Product

The product stage where the design of listening assessment material is assessed. Evaluating a listening assessment requires a comprehensive evaluation should focus on definition of listening, the validity of the assessment instruments and the incorporation of diverse assessment methods to capture listening abilities effectively. Different teachers may have varying definitions of listening capacity, leading to inconsistent measurements for instance, the test misalignment with the intended area of comprehension and unrelated skills (Jeanne, 1988). A clear framework for test validation is essential for supporting claims about test score interpretations (Geranpayeh & Taylor, 2013). Using Rasch analysis can help validate listening assessments by measuring item difficulty and ensuring reliability across different test versions. Then, engaging activities and mindful listening rubrics can enhance the assessment process, making it more interactive and reflective of true listening abilities (Rylander et al., 2023).

Discussion

The findings of this study shed light on the intricate challenges faced by EFL teachers in designing listening assessment materials and underscore the necessity for comprehensive strategies to address these issues. By integrating theoretical perspectives and empirical insights, this discussion explores the nuances of these challenges and identifies actionable solutions to improve listening assessment practices. The analysis draws on key frameworks and theories, emphasizing their relevance in addressing obstacles such as resource scarcity, time constraints, assessment validity, and the lack of inclusivity in test design.

Tailoring Materials to Student Proficiency Levels

One of the most pressing challenges identified in the study is the difficulty teachers face in sourcing and adapting materials to align with their students' diverse proficiency levels. EFL classrooms often comprise learners with varying degrees of listening competence, making it essential to provide tailored assessments. This issue aligns with Vygotsky's Sociocultural Theory (1978), which emphasizes the role of social and cultural contexts in shaping learning experiences. According to this theory, the "zone of proximal development" (ZPD) is critical for effective learning. Listening materials that are either too simplistic or overly complex fall outside the learners' ZPD, thus failing to foster optimal comprehension and skill development. To address this, educators must incorporate authentic and culturally relevant listening materials that reflect students' linguistic and cultural backgrounds. The integration of such content not only enhances engagement but also ensures that assessments resonate with learners' lived experiences. However, achieving this requires institutional support in curating and distributing a diverse range of listening resources. Teacher collaboration, facilitated by shared digital repositories, can also play a pivotal role in bridging resource gaps. By pooling materials that are organized by proficiency levels, educators can efficiently access and adapt content to meet the needs of their students.

The Role of Technology in Mitigating Resource Scarcity

Another persistent challenge is the scarcity of high-quality audio materials. The study revealed that 25% of teachers identified this as a significant obstacle, compounded by the limited availability of appropriate technological tools. This finding is consistent with the TPACK framework (Mishra & Koehler, 2006), which underscores the importance of integrating technological knowledge with pedagogical and content expertise. The lack of access to high-quality audio resources forces teachers to rely on freely available or self-created materials, which often lack the sophistication needed to assess nuanced listening skills effectively. To overcome this challenge, institutions must prioritize the provision of curated digital libraries containing authentic audio recordings and interactive multimedia tools. Such resources should encompass diverse accents, speech patterns, and real-world scenarios to prepare students for the complexities of authentic listening. Additionally, training programs focused on TPACK principles can equip teachers with the skills to leverage technology effectively in their assessment practices. For instance, the use of digital platforms that facilitate the creation of customizable listening tasks can enable educators to address specific proficiency levels and learning objectives. Frameworks like the Technological Acceptance Model (TAM) further highlight the importance of perceived ease of use and usefulness in encouraging the adoption of new technologies. Institutions must therefore ensure that technological tools are user-friendly and aligned with teachers' pedagogical goals. Professional development initiatives should also emphasize the pedagogical integration of technology, moving beyond mere technical training to address its role in enhancing listening comprehension and assessment validity.

Addressing Time Constraints

Time constraints are another critical issue, significantly affecting teachers' ability to design bespoke listening assessments. The demands of lesson planning, classroom management, and administrative responsibilities often leave little room for the meticulous

design of tailored assessment tasks. This finding is consistent with the Time-Management Framework (Macan, 1994), which identifies the importance of structured time allocation in managing competing demands effectively. Teachers often resort to pre-made tests due to limited time, but such materials may not always align with their students' specific needs. To address this, educators can adopt strategies such as pre-selecting adaptable resources and utilizing automated assessment tools. These measures, while effective in saving time, must be balanced with the need to maintain the depth and accuracy of evaluations. Institutions can support teachers by streamlining administrative processes and providing access to ready-made, high-quality assessment templates that can be customized as needed. Moreover, fostering collaboration among educators can alleviate individual time burdens. By working together to develop and share listening assessment materials, teachers can pool their expertise and create a repository of resources that saves time while ensuring quality. Time-management training, informed by Macan's framework, can further empower teachers to prioritize tasks effectively and allocate sufficient time for assessment design.

Ensuring Validity and Reliability in Listening Assessments

The validity and reliability of listening assessments are central to their effectiveness in evaluating students' comprehension skills. Variability in how listening is defined and evaluated often leads to inconsistent outcomes, a concern highlighted in the study. Rasch analysis, as proposed by Rylander et al. (2023), offers a robust methodological approach to enhancing the reliability and consistency of assessment tools. By applying Rasch modeling, educators can ensure that listening tasks accurately measure the intended skills and provide meaningful data on student performance. Bachman and Palmer's model of language assessment (1996) further underscores the importance of practicality, reliability, and authenticity in test design. Listening tasks must reflect real-world communication scenarios to provide an accurate measure of students' abilities. This involves incorporating diverse accents, speech rates, and contextual cues into assessments. The inclusion of such elements not only enhances the authenticity of listening tasks but also prepares students for the linguistic diversity they are likely to encounter outside the classroom. Additionally, addressing cultural and linguistic biases in listening materials is essential for ensuring equity in assessments. Sociocultural Theory (Vygotsky, 1978) supports the inclusion of culturally relevant content, which helps create inclusive assessments that resonate with learners from diverse backgrounds. Institutions must prioritize the development of standardized guidelines that emphasize inclusivity and fairness in listening assessments. Such guidelines can provide teachers with a clear framework for designing tasks that are both valid and reliable.

Enhancing Language Assessment Literacy (LAL)

The lack of teacher training in language assessment literacy (LAL) is a significant barrier to designing effective listening assessments. As Jeanne et al. (1988) argue, variability in assessment practices often stems from gaps in educators' understanding of key principles such as validity, reliability, and fairness. Professional development programs focused on LAL are therefore critical for addressing these gaps. These programs should equip teachers with the skills to navigate the complexities of listening comprehension and design assessments that accurately reflect students' abilities. This includes training in the use of statistical tools for evaluating test validity and reliability, as well as practical strategies for developing fair and inclusive listening tasks. Additionally, institutions must provide ongoing support for teachers to apply these skills in their practice, ensuring that LAL principles are integrated into everyday assessment design.

Leveraging Technology to Create Dynamic Assessments

Technology offers immense potential for creating dynamic and engaging listening assessments that reflect authentic listening experiences. For example, multimedia resources and gamified tasks can make assessments more interactive and aligned with real-world

communication scenarios. By training teachers in TPACK principles, institutions can empower them to integrate technology seamlessly into their practices.

Emerging technologies, such as AI-driven adaptive assessments, also hold promise for addressing challenges related to validity and resource availability. These tools can analyze students' performance in real-time and adjust the difficulty level of tasks accordingly, providing a personalized assessment experience. Further research into the application of such technologies in EFL contexts could reveal innovative solutions for enhancing listening assessment practices.

Implications for Policy and Practice

The challenges identified in this study have significant implications for policy and practice. Stakeholders must prioritize funding for teacher training and the development of high-quality materials to bridge the gaps identified. Allocating resources to these areas is essential for supporting educators in their efforts to create effective listening assessments. Institutions should also establish networks for teachers to share best practices and resources, fostering a culture of collaboration that reduces the burden on individual educators. Policies that emphasize inclusivity and fairness in listening assessments can provide teachers with a clear framework for designing tasks that meet the diverse needs of their students.

Future Research Directions

The findings of this study highlight several avenues for further research. Long-term studies exploring the impact of professional development programs on assessment practices and student outcomes could provide valuable insights. Additionally, investigating the use of emerging technologies, such as AI-driven adaptive assessments, could reveal innovative ways to address challenges related to validity and resource availability. Comparative research on how cultural and linguistic contexts influence assessment practices would also be instrumental in creating more inclusive and adaptable models. Such studies could inform the development of standardized guidelines that emphasize equity and cultural relevance in listening assessments. Designing effective listening assessments in EFL contexts is a multifaceted challenge that demands a comprehensive approach. The issues outlined in this study—resource limitations, time constraints, gaps in assessment literacy, and cultural inclusivity—highlight the need for targeted interventions and systemic support. By investing in professional development, resource availability, and technological integration, institutions can enable teachers to create assessments that capture the complexities of listening comprehension. Future research and policy initiatives should focus on sustainable solutions that address these challenges, ensuring that listening assessments contribute significantly to students' language learning progress.

Conclusion

EFL teachers encounter numerous challenges in designing effective listening assessment materials, emphasizing the need for systemic support to address these complexities. One major difficulty lies in selecting audio materials that match students' diverse listening abilities. Teachers must balance the content's complexity to ensure it is neither overly simplistic nor excessively challenging. This task is further complicated by the varied proficiency levels within classrooms, which demand materials that cater to a broad spectrum of student needs. Compounding this issue is the limited availability of diverse and high-quality audio resources. While materials may exist, many are unsuitable for the specific linguistic and cultural contexts of students, forcing teachers to rely on subpar or self-created resources. This reliance affects the overall quality and relevance of listening assessments.

Time constraints present another significant obstacle for teachers. The demands of teaching schedules often leave little room for the creation of customized listening assessments. Consequently, educators frequently resort to pre-made tests that may not align with their

students' specific learning objectives or cultural backgrounds. This mismatch between test content and student needs diminishes the effectiveness of assessments, undermining their role in fostering language development. To address these challenges, teachers have implemented various strategies. One such approach is dedicating additional preparation time to develop tailored listening materials. By creating their own resources, teachers can ensure the content aligns with their students' proficiency levels and cultural contexts, enhancing engagement and learning outcomes. This strategy is also cost-effective, reducing dependence on commercially available materials that may not be ideal for classroom use. Additionally, teachers conduct vocabulary assessments to gauge students' lexical knowledge, which helps in selecting or crafting listening materials of appropriate difficulty. Choosing relatable and familiar topics further enhances students' engagement and comprehension during listening exercises.

Acknowledgement

The researcher would like to thank all parties who have contributed to this research, especially the participating teachers for contributing to this study.

References

- Al-Jarf, R. (2021). Challenges in Conducting Online Speaking Activities during the Pandemic. <https://www.sciedu.ca/journal>.
- Bizunesh, W., Alemu, A., & Ashenafi, B. (2021). EFL Teachers' Perception and Practice of Assessing Listening Skills, and the Challenges They Face: Secondary Schools in West Jimma in Focus. Jimma University. <https://repository.ju.edu.et/handle/123456789/6215>.
- Bozorgian, H., & Pillay, H. (2013). Enhancing EFL Listening Comprehension Through Integrative Listening Tasks. *The Language Learning Journal*, 41(1), 41-56. doi: [10.1080/09571736.2011.625096](https://doi.org/10.1080/09571736.2011.625096).
- Brown, H.D. (2015). *Teaching by Principles: An Interaction Approach in Language Pedagogy*. (4th Ed.). Prentice Hall Regents.
- Chie, O. (2023). Current Issues in Assessment of Listening Difficulties. *Audiology Japan*. doi: [10.4295/audiology.66.221](https://doi.org/10.4295/audiology.66.221).
- Feyten, C.M. (1991). The Power of Listening Ability: An Overlooked Dimension in Language Acquisition. *The Modern Language Journal*, 75(2), 173-180.
- Field, J. (2019). *Listening in the Language Classroom*. Cambridge Language Teaching Library. Cambridge University Press.
- Geranpayeh, A., & Taylor, L. (2013). *Examining Listening: Research and Practice in Assessing Second Language Listening*. Cambridge University Press.
- Ghanbari, N., & Nowroozi, S. (2021). Exploring Barriers in Online Assessment for EFL Teachers.
- Graham, S., & Santos, D. (2015). Strategies for Second Language Listening: Current Scenarios and Improved Pedagogy. *The Modern Language Journal*, 99(3), 653-669. doi: [10.1111/modl.12247](https://doi.org/10.1111/modl.12247).
- Jeanne, M., Burns, Donald, J., & Richgels. (1988). A Critical Evaluation of Listening Tests. *Intervention In School and Clinic*, 24, 153-162. doi: [10.1177/105345128802400206](https://doi.org/10.1177/105345128802400206).
- Li, J., & Renandya, W. A. (2020). Effective Listening Instruction in EFL Classrooms: Balancing Strategy and Practice. *Asian Journal of Language Teaching*, 28(2), 45-61.
- Lynch, T. (2017). *Teaching Second Language Listening*. Oxford Handbooks for Language Teachers. Oxford University Press.
- Nurhandayani, S., Eliza, N.F., Zuhad, & Ahmad. (2020). Listening Comprehension Problems Among EFL Vocational School Students. 1(1), 1-11. doi: [10.22236/ELLTERJ.V1I1.4907](https://doi.org/10.22236/ELLTERJ.V1I1.4907).
- Nushi, M., & Orouji, F. (2020). Investigating EFL Teachers' Views on Listening Difficulties Among Their Learners: The Case of Iranian Context. Sage Publication.

- Podder, R. (2011). Barriers and Enablers for Teachers Assessing Listening and Speaking Skills at Secondary Level in Bangladesh. doi: [10.26021/9513](https://doi.org/10.26021/9513).
- Rost, M. (2021). Teaching and Assessing Listening in the Second Language Classroom. Pearson Education.
- Rylander, J., Stewart, T., Emi, I., & Sasao, Y. (2023). Test Validation Through Rasch Analysis for an Academic Listening Course. SSRN. doi: [10.2139/ssrn.4642099](https://doi.org/10.2139/ssrn.4642099).
- Shofatunnisa, D., et al. (2021). Using YouTube Videos for EFL Speaking Assessment: A Qualitative Analysis. <https://www.sciedu.ca/journal>.
- Stufflebeam, D. (2003). The CIPP Model of Evaluation. In T. Kellaghan, D. Stufflebeam, & L. Wingate (Eds.), Springer International Handbooks of Education: International Handbook of Educational Evaluation.
- Syahrin, S., & Salih, A. (2020). Investigating Online Learning Styles and Technology in ESL Classes during COVID-19. <https://www.sciedu.ca/journal>
- Vandergrift, L., & Goh, C. C. M. (2018). Teaching and Learning Second Language Listening: Metacognition in Action. Routledge. doi: [10.4324/9781315462849](https://doi.org/10.4324/9781315462849).
- Wallerstedt, C. (2011). Didactic Challenges in the Learning of Music-Listening Skills. 105-126. doi: [10.1007/978-94-007-1617-9_6](https://doi.org/10.1007/978-94-007-1617-9_6).
- Xu, L., & Yi, X. (2020). Advancements in EFL Listening Assessment: A Review of Challenges and Solutions. Journal of Language Education and Research, 7(3), 255-270. doi: [10.21462/jle.v7i3.310](https://doi.org/10.21462/jle.v7i3.310).
- Yamada, H., & Nakamura, K. (2022). Innovative Tools for EFL Listening Assessment in Digital Learning Environments. TESOL Quarterly, 56(4), 892-910. doi: [10.1002/tesq.3090](https://doi.org/10.1002/tesq.3090).
- Zeng, Y., & Zhang, Y. (2021). Mobile-Assisted Listening Assessments in EFL Classrooms: A Meta-Analysis. Educational Technology Research and Development, 69(5), 2473–2496. doi: [10.1007/s11423-021-10057-9](https://doi.org/10.1007/s11423-021-10057-9).
- Zhang, G., Zeller, N., Griffith, R., Metcalf, D., Williams, J., Shea, C., & Misulis, K. (2011). Using the Context, Input, Process, and Product Evaluation Model (CIPP) as a Comprehensive Framework to Guide the Planning, Implementation, and Assessment of Service-Learning Programs. Journal of Higher Education and Outreach Engagement, 15(4), 57–83.