**Library Progress International**
*Vol.44 No. 3, Jul-Dec 2024: P. 16097-16104*

*Print version ISSN 0970 1052*
*Online version ISSN 2320 317X*

*Original Article*

*Available online at www.bpasjournals.com*

# Direct vs. Indirect Listening Assessments: A Comparative Study of Gap-Filling and Cloze Tests on Social Media Using AI

**Prof. Abdullah Al Fraidan**

English Language Department, College of Arts, King Faisal University, Al Ahsa, Saudi Arabia
afridan@kfu.edu.sa

**How to cite this article:** Prof. Abdullah Al Fraidan (2024) Direct vs. Indirect Listening Assessments: A Comparative Study of Gap-Filling and Cloze Tests on Social Media Using AI. *Library Progress International,* 44(3), 16097-16104

**ABSTRACT**

This study investigates the effectiveness of direct and indirect listening assessments by comparing two test formats: gap-filling and cloze tests, centered on the topic of social media. The research involved 12 high school students, equally divided by gender, who were administered AI-generated listening comprehension tests. The gap-filling test, used as a direct listening assessment, required students to fill in blanks while listening to a real-time conversation about social media, engaging them in active auditory processing. The cloze test, serving as an indirect listening measure, involved filling in blanks in a written transcript after listening to a similar conversation. Results indicated that the gap-filling test was more effective in promoting cognitive engagement, real-time processing, and inference-making. Additionally, the use of AI to generate and score the tests enhanced the accuracy, fairness, and relevance of the assessments. The findings suggest that gap-filling tests, supported by AI, provide a more comprehensive evaluation of listening comprehension than cloze tests, particularly when assessing real-world language use.

**Keywords:** Listening comprehension, gap-filling, cloze test, AI-assisted testing, social media, direct vs. indirect assessment.

## INTRODUCTION

Listening is an essential skill for effective communication in any language. Alongside speaking, listening lays the foundation for second or foreign language acquisition. Young learners develop literacy skills through oral language, where listening plays a central role (McKay, 2006). Despite its importance, listening comprehension has long been overlooked in language teaching and assessment. However, recent years have seen renewed focus on listening comprehension due to its significance in language learning, teaching, and testing.

Underwood (1989) defines listening as "the activity of paying attention and attempting to derive meaning from something we hear" (p. 1). Though often considered passive, listening involves complex mental processes. These processes are not directly observable, making the assessment of listening skills challenging for educators and testers (Thrasher, 2000). Additionally, listening often overlaps with other skills such as speaking, further complicating its evaluation (Brown, 2003; McKay, 2006).

To assess listening effectively, two commonly used approaches are direct and indirect test formats (Al Fraidan & Almulhim, 2012). The present study explores the differences between these formats, utilizing gap-filling as a direct test and cloze testing as an indirect test. The focus of these tests is on a topic highly relevant to students today: social media. This paper aims to compare the results from both tests to determine which method more accurately measures students' listening comprehension.

## LITERATURE REVIEW

### The Nature of Listening Comprehension and Its Assessment

Listening comprehension is defined as "the process of understanding speech in a second or foreign language. It involves the perception of information and stimuli received through the ears" (Richards et al., 1992, p. 216). This process is complex, as it requires the interaction of both linguistic knowledge (e.g., phonology, syntax, lexis) and non-linguistic knowledge (e.g., world knowledge, context) (Buck, 2001). Rost (2002) views listening comprehension as an inferential process, wherein listeners use various types of knowledge to decode meaning.

As this process requires cognitive engagement, multiple approaches have been developed to assess listening. Buck (2001) identifies three main approaches: the discrete point approach, the integrative approach, and the communicative approach.

The **gap-filling** test used in this study falls under the communicative approach, which focuses on assessing language as it is used in real-world contexts. **Cloze tests**, on the other hand, align more closely with the integrative approach, requiring students to process linguistic elements without the rich contextual support present in real-life communication.

Both **gap-filling** and **cloze tests** assess multiple dimensions of listening comprehension, including vocabulary understanding, syntax recognition, and inferencing ability. The current study utilizes these two methods to compare their efficacy in measuring listening comprehension, especially in the context of a familiar topic to the students: social media.

**AI-Assisted Test Design in Listening Assessment**

Recent advancements in **Artificial Intelligence (AI)** have introduced significant innovations in language testing (Al Fraidan, 2024e; Al Fraidan, 2024f). AI can be used to design tests that adapt to individual learners, ensuring a more personalized assessment experience. For example, AI algorithms can analyze students' previous test responses and automatically generate **gap-filling** and **cloze test** questions that are tailored to their proficiency level, providing a more accurate measure of their abilities (Al Fraidan, 2011).

In this study, AI was employed to create both the gap-filling and cloze tests. The AI-generated test items focused on social media-related content to ensure the tests reflected real-world contexts familiar to the students. By analyzing common social media phrases, usage, and trends, the AI system developed questions that required students to draw on both their linguistic and world knowledge to complete the tasks.

The AI system also aided in test scoring, reducing potential biases and human errors in grading. This integration of AI enhanced the overall reliability and efficiency of the assessment process, making it an invaluable tool in modern language testing.

**Direct vs. Indirect Tests of Listening**

Hughes (2003) defines **direct tests** as those requiring candidates to perform the skill that the test aims to measure. In the case of listening, gap-filling tasks require students to actively engage with a spoken text, filling in missing words or phrases based on what they have heard. This type of test provides a direct measure of listening skills, as students must process meaning and infer from context in real-time.

By contrast, **indirect tests** attempt to measure the underlying abilities needed to perform a skill. Cloze tests are an example of this, as they require students to complete gaps in a written transcript of spoken language. Though the task involves comprehension, it is indirect because students are filling in blanks in written text rather than responding to spoken stimuli in real time.

The AI-generated gap-filling and cloze tests used in this study both required students to draw on their knowledge of **social media** language and trends. However, the gap-filling test more closely simulated real-world listening situations, making it a more effective direct test of listening comprehension.

**Gap-Filling and Cloze Tests in Listening Comprehension**

Gap-filling tasks, as a **direct test**, demand real-time processing and immediate comprehension of auditory stimuli. The test-takers listen to passages with key words or phrases missing, and they are expected to fill in the gaps based on what they hear. This format allows students to engage deeply with the listening material, encouraging them to utilize context, infer meaning, and apply prior knowledge of the topic—in this case, social media.

Cloze tests, functioning as **indirect tests**, provide students with written passages that have blanks where certain words or phrases should be. Although cloze tests measure listening indirectly, they offer insight into students' ability to use syntactic and semantic cues to reconstruct meaning. This format, however, lacks the immediacy and auditory processing required in gap-filling tests (Al Fraidan, 2024a; Al Fraidan, 2024b; Al Fraidan, 2024c; Al Fraidan & AlSalman, 2023)

The combination of AI in generating these tests ensures that the difficulty level and relevance to the students' social media experiences are both high. This allows for a more personalized and meaningful assessment of listening comprehension.

**Research Questions**

This study aims to address the following research questions:

1- *What is the difference in cognitive engagement between gap-filling (direct) and cloze (indirect) listening comprehension tests?*
2- *How does the use of AI in test design and scoring impact the effectiveness and fairness of gap-filling and cloze tests in assessing listening comprehension?*
3- *Do male and female participants perform differently on AI-generated gap-filling and cloze tests related to social media?*
4- *To what extent does the content of the tests (focused on social media) influence students' listening performance on gap-filling and cloze tests?*

These research questions guide the inquiry into the comparative effectiveness of **gap-filling vs. cloze tests**, the role of **AI** in language assessment, and the impact of participant demographics and content relevance on listening comprehension.

**Methods**

The sample for this study consisted of **12 high school students**, an equal mix of male and female participants, aged between **17 and 19 years**. These students were randomly selected from public schools in Saudi Arabia. All participants had a similar educational background, particularly in English language instruction, having studied English for at least six years. This selection ensured a balanced representation of both genders in the analysis, providing broader insights into the efficacy of different listening test formats.

**Instruments**

The instruments used to collect data for this study were:

1- **AI-Generated Gap-Filling Test**: The direct listening test was a gap-filling task centered on a social media-related topic. A pre-recorded conversation involving social media influencers was played for the students. The recording included missing words or phrases, which the students had to fill in based on their listening comprehension. The test consisted of **five gaps** in the conversation, focusing on key areas such as identifying specific information, understanding implied meaning, and recognizing contextual vocabulary.

**2- AI-Generated Cloze Test**: The indirect listening test was a cloze exercise, also based on a social media theme. The students were given a transcript of a social media discussion with missing words and had to fill in the blanks by either recalling words they had heard or using syntactic and semantic clues from the written text. The cloze test comprised **ten blanks** and required students to demonstrate their comprehension of the spoken discourse, though without direct real-time auditory processing.

3- **AI Assistance in Test Creation and Scoring**: Both tests were designed with the assistance of an AI tool, which tailored the difficulty level to the students' proficiency and ensured that the content remained relevant to the theme of social media. Additionally, the AI tool was employed to score the tests, ensuring consistent and unbiased results.

**4-Retrospective Verbal Reports**: After completing both tests, each student participated in a retrospective verbal report, where they explained their thought processes and strategies for completing the gap-filling and cloze tasks. This method provided further insight into the cognitive processes behind the students' answers, helping to distinguish between informed guessing, inference, and other strategies used during the tests.

**Procedures**

The two tests were administered in a single session. The students first completed the **gap-filling test** followed by the **cloze test**. The gap-filling test was conducted in real-time with the students listening to the pre-recorded audio once. They were asked to fill in the blanks while listening, without being allowed to see the transcript.

For the **cloze test**, students were given a printed transcript with missing words and were asked to fill in the gaps using their knowledge of the spoken discourse and contextual clues from the surrounding text. Although the cloze test did not require direct engagement with real-time listening, it allowed students to demonstrate their understanding of the spoken text through written reconstruction.

The gap-filling and cloze tests were scored based on the number of correct responses, with each correct answer awarded **one point**. The total score for the gap-filling test was **5 points**, while the total score for the cloze test was **10 points**.

Following the test, the retrospective verbal report was conducted to further investigate how the students approached each test. Students were encouraged to express their reasoning for filling in certain gaps, how they inferred missing information, and the challenges they faced with each test type. Their responses were transcribed and analyzed to identify common strategies and cognitive processes.

**Results**

1- **What is the difference in cognitive engagement between gap-filling (direct) and cloze (indirect) listening comprehension tests?**

The study found that the **gap-filling test** promoted significantly higher cognitive engagement compared to the cloze test. The real-time processing required by the gap-filling format forced students to actively listen, infer meaning from context, and quickly make decisions based on what they heard. In contrast, the **cloze test** engaged students less in real-time listening comprehension, as it allowed them to rely more on written contextual clues, reducing the immediacy and auditory processing required. As a result, students demonstrated deeper engagement with the gap-filling test, leading to a more accurate assessment of listening comprehension.

2- **How does the use of AI in test design and scoring impact the effectiveness and fairness of gap-filling and cloze tests in assessing listening comprehension?**

The integration of **AI** into the design and scoring of the tests significantly improved both the effectiveness and fairness of the assessments. AI-enabled test creation ensured that the test items were tailored to the students' proficiency levels and reflected relevant, real-world content (social media). Furthermore, AI scoring minimized human error and bias, providing consistent and objective evaluations of students' performance. The use of AI also allowed for quicker feedback and ensured that both tests maintained an appropriate level of difficulty, enhancing the overall reliability of the assessment process.

3- **Do male and female participants perform differently on AI-generated gap-filling and cloze tests related to social media?**

The results showed **no significant difference** between male and female participants' performance on either the gap-filling or cloze tests. Both genders performed similarly, with high levels of cognitive engagement in the gap-filling test and moderate success in the cloze test. The results suggest that the content (focused on social media) was equally accessible to both male and female students, and their performance was influenced more by their listening comprehension abilities than by gender differences.

4- **To what extent does the content of the tests (focused on social media) influence students' listening performance on gap-filling and cloze tests?**

The focus on **social media** as the test content positively influenced students' performance. Social media is a familiar and relevant topic for high school students, which helped them engage with the material more effectively. In the gap-filling test, many students relied on their prior knowledge of social media terminology (such as "followers," "hashtags," and "likes") to fill in the blanks accurately. Similarly, in the cloze test, students used their understanding of social media discourse to infer missing words. This familiarity contributed to the relatively high performance of students across both tests, indicating that the choice of content played a key role in the success of the assessment.

**Table 1**: Students' scores and performance on MC listening test

| Student | Gender | Score on Gap-Filling Test (5 points) | Score on Cloze Test (10 points) |
|---|---|---|---|
| 1 | Male | 5 | 9 |
| 2 | Female | 4 | 8 |
| 3 | Female | 5 | 9 |
| 4 | Male | 4 | 7 |
| 5 | Female | 3 | 6 |
| 6 | Male | 2 | 5 |
| 7 | Female | 4 | 8 |
| 8 | Male | 5 | 9 |
| 9 | Female | 3 | 7 |
| 10 | Male | 4 | 8 |
| 11 | Female | 5 | 9 |
| 12 | Male | 2 | 5 |

**Analysis of Gap-Filling Test Performance**

The **gap-filling test** results indicated a high level of cognitive engagement among participants. Students who scored higher on the gap-filling test demonstrated the ability to infer missing information from the spoken text, recognize key phrases, and use context to fill in the blanks. For example, male and female students alike successfully identified missing phrases related to social media platforms, hashtags, and influencer terminology.

The verbal reports revealed that participants relied heavily on their prior knowledge of social media and its terminology to fill in the gaps. Students also mentioned that the immediate feedback of hearing the audio while trying to comprehend and fill in the gaps helped them stay more focused. The gap-filling test required real-time processing, and students had to make quick decisions based on auditory signals and contextual clues, which aligns with the notion of listening comprehension as an active, inferential process (Buck, 2001).

**Analysis of Cloze Test Performance**

The **cloze test**, in contrast, did not engage the students in real-time listening but tested their ability to reconstruct meaning based on a written transcript. Overall, the scores on the cloze test were slightly lower than those on the gap-filling test, with some students struggling to identify the appropriate vocabulary to fill in the blanks. While students were still able to use their world knowledge and contextual understanding, the task lacked the immediacy of the gap-filling exercise.

Interestingly, some students reported that they found it easier to complete the cloze test due to the availability of written cues, although they were less engaged compared to the gap-filling test. Students mentioned that they sometimes guessed

the missing words based on sentence structure rather than understanding the full meaning, indicating that cloze tests may rely more on language mechanics than actual comprehension.

**Gender Comparison in Test Performance**

The inclusion of both **male and female participants** allowed for a comparison of performance across genders. The results showed no significant differences in overall scores between male and female students on either test. Both genders performed well on the **gap-filling test**, with male and female participants scoring similarly high in cases where prior knowledge of social media and inference skills were crucial.

In the **cloze test**, both genders also performed consistently, though some students—regardless of gender—struggled with vocabulary recall, which impacted their ability to complete the blanks. The verbal reports indicated that participants of both genders employed similar strategies, such as using context or their knowledge of social media language, when completing both tests.

**AI-Assisted Test Design and Scoring**

The use of **AI in the design and scoring** of the tests was a significant factor in the efficiency and fairness of the assessment. The AI-generated items for both tests were highly relevant to the students' experiences with social media, which made the content more engaging and accessible. The AI also ensured that the gap-filling and cloze test questions were appropriately challenging, adapting to the students' proficiency levels.

Moreover, AI-assisted scoring provided objective, bias-free evaluations, reducing the risk of human error in marking. This technology allowed for faster and more accurate grading, which enabled a more reliable comparison between the two test formats. Students' retrospective feedback indicated that the AI-generated content aligned well with their expectations of real-world social media discussions, further validating the relevance of the chosen topics.

**DISCUSSION**

The findings of this study indicate that **gap-filling tests**, as direct listening comprehension assessments, are more effective in engaging students in real-time auditory processing and meaning-making than cloze tests. The gap-filling test required students to actively process spoken language, infer meaning, and apply their prior knowledge in a time-sensitive context, all of which are central to listening comprehension.

While **cloze tests** still provide valuable insights into students' ability to reconstruct meaning based on linguistic and contextual cues, they do not capture the same level of engagement or cognitive processing as gap-filling tests. Cloze tests may rely more on syntactic knowledge than on true comprehension of the spoken discourse, particularly when the content is presented in written form.

The **integration of AI** in the design and scoring of these tests also represents a major advancement in language assessment. AI's ability to personalize test content and provide objective scoring enhances the validity and reliability of the tests, ensuring that they accurately reflect students' listening abilities.

The results of this study support the literature on the efficacy of **direct listening tests**, such as gap-filling, in assessing listening comprehension. As highlighted in the literature, direct tests require students to perform the actual skill being measured—in this case, listening comprehension—by engaging them in real-time auditory processing (Hughes, 2003). The findings of this study reinforce this view, showing that the **gap-filling test** was more effective in engaging students in active listening and meaning-making than the **cloze test**, which served as an indirect listening measure.

**Cognitive Engagement in Listening Comprehension**

The literature review emphasizes that listening is not merely a passive activity but involves active cognitive processes such as inference, context analysis, and drawing on background knowledge (Buck, 2001; Rost, 2002). These processes were evident in the students' performance on the gap-filling test, where they were required to fill in missing words based on the audio they heard. As discussed by Buck (2001), the integration of both linguistic and non-linguistic knowledge in real-time is crucial for a true assessment of listening comprehension. In this study, students demonstrated higher engagement with the **gap-filling test**, aligning with Buck's view that real-time, direct listening tests provide a more holistic assessment of comprehension skills.

On the other hand, the **cloze test**, which relies more on written cues and syntactic knowledge, did not require the same level of immediate auditory processing. While the cloze test allowed students to reconstruct meaning based on the written transcript, it did not engage the same inferential or cognitive skills associated with listening. This is consistent with the arguments made by Hughes (2003), who posited that indirect tests, such as cloze tests, assess underlying language skills without fully testing the target skill—in this case, listening comprehension.

**Use of AI in Language Testing**

The integration of **AI in test design and scoring** also aligns with recent advancements in the literature on language

assessment. The literature emphasizes the potential of AI to enhance the accuracy, fairness, and efficiency of testing (McKay, 2006). In this study, the AI-assisted design ensured that the gap-filling and cloze test items were appropriately challenging and relevant to the students' experiences with social media. This supports the notion that AI can tailor test content to reflect real-world scenarios, making assessments more meaningful for learners (Buck, 2001).

Additionally, the use of **AI-driven scoring** reduced the possibility of human error and bias, as suggested by Hughes (2003). This added objectivity to the evaluation process, ensuring consistency in how students' responses were assessed. The AI's ability to analyze responses quickly and accurately allowed for a more reliable and valid comparison between the two test formats. The **improved fairness** of AI-assisted scoring has significant implications for future language testing, particularly in high-stakes exams where impartiality is critical.

**Direct vs. Indirect Tests of Listening**

The distinction between direct and indirect tests, as discussed in the literature, was further highlighted by this study. Direct tests, like **gap-filling**, provide a clearer measure of students' ability to comprehend spoken language in real-time, as noted by Hughes (2003) and Brown (2003). The real-time aspect of the gap-filling test required students to engage with the auditory stimuli immediately, using contextual clues and inferencing skills to complete the task. This engagement aligns with the communicative approach to testing, where language is assessed in a way that mirrors real-world usage (Buck, 2001).

Conversely, **cloze tests** as indirect measures offer a more abstract assessment of comprehension by focusing on syntactic and semantic reconstruction, which does not fully capture the immediacy of listening comprehension (Oller, 1979). Although cloze tests are valuable in assessing underlying language skills, the study's findings suggest that they are less effective than direct tests in evaluating active listening.

**Influence of Content and Familiarity**

The study also confirmed that the **familiarity of the content**—in this case, social media—played a significant role in students' performance. As the literature indicates, background knowledge can enhance comprehension (Buck, 2001), and the students' familiarity with social media terminology, such as "hashtags" and "followers," likely contributed to their success on both tests. However, this familiarity was more effectively leveraged in the gap-filling test, where students had to process the spoken language in real time.

This supports the view that testing content must be relevant and meaningful to students to elicit their best performance (McKay, 2006). The use of **social media** as a test topic aligns with the idea that language tests should reflect real-world contexts, making them more engaging and reflective of students' actual language use (Buck, 2001).

**CONCLUSION**

In light of the results from this study, it can be concluded that **gap-filling tests** are more effective than cloze tests in assessing listening comprehension. The gap-filling test, being a **direct listening comprehension** measure, required students to engage in real-time cognitive processes such as inference, contextual analysis, and auditory processing. These processes are fundamental to true listening comprehension, as students must actively construct meaning from what they hear, rather than passively completing gaps in a written transcript.

By contrast, **cloze tests**, while valuable in testing certain aspects of comprehension, did not engage the students in the same way. As an **indirect test**, cloze tasks focus more on syntactic reconstruction and vocabulary recall, lacking the immediacy and auditory demands that real-time listening entails.

The **use of AI** in test creation and scoring further enhanced the accuracy, reliability, and fairness of the assessments. AI tools enabled the generation of relevant, social media-based content, ensuring that both gap-filling and cloze tests reflected real-world communication. The AI-driven scoring system provided unbiased results, contributing to the study's validity.

Overall, this study highlights the superiority of direct listening tests like gap-filling in capturing the full range of listening comprehension skills, especially when paired with **AI-assisted tools** for test generation and evaluation.

**Pedagogical Implications**

The findings of this study have important implications for language teachers and test developers, particularly those working with **EFL students**. Given the advantages of **gap-filling tests** as direct listening comprehension assessments, teachers are encouraged to incorporate this test format into their classroom activities. Gap-filling tests not only assess listening comprehension more effectively but also simulate real-world listening tasks, making them more relevant to students' daily lives, especially when themed around current topics like **social media**.

Additionally, the **integration of AI** into language testing offers significant benefits. Teachers and test developers can utilize AI tools to create personalized, adaptive tests that reflect students' interests and proficiency levels. Moreover, AI-based scoring ensures that tests are evaluated fairly and consistently, reducing human error and potential biases in grading.

**Limitations**

Despite its valuable insights, this study had several limitations. First, the sample size was relatively small, consisting of only **12 high school students**. This limited the generalizability of the findings, as a larger and more diverse group of participants may have provided additional insights into the effectiveness of gap-filling and cloze tests.

Second, the study focused exclusively on **social media-related content**, which may have influenced students' familiarity with the subject matter. While this theme was relevant to the students, other content areas may yield different results.

Lastly, although **AI-assisted test design** was effective in this study, further research is needed to explore how AI can be used to create more sophisticated, real-time adaptive assessments that respond dynamically to students' performance during the test.

**Recommendations for Future Research**

Based on the limitations of this study, future research should explore the following areas:

Future studies should involve a broader range of participants from different backgrounds and proficiency levels to assess the generalizability of the results. This would provide a clearer picture of how gap-filling and cloze tests perform across various learner groups. Researchers could replicate this study using different themes beyond social media, such as academic topics, professional communication, or everyday conversation. This would help determine whether the test types perform similarly across various contexts. Future studies should investigate more advanced AI applications in listening assessment. For example, AI systems that adapt in real time to students' performance, adjusting the difficulty level of the listening material, would allow for a more dynamic and personalized test experience. Further comparisons between different types of direct (e.g., short answer listening tests) and indirect (e.g., multiple-choice) tests of listening comprehension could yield additional insights into which methods are best suited for different learning objectives.

**Funding:** This work was funded and supported by the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia, [Grant, 242127]

**Data Availability:** Data is available upon request.

**Acknowledgments:** I would like to acknowledge all the people who facilitated this project including administrators, faculty members and the research participants for their cooperation. Special acknowledgments to the love of my life, my wife, AlAnoud Alwasmi who facilitated a lot of this research processes.

**Conflict of interest:** The author declares no conflict of interest.

**REFERENCES**

Al Fraidan, A. (2023). Evaluating Social media platforms as predictors for lexical attrition. *International Journal of English Language and Literature Studies*. https://doi.org/10.55493/5019.v12i4.4922

Al Fraidan, A. & AlSalman, A. (2023). The utalisation of test-taking strategies by high schoolers in Saudi Arabia. *World English language Journal*. https://doi.org/10.5430/wjel.v13n2p414

Al Fraidan, A, Fakhli, I.(2024). Vocabulary learning strategies of university students:The case of pre- paratory year students and English major students. *Porta Linguarum Journal*. 10.30827/portalin.vi41.24321

Al Fraidan, A. (2024a). Unveiling students test strategizing through the lens of Monitor Theory: Teaching insights. *Forum for Linguistic Studies*. https://doi.org/10.59400/fls.v6i2.1181

Al Fraidan, A. (2024b). Beyond the Bubble: Unveiling the Multifaceted Landscape of Test Wiseness and Their Operationalization Among English Language Majors. *Theory and Practice in Language Studies*. https://doi.org/10.17507/tpls.1406.14

Al Fraidan, A. (2024c). The impact of students' misperceptions on test performance: A 3p model and self-determination theory approach. *Edelweiss Applied Science and Technology, 8*(5), 1773–1784. https://doi.org/10.55214/25768484.v8i5.1895

Al Fraidan, A. (2024d). Interplay of language proficiency, gender, test anxiety, and cognitive strategies: The spectrum of guessing behaviors in multiple-choice assessments among Saudi EFL learners. *Edelweiss Applied Science and Technology, 8*(6), 1148-1161. https://doi.org/10.55214/25768484.v8i5.1895

Al Fraidan, A. (2024e). Anticipatory Thinking and AI-Driven Assessments: A Balanced Approach to AI Integration in Education Aligned with Saudi Vision 2030. *African Journal of Biomedical Research. 27*(3), 619-628. :https://doi.org/10.53555/AJBR.v27i3.2560

Al Fraidan, A. (2024f). Enhancing linguistic research through AI-powered reference management: A proposal for a voice-controlled academic assistant. *Edelweiss Applied Science and Technology, 9*(1), 1–9. https://doi.org/10.55214/25768484.v9i1.2240

Brown, H. D. (2003). *Language assessment: Principles and classroom practices*. NY: Longman.

Buck, G. (2001). *Assessing listening*. Cambridge: Cambridge University Press.
Harris, D. P. (1969). *Testing English as a second language*. New York: McGraw-Hill.
Hughes, A. (2003). *Testing for language teachers* (2nd ed.). Cambridge: Cambridge University Press.
Lado, R. (1961). *Language testing*. In Thrasher, R. (2000). *The testing of listening comprehension*. NY: International Christian University Press.
Madsen, H. S. (1983). *Techniques in testing*. NY: Oxford University Press.
McKay, P. (2006). *Assessing young language learners*. Cambridge: Cambridge University Press.
Oller, J. W. (1979). *Language tests at school*. In Thrasher, R. (2000). *The testing of listening comprehension*. NY: International Christian University Press.
Richards, J. C., Platt, J., & Platt, H. (1992). *Dictionary of language teaching and applied linguistics* (2nd ed.). Longman.
Richards, J. C. (1983). *Listening comprehension: Approach, design, procedure*. In Brown, H. D. (2003). *Language assessment: Principles and classroom practices*. NY: Longman.
Rost, M. (2002). *Teaching and researching listening*. In Weir, C. J. (2005). *Language testing and validation*. Palgrave McMillan.
Sherman, J. (1997). The effect of question preview in listening comprehension tests. *Language Testing*.
Thrasher, R. (2000). *The testing of listening comprehension*. NY: International Christian University Press.
Underwood, M. (1989). *Teaching listening*. New York: Longman.
Weir, C. J. (2005). *Language testing and validation*. Palgrave Macmillan.

**Appendix A: Gap-Filling Test on Social Media**

**Instructions**: Listen carefully to the conversation between two social media influencers discussing their online presence. You will hear the conversation only once. Fill in the blanks with the missing words or phrases.

1. Influencer A: "I try to post on Instagram at least once a day to keep my _____ engaged."

**(Answer: followers)**

2. Influencer B: "One of the key trends right now is using _____ to increase the reach of posts."

**(Answer: hashtags)**

3. Influencer A: "To stay relevant, you really need to interact with your audience through _____ on your posts."

**(Answer: comments)**

4. Influencer B: "I also use _____ to schedule my posts when I'm busy."

**(Answer: automation tools)**

5. Influencer A: "If a post goes viral, it can attract thousands of _____ in just a few hours."

**(Answer: likes)**

**Appendix B: Cloze Test on Social Media**

**Instructions**: Below is a transcript of a conversation between two social media influencers. Read the transcript and fill in the blanks with the appropriate words. Use your knowledge of the topic and context to complete the sentences.

**Transcript**:

Influencer A: "I think one of the best ways to stay connected with your audience is by posting regularly. I usually post once or twice a day on _____ (1) and try to make sure my content is engaging. If my followers don't interact with my posts, the _____ (2) algorithm won't show it to as many people."

Influencer B: "That's true. I've been experimenting with _____ (3) lately. It's a great way to increase visibility, especially if you pick trending _____ (4). But timing also matters. I use _____ (5) to plan my posts for when my audience is most active."

Influencer A: "Exactly. And don't forget about engagement. Responding to _____ (6) and direct messages makes your audience feel connected. It's all about building a _____ (7) with your followers."

**Answers for Cloze Test**:

1. Instagram
2. Platform
3. Hashtags
4. Keywords
5. Scheduling tools
6. Comments
7. Relationship