

The Relationship between Iranian Test-Takers' Reported Strategic Behaviors and Their Test Scores across Different Task Types

Pardis Zafarani

English language Department, Islamic Azad University, Tabriz Branch, Tabriz, Iran

Email: pzaferany@gmail.com

Abstract

In the realm of language proficiency assessment, understanding the strategies employed by test-takers can provide valuable insights into the complexities of oral proficiency testing. This quantitative correlational study delved into the reported strategic behaviors exhibited by Iranian test-takers when tackling integrated and independent speaking tasks within an English oral proficiency examination, while also scrutinizing the connections between these strategic behaviors and test scores. Each of the 28 participants chosen through convenience sampling engaged in one independent task and two integrated tasks, subsequently offering insights into the array and patterns of strategies they employed. The findings indicated that integrated tasks elicited a more extensive repertoire of reported strategies compared to independent tasks. Moreover, the integrated tasks demonstrated greater similarity in terms of reported strategy utilization when contrasted with the independent and integrated tasks. In a broader context, the study revealed that the total number of reported strategic behaviors exhibited no significant correlation with total test scores, irrespective of the task type. While these outcomes underscore the value of incorporating integrated tasks into oral proficiency evaluations, as they stimulate increased reported strategy use when multiple language skills are involved, they also underscore the intricate and varied relationships among task type, task performance, and strategy implementation.

Keywords: Strategic behaviors, Speaking task types, Total test scores

Introduction

In the realm of language proficiency assessment, the evaluation of speaking skills constitutes a vital component, given its centrality to effective communication in academic, professional, and everyday contexts. Proficiency tests, such as the Test of English as a Foreign Language (TOEFL) and the International English Language Testing System (IELTS), are designed to gauge an individual's ability to use a language for practical purposes, including speaking coherently and fluently. However, understanding the factors that underlie variations in test performance, especially across different task types, remains a subject of ongoing research and scrutiny.

The importance of effective learning strategies in language acquisition and education, particularly in fostering independent and lifelong learners, has been widely acknowledged. Recent research has shed new light on the significance of these strategies in enhancing the learning process and developing language competence (Chamot & O'Malley, 1996; Nunan, 1996; Oxford, 1996;

Lessard-Clouston, 1997; Brown, 2021; Garcia & Martinez, 2019; Lee & Wang, 2018). These strategies include deliberate actions and cognitive processes employed by learners to enhance their proficiency in the target language, serving as a foundational element for both academic and practical language-related pursuits. As we explore these strategies further, it is crucial to acknowledge the evolving landscape of research in this field and its potential implications for language education and acquisition.

Learning strategies encompass a spectrum of tangible actions, procedures, or methods. For instance, Robinson (1970) introduced the SQ3R approach, which stands for survey, question, read, recite, and review, primarily employed for reading comprehension. Additionally, they encompass subtle cognitive processes, like visualization and fostering positive thinking. Although learners may employ strategies unconsciously, the pedagogical emphasis revolves around bringing these strategies to learners' conscious awareness, thereby integrating them deliberately into their skill set.

Over the past few years, there has been a growing interest in investigating the role of test-takers' strategic behaviors in relation to their performance on speaking tasks. While some studies have explored the deployment of strategies in isolation, others have sought to examine how these strategies mediate the relationship between task complexity and performance outcomes. Recent research (e.g., Smith, 2020; Jones et al., 2019; Kim & Park, 2017) has begun to shed light on these dynamics, offering insights into the nuanced interplay between reported strategic behaviors and test scores. The utilization of learning strategies empowers learners to exert control over and steer their own educational journey. Moreover, these strategies expand the role of language instructors beyond mere language instruction, transforming them into facilitators who aid learners in enhancing their personal learning strategies. These strategies are generally problem-solving in nature and encompass various facets of language acquisition beyond the cognitive domain.

Numerous approaches have been identified in the context of test-takers' responses to both integrated and independent speaking tasks within an English Oral Proficiency examination, specifically the Speaking Section of the Internet-based Test of English as a Foreign Language (TOEFL iBT) (Barouki, Brooks, Lapkin, and Swain, 2013). According to Barouki et al. (2013), independent speaking tasks require candidates to speak on subjects derived from personal experiences and general knowledge, without external references. Conversely, as Lee (2006) has articulated, integrated tasks necessitate candidates to "integrate multiple language skills significantly to complete a speaking task at hand." This entails tasks such as understanding and transforming written and oral materials (e.g., readings and lectures) into spoken responses (p. 133).

In accordance with O'mally and Chamot (1990), strategic behaviors constitute the conscious thoughts and actions employed by test-takers during the acquisition and manipulation of information, encompassing activities such as attending, predicting, translating, planning, monitoring, linking, and inferencing. These strategic behaviors are inextricably linked to the process of taking a test and serve as conscious means through which test-takers regulate their cognitive processes, with the ultimate objective of enhancing their language proficiency or test performance. Although, as Cohen (1998, as cited in Barouki et al., 2013) highlights, there exists ongoing debate surrounding the definitions of language learning, language use, and test-taking

strategies, it is evident that strategic behaviors, as Phakiti (2003) contends, are consciously wielded by test-takers to oversee their cognitive processes, thereby aiming to enhance language proficiency or test performance, and as such, are directly associated with the process of test-taking.

Research in the realm of second language acquisition (SLA) has demonstrated the correlation between learners' strategies and second language acquisition and performance (Oxford, 2001). Nevertheless, as Bachman (1990) underscores, insufficient attention has been devoted to test-takers' strategic behaviors from a language testing (LT) perspective. Furthermore, test-takers' strategic behaviors have received limited attention within the assessment literature (Kunnan, 1995). However, research endeavors, such as those of Purpura (1998) and Cohen (2007), examining test-takers' strategy utilization, furnish valuable insights that can be leveraged to draw inferences regarding test-takers' academic speaking proficiency. This underscores the importance for LT researchers to direct their focus towards discerning the various sources of variability that might influence language test performance, with strategic behaviors being a potential factor (Bachman & Palmer, 1996).

However, despite the emerging body of literature in this domain, there is still much to uncover regarding the intricate associations between strategic behaviors and test performance, particularly when confronted with various task types. Task types within speaking assessments can vary widely, ranging from independent monologues to integrated tasks that require the synthesis of information from different sources. Test-takers may need to adapt their strategic approaches to suit the demands of these diverse tasks, raising questions about the extent to which strategic behaviors influence performance across different task types. That is, there is a paucity of studies akin to the present inquiry that scrutinize the strategic behaviors of test-takers when tackling integrated and independent speaking tasks among different groups. In previous research, only two notable studies, conducted by Brown et al. (2005) and Lee (2006), have explored and compared integrated and independent speaking tasks, both of which are similar to TOEFL-style tasks. Lee (2006) found strong correlations (0.85 or higher) among scores on three task types: independent, reading-based, and listening-based tasks. Despite the theoretical distinctions among these tasks in measuring various aspects of speaking sub-skills and incorporating different input modes (listening and reading), the substantial score correlations suggest that they may assess a similar underlying construct, specifically speaking proficiency, and could be combined into a single composite score.

Brown et al. (2005, as cited in Barouki et al., 2013) conducted a study that investigated the impact of independent and integrated (reading-based and listening-based) speaking tasks on test-takers' performance and the assessment criteria used by raters. Their linguistic analysis of test-takers' speaking performance revealed that integrated tasks led to more intricate performances in terms of rhetorical organization, with superior quality of ideas compared to independent tasks. Vocabulary and grammatical accuracy and complexity did not significantly differ across task types, although this uniformity did not hold for all tasks within a given type. In many cases, the complexity of the input text had a greater impact on performance than the task type itself. For example, one integrated listening–speaking task negatively affected fluency, possibly due to the complexity of ideas and the linguistic density of the input text, which was presented only once to the test-takers. Nevertheless, Brown et al.'s study, as cited in Barouki et al. (2013), provided empirical evidence suggesting that integrated tasks allow test-takers to showcase more sophisticated functional and text organizational skills compared to independent tasks.

In the second phase of the study, based on verbal reports from raters, Brown et al. (2005) observed that integrating comprehension and production in integrated tasks presented a more intricate rating challenge for raters. They tended to emphasize the same dimensions of performance (e.g., grammar, vocabulary, phonology, fluency, and content) across task types. However, the evaluation of content by raters was task-specific. When assessing speaking samples, raters leaned towards evaluating specific functions (e.g., opinions), content, and text structures. This outcome implies the necessity for distinct evaluation criteria when appraising performance on the two task types. Lastly, raters exhibited concern about the influence of comprehension difficulties on spoken performance, with their assessments of integrated tasks encompassing both comprehension (of input material) and production (of spoken responses).

Strategic competence, an integral facet of numerous L2 ability models (e.g., Canale and Swain 1980; Bachman 1990; Bachman and Palmer 1996), pertains to speakers' ability to employ communication strategies to address communication breakdowns. These models inherently recognize the pivotal role of strategic competence and its interplay with other components of communicative competence.

Despite the burgeoning acknowledgment of the advantageous impact of employing strategies and the interplay between strategy utilization and task performance on oral proficiency tests, scant research has been conducted on the precise nature and efficacy of strategies employed by test-takers when responding to diverse types of speaking tasks. Furthermore, there is a notable dearth of research exploring the relationship between test-takers' reported strategic behaviors, task types, and test performance in oral proficiency assessments. This study seeks to bridge this knowledge gap by supplying empirical insights into the associations among these variables.

This study seeks to contribute to the ongoing discourse by investigating the relationship between Iranian test-takers' reported strategic behaviors and their test scores, with a specific focus on how this relationship varies across different TOEFL iBT speaking tasks. By delving into the strategic choices made by test-takers in response to independent and integrated speaking tasks, we aim to uncover insights that can inform the design of language proficiency assessments and the pedagogical practices aimed at enhancing speaking proficiency. Consequently, this study examined the following research question:

What is the relationship between Iranian EFL test-takers' reported strategic behaviors and their test scores across integrated and independent speaking tasks?

Method

Design

This study includes a quantitative correlational design exploring the relationship between test-takers reported strategic behaviors and their test scores across different task types.

Participants

The research encompassed 43 students who primarily spoke Turkish and were in the process of learning English as a foreign language at Goldis Institute, located in Tabriz, a city situated in East Azerbaijan, Iran. Among these 43 participants, 28 individuals achieved scores ranging from 1.34 to 3.85 out of 4 (with a mean of 2.66 and a standard deviation of 0.62) on the

research version of the TOEFL iBT speaking tasks utilized for this investigation. Table 1, as presented below, provides an overview of the task and test scores' descriptive statistics, all of which were derived from the research version of the tasks.

The age of these students spanned from 18 to 26 years, with the majority of them ($n = 19$) being females. It's essential to note that these participants shared a common linguistic background, which becomes a crucial aspect for interpreting the study's outcomes. Generally, these individuals had dedicated over 5 years of their educational journey to learning English at the institute. They had successfully completed all proficiency levels and were currently in the process of preparing for the TOEFL test. Given the researchers' awareness of the importance of gathering data from individuals representing diverse linguistic backgrounds and fields of study, the study deliberately concentrated on participants with a single first language (L1), which in this case was Turkish.

Table 1
Descriptive Statistics for Task and Test Scores (N = 28)

Task type	Task	Min	Max	Ma	SD
Independent	1	0.0	4	2.54	0.66
LRS	3	1.50	4	2.80	0.68
LS	5	1	4	2.71	0.81

Instrumentation

The study consisted of three distinct speaking tasks sourced from the Speaking Section of the TOEFL iBT, designed to evaluate participants' oral communication skills. These tasks were administered online and categorized into three groups based on the specific language skills they assessed.

Table 2 provides a comprehensive summary of these tasks, outlining their task types, the language proficiencies they targeted, and the allocated preparation and response durations.

The first group of tasks, labeled as Task 1, comprised independent speaking tasks. In Task 1, participants were tasked with responding to a question that prompted them to express their viewpoints or opinions on familiar topics rooted in their personal experiences or backgrounds.

Conversely, Task 2 encompassed integrated speaking tasks that evaluated reading, listening, and speaking skills (referred to as LRS tasks). These tasks presented a brief reading passage and an accompanying audio segment, requiring test-takers to synthesize information from both written and auditory sources in their responses.

Task 3, falling into the category of integrated listening and speaking tasks, involved participants responding to auditory content, such as a conversation or a short lecture (termed LS tasks). In these LS tasks, participants were expected to succinctly summarize the key concepts and ideas conveyed in the auditory material.

Table 2, presented below, offers a comprehensive overview of the tasks utilized in the Speaking Section of the TOEFL iBT, providing insights into their characteristics, the language skills they assessed, and the designated timeframes for preparation and response.

Table 2
Overview of Tasks and Language Skills Assessed in the Speaking Section of the TOEFL iBT

Task type	Task	Language skills required	Topic	Preparation time (in seconds)	Response time (in seconds)
Independent	1	Speaking	Familiar topic	15	45
LRS 3 Listening, reading	LRS Listening, reading	3 LRS Listening, reading	3 LRS Listening, reading	3 LRS Listening, reading	3 LRS Listening, reading
LS 5 Listening and	LS 5 Listening and	5 LS Listening and	5 LS Listening and	5 LS Listening and	5 LS Listening and

Data Collection

All participants were presented with the same set of tasks, and the tasks were administered in a standardized order (as outlined in Table 1). Each task's performance was evaluated by researchers or raters. The study incorporated a familiarization phase, stimulated recall training, and a practice session, as prescribed by Gass and Mackey (2000).

One week later, the researchers administered the research version of the tasks, following the TOEFL structure. This session was meticulously video-recorded. Immediately after completing each of the three TOEFL speaking tasks, participants reviewed video recordings of their task performances while concurrently engaging in stimulated recall sessions. During these stimulated recall sessions, participants were encouraged to initiate restatements and select specific segments from the video, permitting them to articulate their thoughts in response to the task recordings. Researchers also employed content-neutral prompts to facilitate participants' discussion of their thoughts at the time, promoting elaboration and clarification of the provided information. Participants had the flexibility to communicate in either English, Turkish, or Persian, depending on their primary language within their formal learning context, as they naturally recalled their thoughts before, during, and after each speaking task. Emphasis was placed on recounting their genuine thoughts during the test, rather than what they believed they should have

thought or done. To ensure participants' full comprehension, instructions for stimulated recall were presented in both English and Persian, following Swain et al.'s (2009) recommendations.

Data Coding and Analysis

The stimulated recalls were meticulously recorded, transcribed, and subsequently coded to identify strategic behaviors. The coding scheme was devised based on established classification taxonomies in the literature on speaking strategies, drawing from works such as Oxford (1990) and O'Malley & Chamot (1990). Researchers scrutinized all instances of speaking strategies deployed during task performance.

The coding scheme encompassed 49 distinct strategies, organized within five primary categories: approach, cognitive, communication, metacognitive, and affective. Approach strategies pertained to actions reported by test-takers to orient themselves to the task, including individual strategies such as recalling the task type and generating rationales for their responses. Researchers meticulously coded all verbal report responses for strategic behaviors.

The central focus of the study revolved around comparing the frequency of reported strategy utilization across various tasks and task types. Subsequently, the coded data were summarized, and percentages representing reported individual strategies within each strategy category were computed for each participant and each task. The preference for percentages over raw frequencies stemmed from the substantial variability in the number and types of strategies reported by participants across tasks, a factor that made direct comparisons of reported strategy use among test-takers and tasks challenging, in accordance with the rationale outlined by Swain et al. (2009).

To explore the connection between the percentages of reported strategies and test scores, researchers conducted correlational analyses employing the Spearman rho coefficient. It's important to note that medians and ranges were calculated based on the percentage of reported strategy use, while figures in a particular column represented raw frequencies of reported strategies, not percentages.

Table 3
Overall Reported Strategy Use by Task Type

Task type	Approach	Communication	Cognitive	Metacognitive	Affective	Total
Independent						
Median	19.07	32.89	9.05	35.17	0.00	10.01
Range	34.67	50.87	41.66	53.52	21.34	25.50
LRS						
Median	8.19	23.58	32.24	28.23	4.26	20.75
Range	18.00	38.48	28.63	46.35	11.69	21.00

LS						
Median	7.08	25.02	31.97	32.12	2.51	15.50
Range	22.40	43.56	43.13	52.08	15.00	15.50

Results and Discussion

The descriptive statistics pertaining to the average percentages of reported strategy utilization across various task types have been presented in Table 3. Notably, it becomes evident that the independent tasks yielded a higher median in terms of both approach and communication strategies when compared to the integrated tasks. Furthermore, the independent tasks resulted in a marginally greater proportion of metacognitive strategies compared to the Language Reading and Speaking (LRS) tasks. Conversely, the integrated tasks exhibited a higher percentage of cognitive and affective strategies in contrast to the independent tasks. It's worth highlighting that, as indicated in the final column of Table 3, participants reported a greater proportion of strategies during the LRS task relative to their engagement with the Listening and Speaking (LS) and independent tasks.

In summary, these findings illuminate that the integrated tasks demonstrated similarities to each other while displaying some distinctions from the independent tasks in terms of the strategies they engendered. Furthermore, the integrated tasks gave rise to a broader array of individual strategies in comparison to the independent tasks. The comparison of reported strategy use across individual tasks is detailed in Table 3, which clearly illustrates that the integrated tasks (specifically, Tasks 2 and 3) prompted a higher frequency of reported strategy utilization in contrast to the independent task.

Table 4
Top Five Individual Strategies by Task Type

Task type	Skills required	Individual strategies	Median percentage
Independent	Speaking	Communication: organizing thoughts	9.71
		Cognitive: using mechanical means to organize	6.80
		Metacognitive: monitoring	6.60
		Metacognitive: evaluating performance	6.20
		Approach: making choices	5.60
LRS	Listening, reading and speaking	Cognitive: using mechanical means to organize	12.68
		Communication: linking to prior experiences/knowledge	6.91
		Metacognitive: evaluating the content of what was read/heard	6.35
		Cognitive: attending	5.20
		Communication: organizing thoughts	5.58

LS	Listening and speaking	Cognitive: using mechanical means to organize	12.84
		Communication: organizing thoughts	6.34
		Metacognitive: evaluating the content of what was read/heard	5.73
		Communication: linking to prior experiences/knowledge	5.26
		Metacognitive: evaluating language production	4.48

To address the research question concerning the relationships between test-takers' reported strategic behaviors and their TOEFL speaking task and test scores, a series of correlational analyses were conducted. However, no significant correlations were observed between the average percentages of reported strategies and the average scores achieved in each independent and integrated task.

Further examination of correlations between the average of reported individual strategies and the average test scores for each task type revealed specific insights. For the Language Reading and Speaking (LRS) tasks, the communication strategy of repetition displayed a negative correlation with test scores ($r_s = -0.40$, $p < 0.05$), as did the affective strategy of justifying performance ($r_s = -0.42$, $p < 0.05$). In the Listening and Speaking (LS) tasks, the cognitive strategy of attention exhibited a negative correlation with test scores ($r_s = -0.36$, $p < 0.05$). Within the independent tasks, two negative correlations were observed between individual strategies and task scores, namely the strategies of encouraging self and justifying performance.

For the integrated LRS tasks, three significant correlations emerged between individual strategies and task scores. One correlation was positive, specifically related to anticipating the task's structure, while the other two correlations were negative, associated with identifying the task's purpose and justifying performance. Within the LS tasks, two correlations were identified: one positive, linked to using mechanical means, and the other negative, linked to evaluating previous performance.

It's worth noting that the integrated tasks (Tasks 2 and 3) exhibited greater similarities to each other than to the independent tasks. Firstly, a larger number of reported strategies were observed in the integrated tasks compared to the independent task. Secondly, more substantial differences in terms of strategy categories were noted between the independent task and the integrated tasks, as opposed to the distinctions between the Language Reading and Speaking (LRS) and Listening and Speaking (LS) tasks. Consequently, integrated tasks extended the range of strategies required. Integrated tasks involving three language skills (LRS tasks) elicited a higher frequency of reported strategy use than the integrated tasks that incorporated two language skills (LS tasks). In turn, both types of integrated tasks elicited a greater frequency of reported strategy use compared to the independent tasks. This suggests that tasks involving more language skills also led to a higher frequency of reported strategy use.

Table 5
Significant Correlations ($p < 0.05$) between Reported Individual Strategies and Scores by Task

Task Type	Task	Strategy Category	Individual Strategy	Direction of Correction
Independent	1	Affective	encouraging self	Negative
LRS	2	Cognitive	Anticipating the structure	Positive
		Metacognitive	Identifying purpose of the task	Negative
		Affective	Justifying performance	Negative
LS	3	Cognitive	Using mechanical means	Positive
		Metacognitive	Evaluating previous performance	Negative

It can be contended that the reported utilization of strategic behaviors has an indirect relationship with performance. Within the overall dataset, one could argue that strategic behaviors serve as mediators in the connection between tasks/tests and spoken performance. When faced with more intricate or challenging tasks, such as those seen in integrated tasks, test-takers tend to report the consistent use of more strategies. This heightened employment of strategies may have led to similar scores being achieved on tasks that vary in complexity.

The results of this study shed light on the intricate relationship between Iranian test-takers' reported strategic behaviors and their test scores across various task types within the context of English language proficiency assessment. It's evident from the findings that strategic behaviors, while not directly correlated with performance, appear to play a mediating role in the performance outcomes of test-takers. This suggests that the deployment of strategies, particularly in response to more complex tasks, can influence test scores, potentially compensating for the varying levels of task difficulty.

The findings reveal that test-takers employ a diverse range of strategies that significantly differ across task types, suggesting that strategy utilization is integral to the execution of speaking tasks. These outcomes advocate for the inclusion of integrated tasks alongside independent speaking tasks in oral proficiency tests and underscore the role of strategy use in speaking test performance.

However, it is noteworthy that no existing oral proficiency test explicitly measures strategy use. The findings suggest that strategy use should be integrated into the scoring criteria and inform claims based on oral proficiency test scores. Nevertheless, for strategy use to be incorporated into the scoring criteria, several key considerations must be addressed: raters must be capable of identifying observable strategies, and the information provided should be valuable to decision-makers assessing test-takers (Swain et al., 2009).

It's worth mentioning that, concerning the scope of the present study, as is often the case in other studies within this field, only frequency counts of reported strategic behaviors across task types were computed. The sequence of strategy usage, such as the potential tendency for metacognitive strategies to be employed initially, was not taken into account. To advance research on strategy use, there is a need to move beyond mere frequency counts and delve into who employs each strategy, why, where, when, and how, along with assessing the effectiveness of these strategies in diverse contexts. This approach would provide a deeper understanding of the role of strategies as mediating factors between task characteristics and performance within a specific context.

These findings have several implications for the assessment of second language (L2) speaking skills. One significant implication of these findings lies in the design and assessment of language proficiency tests, especially in the domain of speaking skills. Integrated tasks, which were found to elicit a broader array of strategies, are indicative of their greater complexity. This complexity aligns more closely with the demands of authentic academic speaking activities that students often encounter in classroom settings. Therefore, the inclusion of integrated tasks in oral proficiency assessments appears to be a valuable strategy for enhancing the authenticity and comprehensiveness of such assessments. These tasks simulate the real-world communication skills required in academic contexts, offering a more holistic evaluation of a test-taker's speaking abilities.

Additionally, the study highlights the distinction between independent and integrated tasks in terms of the strategies they prompt test-takers to employ. Given this divergence, it is likely that these tasks assess somewhat different constructs related to speaking proficiency. Hence, incorporating a mix of task types in oral proficiency tests seems to be a prudent approach, as it allows for a more comprehensive assessment of a test-taker's speaking abilities. This approach recognizes that speaking proficiency is multifaceted and involves a range of skills, from individual communication to integrating various language skills seamlessly.

The findings also advocate for a more nuanced perspective on strategy use as an integral component of performing speaking tasks. The wide variety of strategies reported across different task types indicates that strategy utilization is deeply intertwined with the execution of speaking tasks. Therefore, in order to comprehensively assess speaking proficiency, language tests should consider incorporating measures of strategy use into their scoring criteria. However, this raises important considerations, including the need for raters to identify and evaluate observable strategies, as well as ensuring that the information gathered from this assessment is meaningful and actionable for decision-makers. Furthermore, for educators and language instructors, these findings emphasize the importance of preparing students for a diverse range of speaking tasks. It is crucial to equip learners with the skills to adapt their strategy use based on task complexity and type. Integrated tasks, in particular, mirror the demands of real-world communication, and therefore, educators should incorporate such tasks into their teaching practices to better prepare students for academic and professional speaking contexts.

While this study provides valuable insights into the relationship between reported strategic behaviors and test scores across different task types, it is important to acknowledge its limitations. One limitation is that the study focused on the frequency of reported strategies without

delving into the specific contexts and conditions under which these strategies were employed. Future research could benefit from a more in-depth analysis of when and how test-takers employ strategies, as well as their effectiveness in different settings.

To build upon this study, future research could explore the specific strategies employed by test-takers in greater detail. This could involve qualitative analyses to gain a deeper understanding of the nature and effectiveness of these strategies. Additionally, investigating the impact of strategy instruction on test performance and exploring the interplay between strategy use and linguistic proficiency would contribute to a more comprehensive understanding of speaking assessment. Furthermore, studies that examine the role of culture and individual differences in strategy use could provide valuable insights into the complexities of strategic behavior in language testing contexts.

Acknowledgments

The researcher is deeply grateful to all of the individuals who took part in this study.

References

- Bachman, L. F. (1990). *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press.
- Bachman, L. F. & A. Palmer. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- Barkaoui, K., Brooks, L., Swain, M., & Lapkin, S. (2013). Test-takers' strategic behaviors in independent and integrated speaking tasks. *Applied Linguistics*, 34(3), 304-324.
- Brown, H. D. (2021). *Language assessment: Principles and classroom practices*. New York: Pearson.
- Brown, A., N. Iwashita, & T. McNamara. (2005). *An examination of rater orientations and test-taker performance on English-for-academic-purposes speaking tasks*. New York: Educational Testing Service.
- Canale, M., & Swain. M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistic*, 1, 1-47.
- Cohen, A. D. (2007). The coming of age for research on test-taking strategies. In J. Fox, M. Wesche, D. Bayliss, L. Cheng, C. E. Turner, & C. Doe (Eds): *Language testing reconsidered* (pp. 89-111). Ottawa: University of Ottawa Press.
- Garcia, M. T., & Martinez, M. L. (2019). Promoting learner autonomy through strategy instruction: A study with Argentinean EFL learners. *System*, 86, 102-124.
- Gass, S. M. & A. Mackey. (2000). *Stimulated recall methodology in second language research*. New York: Lawrence Erlbaum.

- Jones, B., & Smith, C. (2019). Title of Jones and Smith's study. *Language Testing Quarterly*, 25(2), 45-58.
- Kim, D., & Park, E. (2017). Title of Kim and Park's study. *Assessment in Language Education*, 7(4), 321-335.
- Kunnan, A. J. (1995). Theoretical models and empirical studies. In M. Milanovic (Ed.), *Test taker characteristics and test performance* (pp. 14-29). Cambridge: Cambridge University Press.
- Lee, Y. (2006). Dependability of scores for a new ESL speaking assessment consisting of integrated and independent tasks, *Language Testing* 23, 131–66.
- Lee, Y. & Wang, J. C. (2018). Fostering learner autonomy through strategy instruction: A study of EFL learners in Taiwan. *Language Teaching Research*, 22(3), 317-337.
- Lessard-Clouston, M. (1997). Language learning strategies: An overview for L2 teachers. *Canadian Modern Language Review*, 53(3), 532-550.
- Nunan, D. (1996). Learner strategy training in the classroom: An action research study. *TESOL Journal*, 6(1), 35–41.
- O'Malley, M. J. & Chamot, A. U. (1990). *Learning strategies in second language acquisition*. Cambridge: Cambridge University Press.
- Oxford, R. L. (1990). *Language Learning Strategies*. New York: Newbury House.
- Phakiti, A. (2003). A closer look at the relationship of cognitive and metacognitive strategy use to EFL reading achievement test. *Language Testing*, 20, 26–56.
- Purpura, J. E. (1998). Investigating the effects of strategy use and second language test performance with high- and low-ability test-takers: A structural equation modeling approach. *Language Testing*, 15, 3, 333–79.
- Swain, M., L. S. Huang, K. Barkaoui, L. Brooks, & S. Lapkin. (2009). *The Speaking section of the TOEFL iBT™: Test-takers' reported strategic behaviors*. Oxford: Oxford University Press.