Original Paper

Test-taking Strategies on Reading Comprehension Tests: A Review of Major Research Themes

Mohammed S. Assiri[1*] & Emad A. Alodhahi[2]

[1] English Department, King Khalid University, Abha, Saudi Arabia

[2] English Language Center, Institute of International Languages, Riyadh, Saudi Arabia

[*] Mohammed S. Assiri, English Department, King Khalid University, Abha, Saudi Arabia

Abstract

There is a growing body of thought and research on strategy use on tests of reading comprehension. Nevertheless, there have been few research reviews that have treated major themes involved in thought and research on test-taking strategies, specifically in the context of reading comprehension. Hence, this paper reviews the themes that are central to the discussion of strategy choice and use on reading comprehension tests. Research themes that form the foci of the paper include discussion of test-taking strategies as they relate to the process of reading test-taking, formats of reading tests, validation of reading tests, level of language proficiency, and performance on tests of reading comprehension.

Keywords

test-taking strategies, reading comprehension, language tests, test format, test validation, language proficiency, test performance

1. Introduction

The last three decades have witnessed a proliferation of research on both language learning and language use strategies. Not until very recently have scholars and researchers become increasingly interested in the kinds of strategies test takers employ when responding to a form of language assessment. This has led to ongoing research to explore the nature of these strategies and how they relate to a multitude of factors involved in the processes of language test-making and -taking. The main goal of test-taking strategy research "is to arrive at a series of empirically validated suggestions for what respondents need to do in order to enhance their performance on tests" (Cohen, 2012, p. 99). The motive to write this review was substantiated by the authors' notice that both thought and research on the use of test-taking strategies on reading tests have generally followed certain strands worth being

spotlighted in a review article.

Accordingly, this paper offers an overview of the major themes that have dominated both the discussion and research on strategy use on reading tests. In so doing, the paper is organized into five sections. In the first section, a discussion of how strategies in general are used on reading tests is presented. In section two, it is explicated how test takers employ textual and technical aspects of reading tests to decide how to go about responding to question items. The third section discusses the premise of establishing test validity on the basis of test-taking strategy use. Section four centers on differences between high- and low-proficiency readers in how they approach reading tasks. In the last section, a discussion is made of how test-taking strategies, in their two forms (i.e., test-management and test-wiseness), mediate between language competence and test performance. The paper concludes with a summary of the research themes discussed throughout the paper.

2. Five Major Research Themes

2.1 Reading Test-Taking and Strategy Use

The language use strategies employed by learners on the various forms of language assessment are referred to as test-taking strategies (Cohen, 1998). To be specific, test-taking strategies are techniques that test takers resort to with the aim of getting correct answers on a given form of language assessment (cf., Cohen, 1986; Cohen & Upton, 2006). The successful use of these strategies does not necessarily imply mastery of the testing task at hand, as Cohen (1986) explains clearly when suggesting that test takers may get their answers to a multiple-choice reading test right "without fully or even partially understanding the text" (p. 132). In a later article, Cohen (1992) notes that test-taking strategies represent processes that test takers can have control over by selecting what they believe would help them tackle a test question, suggesting that test-taking strategies are conscious processes. He adds that these strategies can either be a short move (e.g., looking for a clue that links the information in the question to that in the reading text) or a long one (e.g., reading the whole text after reading the questions).

Scholars (e.g., Allan, 1992; Cohen, 2014; Phakiti, 2008) have referred to two categories of test-taking strategies: test-management strategies and test-wiseness strategies. The former call for logical and purposeful response behaviors, are reflective of the underlying competence, and are responsive to the underlying construct being assessed, whereas the latter involve the use of textual and/or technical aspects of the test to get the right answers, are not reflective of the underlying competence or responsive to the underlying construct being assessed. From a different perspective, Rupp, Ferne and Choi (2006) classify the test-taking strategies used on reading tests into general strategies that can be applied to any test format, text-related strategies that test takers employ with the text, and item-related strategies that test takers use with the question items. Some research evidence suggests that the extent to which test takers make use of test-management versus test-wiseness strategies can substantially determine the quality of their test performance (e.g., Wu & Stone, 2015).

Test-taking strategies are generally viewed as compensatory since they are typically used by test takers to make up for some deficiency either in the language ability necessary to perform the test tasks, in the skill to take the test, or in both. In view of that, Cohen and Upton (2006) suggest that test-taking strategies can be explained in terms of the strategic competence framework proposed by Bachman and Palmer (1996, 2010). According to this framework, test respondents go through four metacognitive processes when reacting to a testing task. Test takers first assess the goals of the testing task and determine what aspects of knowledge it draws on (assessment); then, they discern what to do in response to the task (goal setting); after that, they relate the required information in the task to their knowledge and decide about how to act (planning); and finally, they put what they have decided to do into action through the actual provision of the answer (execution).Test performance of two test takers, who might be at the same level of language competency, can be set apart by the extent to which they engage in these processes and manage the use of test-taking strategies on the test (Bachman & Palmer, 1996, 2010).

Reading has been considered by reading experts an interactive and constructive process in which the reader interacts with the text and simultaneously uses a variety of means available to him (e.g., background knowledge and contextual cues) to construct meaning of it (Alexander & Jetton, 2000; Dole, Duffy, Roehler, & Pearson, 1991; Powers & Wilson Leung, 1995). The study of strategy use among readers has offered insights about the overall cognitive and metacognitive processes involved in reading comprehension (Singhal, 2001). Reading strategies represent conscious processes readers utilize to enhance their understanding of a given reading task (e.g., skipping unknown vocabulary) (Birch, 2002). In general, the previous studies of reading strategies among L1 and L2 readers at various levels of proficiency in a wide variety of learning contexts point out the crucial role of such strategies in developing necessary reading skills (cf., Alfassi, 2004; Mokhtari & Sheorey, 2008; Pressley & Afflerbach, 1995). As for gender differences in reading comprehension, a number of studies have found no significant differences between males and females in the use of reading skills and strategies (e.g., Logan & Johnston, 2010; McGeown, Goodwin, Henderson, & Wright, 2012; Phakiti, 2003). Recently, there has been more emphasis on how prior knowledge can maximize strategy use in such a way that it endorses comprehension on reading tasks (e.g., Mokhtari, 2018).

The role of metacognition in reading exhibits itself in how readers are aware of their abilities in relation to the complexity of the reading task, factors pertinent to the difficulty of the reading task (e.g., text familiarity), their repertoire of reading strategies, how the selected strategies are to be used, and the extent to which comprehension is going on (Baker, 2008; Baumann, Jones, & Seifert-Kessel, 1993; Pressley & Gaskins, 2006). Metacognitive awareness and control of reading strategies are essential to achieve an optimum level of reading (Mokhtari, Sheorey, & Reichard, 2008).The relationship between metacognition and comprehension can be described as "reciprocal causation" such that "improvements in metacognition contribute to improvements in comprehension, which, in turn, contribute to further improvements in metacognition" (Baker, 2008, p. 34). In this sense, reading becomes more of

209

"constructively responsive", in which process readers adapt their choice and use of strategies to the demands of the textual information (Pressley & Afflerbach, 1995, p. 2). These perspectives have their origin in L1 reading research; for example, in a number of studies (e.g., Baker & Brown, 1984; Cain, Oakhill, & Bryant, 2004; Roeschl-Heils, Schneider, & van Kraayenoord, 2003), it has been found that the extent to which readers made use of their metacognitive awareness of their reading abilities and cognitive strategies determined the effectiveness of their reading performance. In studies of L2 readers, metacognitive awareness of reading strategy choice and use has shown to be a sign of developing the necessary level of comprehension and so was a characteristic of skilled readers (e.g., Assiri, 2014; Barnett, 1988; Sheorey & Mokhtari, 2001; Phakiti, 2003; Yang, 2002). This major finding showed to be the case with various types of texts: narrative, expository, and graphic (Matheson, 2018).

It is important here to distinguish between test-taking strategies and reading strategies, as these two categories of strategies show some overlap and so can easily be confused in the context of reading assessment. First, test-taking strategies are not specific to any language skill, although it is true that each language skill has some specific test-taking strategies. Second, while reading strategies are generally used when readers engage in a reading activity and thus "are related to text comprehension" (Singhal, 2001, p. 1), test-taking strategies are only used when dealing with a test or assessment task; that is, they are more "driven by the test questions" (Farr, Prichard, & Smitten, 1990, p. 218). Because test respondents approach a given reading-test task with the aim of getting it right, they would certainly turn to strategies that involve reading. However, this does not imply that all strategies used on reading tests necessarily relate to the reading process (Allan, 1992) as a considerable number of these strategies are either test-management or test-wiseness strategies or a combination of both. For example, in contrast to non-testing situations, a multiple-choice reading task calls for "a continual, conscious, and linear engagement in problem-solving activities" (Rupp et al., 2006). In practice, Cohen and Upton (2006) observed that their sample of 32 test takers made far more frequent use of test-taking strategies in comparison to reading strategies. This is not to downplay the importance of reading strategies to perform well on testing tasks that involve summarization, for instance (e.g., Szűcs & Kövér, 2016).

2.2 Test-Taking Strategies and Test Format

It is intriguing to know what test-taking strategies might work with one test format, but not with another. One issue that has triggered extensive research regarding strategy use on reading tests centers on the extent to which questions on standardized tests are answerable without the texts. For example, Powers and Wilson Leung (1995) had a group of L1 readers answer three sets of reading comprehension questions without the passages, and at the same time mark on a checklist the strategies they were using. It was found that although respondents had difficulty with some questions that were more passage dependent, they mostly employed their ability of verbal reasoning through using the questions as building blocks to develop a mental schema of the text.

Nevo (1989) describes the format aspects of a multiple-choice reading test that can influence strategy use significantly, including the level of text familiarity and task complexity. Research on the effect of

210

task difficulty on test performance has focused on strategy use to determine which tasks and question items are especially challenging to test takers. For example, Cohen and Upton (2006) observed that the choice of language to report the use of certain test-taking strategies among their respondents, while performing the TOEFL-iBT reading tasks, reflected the level of difficulty of the items with which these strategies were associated (e.g., wrestling with the question intent). Assiri (2011) investigated test-taking strategy use on the reading section of the TOEFL iBT. The researcher used a procedural integration of stimulated recall, self-observation, and retrospective interview (cf., Assiri, 2016) with 25 Arab ESL learners. Test takers were found to use strategies depending on item format and difficulty in ways that allowed them to achieve different goals, adapt strategies to various task items, and apply strategies using several textual and technical means. Using eye-tracking technology in his study of the relationship between task complexity and task-based reading performance, Jung (2017) found that the complex texts called for test takers' use of reading strategies indicative of highly careful reading and intensively textual processing. In another study, it was found that test takers adapted their strategy use to the increasing task difficulty in a manner that capitalized on comprehension and/or test-wiseness (Wu, Chen, & Stone, 2017).

In other studies using multiple-choice reading tasks, there was a notable tendency among respondents, specifically those with low proficiency, to guess the key answers from the options without referring to the text, match the content of the item stem and options with that of the passage (e.g., Rupp et al., 2006), or eliminate what they perceived not to be the key answers among the options (e.g., Storey, 1997). Farr et al. (1990) found that test takers used the multiple-choice questions to guide their dealing with the text in order to identify the key answers. In another study, Rupp et al. (2006) examined response strategies used by a group of 10 ESL readers on three multiple-choice reading tests for academic purposes. The researchers observed that respondents moved from identifying the theme to locating specific details to answer the questions, used the questions to guide their scanning of the text and locating the key information, used the order of the questions to identify the location of the respective key information in the text, and made use of rational elimination based on prior knowledge or clued-up guessing. All things considered, Rupp and his colleagues concluded that [multiple-choice] questions might function well as separable measures of how difficult different aspects of texts are for test-takers or of how well test-takers engage in lower-order component processes rather than as composite measures of higher-order reading comprehension, which they may be sometimes colloquially assumed to be (p. 468).

Rafi and Islam (2017) identified seven categories of test-taking strategies as option-selecting strategies, question-rereading strategies, option comprehension strategies, answer-checking strategies, option consideration strategies, cognitive strategies, and clue-finding strategies (p. 46). Interestingly, test takers may use the same test-taking strategies on multiple-choice tests regardless of whether or not the texts on such tests are familiar to them (Lee, 2015).

In studies where L2 learners were asked to respond to open-ended questions on reading passages, some

respondents were observed to find where the answer was most likely to be in the passage and write the whole sentence or context containing the answer in response to the given questions (e.g., Cohen & Aphek, 1979). In other studies where cloze tasks were used, low proficient students were shown to use more of micro-level processing when a half of each deletion was given to them by trying to guess the deleted word using the remnant of it or other local hints rather than a more global understanding of the context surrounding the word or using macro-level processing (e.g., Stemmer, 1991). Text authenticity has also shown to be a determinant of strategy use. For example, Abanomey (2002) found that text authenticity did not affect how many test-taking strategies were used by respondents as much as how these strategies were used. That is, while the authentic texts invoked the use of bottom-up strategies, the inauthentic texts called upon top-down strategies. The researcher ascribed this observation to the fact that inauthentic texts do not possess the kind of textual features which draw on bottom-up strategies as do authentic materials.

2.3 Test-Taking Strategies and Test Validation

Another intriguing facet to the study of strategy use on reading tests has to do with how it can inform efforts put into test validation. Test validation rests on the question as to whether examinees' response behaviors on a given test conform to the test purposes and its intended uses (cf., Cohen, 2006). Bachman (1991) maintains that our study of test-taking strategies for the purpose of test validation of language tests provides us with a lens into test performance, and so, reflects the extent to which our test tasks are similar to real-world uses of language. Test validation can be performed during the pilot phase of test development by having a sample from the target population of examinees take the test and observe their test-taking strategies (Cohen & Upton, 2006).

As noted in the previous section, the kinds of test-taking strategies to use on a given test depend on the task types and question items on the test. Accordingly, an incorrect answer to a question item on a test could point out that either the respondent himself failed to answer correctly or the test format influenced the respondent toward the provision of the incorrect answer (Cohen, 1998). Grotjahn (1987), Klein-Braley (1985), and Klein-Braley and Raatz (1984) were among the early efforts to use test-taking strategies as the basis for validating and refining language tests, specifically cloze tasks. The outcomes of such efforts have led to the development of cloze tests that use rational deletion and offer adequate sampling of the language components to be assessed. Anderson, Bachman, Perkins and Cohen (1991) attempted to validate a TOEFL reading test in a multiple-choice format by focusing on the relationship between strategy use and the item types and performances on the test. The researchers found that the question type determined the choice and use of the test-taking strategies to tackle it; for example, wherever respondents were asked to make inferences, some chose to relate and match the content of the question to that of the text. And, wherever a small number of strategies were used, the question items with which such strategies were used were shown to be too easy, too difficult, or less discriminable.

Judgment about the validity of tests does not only rely on the behavior of the individual question items on these tests, but also on the overall format of the tests. In this regard, Tsagari (1994) investigated how

212

the free response format for assessing reading compares with the multiple-choice format, with a group of ESL learners. On the basis of the examinees' strategy use, Tsagari concluded that the two formats measure reading comprehension differently in that each format seems to tap into distinct reading abilities. In a study which aimed to validate a reading test with 13 adult learners of French, Wijgh (1995) observed that test takers' strategy use did not match the test constructors' intentions for the test questions, which led him to conclude that either the test takers were not skilled enough to use appropriate strategies or the question items themselves failed to call for adequate strategy use. With the aim of determining what a cloze task truly measures, Storey (1997) had a group of Chinese women answer multiple-choice, discourse cloze tests using rational deletion. The researcher noticed that on tests where the deletions involved discourse markers, respondents were prone to detect the line of argumentation and employ the rhetorical organization to supply the deletions. However, where the deletions were cohesive devices, respondents could just rely on local clues to figure out the deletions. After all, the researcher concluded that discourse cloze has the capacity to call upon processing strategies used at both local and global levels of the text.

In two of three more recent studies, Lumley and Brown (2004, 2006) looked into the validity of the integrated reading and writing tasks on the TOEFL iBT with 60 respondents from three language backgrounds. On the basis of strategy reports collected from the respondents, the researchers were able to identify serious flaws with these tasks in terms of the difficulty of deciding among raters about whether the responses to the writing tasks were in the participants' words or language they copied from the reading texts. In their study, Cohen and Upton (2006) sought to investigate the extent to which the reading section of the TOEFL iBT truly assesses the academic reading skills prospective ESL students need to have at the university level. As a result, the researchers noted that test takers dealt with the whole section as demanding of masterful test-taking strategies, and that neither *inferencing* nor the *reading to learn* task types required reading skills distinct from each other. Even so, they concluded that the reading section of the TOEFL iBT adequately measures academic reading skills required at the university level. In a more recent study, Dawadi and Shrestha (2018) examined the construct validity of an English reading test used as part of the Nepalese school leaving examination. The researchers found huge discrepancies between test takers' self-reported strategies and eight experts' judgments of the skills measured by the test.

One approach to test validation seeks to ensure that test takers have to rely on the skill or knowledge represented by the test construct to answer the test more than on their test-wiseness. For example, P. Yang (2000) set out to examine the extent to which test-wiseness impacts performance on the computer-based TOEFL. First, Yang had his respondents answer an adaptation of Rogers and Bateson's (1991) Test-wiseness Test and a TOEFL practice test. Based on their scores on the test-wiseness test, two groups were identified—one as test-wise and the other as test-naïve. Then, respondents were asked to report their strategy use with selected items from both the test-wiseness test and the TOEFL practice test. It was found that test-wiseness could help with at least half of the items from the listening and

reading sections of the test and that test-wise examinees could follow systematic ways in tackling those question items amenable to test-wiseness. In this respect, Cohen (2006) suggests that test makers should strive to ensure that their tests are not susceptible to test-wiseness if these tests are to be optimally challenging for examinees. In general, test-wiseness strategies fit in the description that Powers and Wilson Leung (1995) offer when stating that "[s]trategies that raise test scores but bear little if any relationship to what the test was designed to measure may diminish the predictive power of a test or dilute the meaning of scores derived from it" (p. 105). This is what also motivated Yamashita (2003) to recommend that test takers' perceptions and opinions about question items be taken into account since test takers are expected to reveal the kinds of strategies that may work for them with given question items, without these strategies being necessarily reflective of the trait or skill being assessed.

2.4 L2 Proficiency and Test-Taking Strategy Use

Based on an extensive review of research on reading strategies, Singhal (2001) suggests that the use of reading strategies is strongly associated with the level of proficiency such that highly proficient readers use a variety of strategies more frequently and effectively when compared to less proficient readers. Similarly, Tian (2000) points to the major findings from comparative studies of reading strategy use among varied levels of proficiency, which can be summarized as follows: while proficient readers work towards forming a global understanding of what they read using higher-order processing skills, less proficient readers work on a more local level using lower-order processing skills; proficient readers show high flexibility in their strategy use when less proficient readers tend to be more rigid in this respect; and highly proficient readers utilize more active and ongoing monitoring while their less proficient counterparts fail to execute adequate level of monitoring and so are less able to evaluate and fine-tune their strategy use.

Nevo (1989) was among the first scholars to make the assumption that because reading-test taking represents a problem-solving situation, high-proficiency test takers employ strategies that increase their chance of getting a question item correct to a greater extent in comparison to their low-proficiency counterparts. She based this assumption on a note she made earlier in her article, suggesting that readers' ability to deal with problem-solving situations in language use correlates with their levels of language proficiency, specifically knowledge of grammar and vocabulary. Also, Purpura (1999) highlights the role of language ability in the efficient use of metacognitive strategies on reading tests. In fact, the two discrete levels of proficiency exhibit different models of cognitive strategy use, with that of the high-proficiency group being specially complex and quite indicative of advanced levels of memory activation and information retrieval (Purpura, 2013). Besides task difficulty, Phakiti (2003) counts proficiency level as another factor that affects cognitive and metacognitive strategy use; for example, Phakiti (2003) noticed among his respondents that although both high- and low-proficiency test takers exhibited response behaviors so automatic to the extent that they were not aware if they had used metacognitive strategies like *checking* and *monitoring*, the low-proficiency test takers employed

214

these two strategies to a lesser degree.

Early research on the differential use of strategies on reading tests pointed out that the main difference between L1 and L2 readers lies in L1 readers' superior ability to make use of intra- and inter- sentential, and semantic clues, as measured by means of oral miscue and cloze tasks (e.g., Cziko, 1978, 1980; Douglas, 1981; Hauptman, 1979). At an EFL setting, Mangubhai (1990) used cloze reading procedures along with think-aloud protocols to investigate strategy use among young learners at three different levels of proficiency. The subjects were six learners in year eleven—two High, two Middle, and two Low achievers according to their scores on a national EFL examination. The subjects demonstrated differences in their strategy use between the high and the low levels of proficiency in that the high achievers used such strategies as *look at larger context after generating the word, refer to prior knowledge, rephrase the sentence in order to generate the word, evaluate guesses for their correctness,* and *analyze the passage using prior and contextual knowledge in order to generate the word*, whereas the low achievers used such strategies as *look at the immediate context and generate randomly and/or reject words on syntactic or semantic grounds* (p. 133). Strong positive correlations were observed between the respondents' levels of proficiency, the total percentages of effective strategies they used, and their scores.

In another study combining open-ended and multiple-choice reading tasks, Gordon (1987) used think-aloud protocols to look into response behaviors among 30 tenth-grade EFL learners. The findings from this study suggested that the respondents who were at a low level of proficiency focused on isolated or fragmented elements of the text where highly proficient respondents related the meaning of intact, individual sentences to the whole text. In relation to this, while the low-proficiency students used strategies such as *matching words in the options to words in the text, copying information from the text,* and *translating word for word*, the high-proficiency students used strategies such as *predicting information* and *making inferences* (as cited in Cohen, 1998, p. 100).

It is very often the case that in reading test-like situations such as cloze tasks, those who are highly proficient exhibit skillfulness in using text-level comprehension to guide their completion of the cloze items (Bachman, 1985), or at least use other problem-solving strategies such as rational guessing when faced with challenging deletions (Cohen, 1984). On the other hand, those test takers who are less proficient tend to make heavy use of translation and very localized clues to solve deletions (Cohen, 1984). It is worth noting here that the tendency to employ rational guessing was observed to be a characteristic of learners' reading in their L1 in contrast to their reading of L2 texts, in which case they were observed to make more use of random guessing (Nevo, 1989). In her study, Nevo (1989) sought to examine the use of test-taking strategies on a multiple-choice test of reading among 42 Hebrew tenth graders studying French. The subjects responded to a multiple-choice reading test administered first in Hebrew and then in French. While responding to the test items, the respondents were to mark each strategy they used on a strategy checklist on an item-by-item basis. The researcher found that the respondents' higher proficiency in L1 when compared to L2 enabled them to use effective strategies or

strategies that often led them to provide correct answers more on the test in Hebrew (L1) than on the one in French (L2).

Dollerup, Glahn, and Rosenberg-Hansen (1982) propose a distinction between two modes of taking standardized reading tests as "mainline" versus "fragmented". According to this view, mainline reading involves the test taker's skimming of the text to develop the main idea and then answer the question items on this basis, whereas fragmented reading is characterized by switching between the text and the question items (p. 96). Therefore, on reading tests, while highly proficient test takers are expected to be mainline readers, less proficient ones follow more of a fragmented way of taking the test. This may explain why proficient readers have been shown to use fewer strategies on standardized reading tests than those who are less proficient. An example comes from a study in which Yamashita (2003) had 12 Japanese EFL students, at the university level, complete a cloze test using rational deletion and at the same time verbalize their thoughts. The results showed that the highly proficient readers handled the deletions one after another drawing on information at both the textual and the clausal levels, whereas their low proficient counterparts switched back and forth among deletions and used mostly clause-level information.

In an attempt to investigate how each of strategy use, proficiency level, and level of language aptitude relates to one another, Yoshizawa (2002) had a group of 54 Japanese adult ESL learners. The participants were instructed to report the kinds of text-processing strategies they normally use when performing L2 listening and reading tasks by completing a questionnaire. A language aptitude battery was used to measure their foreign language aptitude and a test from the TOEFL Institutional Testing Program was used to assess their English proficiency. The researcher found that the participants exhibited progressively more effective strategy use across their proficiency levels from low to high.

In Cohen and Upton (2006), the researchers were interested in the kinds of reading and test-taking strategies that test takers tend to use on the TOEFL-iBT reading section. The participants were 32 mostly graduate students, representing seven language backgrounds. The data on the test takers' strategy use were collected via think-alouds in which each test taker verbalized his or her thoughts while responding to reading tests containing two 600-700 word readings—each with 12-13 question items. Cohen and Upton made the observation that their respondents used predominantly more test-management strategies than test-wiseness strategies, which the researchers linked to the high proficiency of the test takers. On this basis, one can deduce that test takers who are at a low level of proficiency would be more likely to use more test-wiseness than test-management strategies. Assiri's (2011) study revealed that high-test performance and scoring were characterized by superior skills of comprehension and test-management as well as high levels of strategic awareness and monitoring. Conversely, low test performance and scoring were associated with poor skills of comprehension and excessive use of test-wiseness.

2.5 Test-Taking Strategies and Test Performance

In their model of communicative language ability, Bachman and Palmer (1996, 2010) clearly

demonstrate how strategy use, prompted by strategic competence as a mediating component between competence and performance, influences how competence contributes to performance. Whether or not strategy use can make a difference in scores on language tests used to be an issue calling for research into how the use of test-taking strategies shapes test performance. Meanwhile, test constructors relentlessly expressed their refutation to any claims about the possibility of gaining high scores by means of test-taking strategies (Tian, 2000). Unfortunately, the distinction between the two types of test-taking strategies (i.e., test management and test wiseness) was not adequately addressed in the debate.

Presumably, the test-taking process is similar to a problem-solving situation, in which case the use of problem-solving strategies is natural, and so reflects the authenticity of the test. Problem-solving strategies on language tests can either be test-management strategies or test-wiseness strategies; hence, while the use of the former is a sign of a skilled response behavior, the latter can be indicative of a poor response behavior, an invalid test item, or both. On the other hand, while test-management strategies can be mastered through test preparation practices, test-wiseness strategies are more linked to problem-solving abilities of test takers, and so do not easily lend themselves to training or instruction. To the contrary of using test-management strategies, as Allan (1995) suggests, the use of test-wiseness strategies is idiosyncratic and so results in unfair testing and undeserving achievement, assuming that the given test is amenable to test-wiseness.

The issue of how strategy use relates to test performance depends to a large extent on the nature of the test format and tasks. For example, it has been observed across a number of studies of response behaviors on cloze tasks that test takers could still manage to obtain high scores and never had to read the whole text or even understand its main idea (Cohen, 1984). Test takers in other studies of performance on cloze tasks have been observed to first use the local clues to solve as many deletions as they could, and then they moved onto forming a general idea about the text and used more global clues in order to tackle the unsolved deletions (e.g., Kleiman, Cavalcanti, Terzi, & Ratto, 1986). Yet, in another study of strategy use on cloze tasks, Homburg and Spaan (1982) reported that their respondents' use of such strategies as identifying parallel elements, discourse chunking, cataphoric reading, and anaphoric reading correlated with their success in identifying the correct completions and that those who made effective use of cataphoric reading were better able to figure out the main idea of the text than those who did not.

Cohen (1984) also refers to other studies in which test takers were asked to respond to multiple-choice questions in the absence of the reading passages and how these test takers could still manage to score high above the chance level (i.e., 25% with question items with four alternatives). In his study of strategy use on summarization tasks, Cohen (1994) found that the time some test takers spent going through and applying strategic processing far exceeded the time they spent writing their summaries and that they oftentimes chose to add whole blocks of the text being summarized to their summaries. There is also considerable evidence regarding positive transfer of strategies from L1 to L2 as far as

217

performance on reading tests is concerned. In this regard, Nevo (1989) identified two strategies as the most frequently used ones in both L1 (Hebrew) and L2 (French) among her respondents, including reading the questions first and then looking for the key information in the text and matching clues from the question items to those in the text to locate the key information.

As has been confirmed across studies of strategy use on language tests, test-taking strategies are by and large a function of the test format and situation; in other words, "[a]s long as the task is part of a test, students are bound to use strategies they would not use under non-testing conditions" (Cohen, 1992, p. 99). Bachman and Palmer's (1996, 2010) view of the mediating role of strategic competence between language knowledge and language use is obviously warranted in the context of language testing when we consider how test-taking strategies can either facilitate or debilitate performance on language tests. For example, a test taker who chooses to read the questions first on a reading standardized test has a higher chance of completing the test more quickly and efficiently than one who chooses to read the text first and then proceeds to the questions, assuming that both test takers share the same proficiency level (Cohen, 1992). Clearly, the use of the first strategy can be said to facilitate test performance, whereas the other strategy can considerably debilitate performance under conditions of timed testing and/or deficient competence.

Generally, within the broad field of strategic competence, successful use of strategies demands that strategies be relevant to the nature of the task at hand, strategies be in sync with learner characteristics, and a learner be aware if a strategy is to be used by itself or combined with other strategies and how either form ought to be used (Anderson, 2005). Even those facilitative strategies can be more or less effective depending on when and how they are used (Cohen, 1992). Therefore, what Purpura (1998) concluded with regard to how the use of cognitive and metacognitive strategies influences test performance applies to the case of using test-taking strategies; that is, the nature of the test task on which a given strategy is used as well as how this strategy is used both determine the extent to which this strategy can benefit test performance. Along the same lines, Nikolov (2006) noted that effective use of test-taking strategies counts on the degree of compatibility among strategies, the test task being performed, and the commands of trends that test takers have developed with the use of these strategies.

Anderson (1991) also points out that the manner in which strategies are used in isolation or in conjunction with other strategies determines their beneficial effect on test performance and scores. He based this conclusion on a study in which, Anderson (1991) looked at individual differences in the use of reading and test-taking strategies among 28 adult ESL learners at three different levels of proficiency. Each respondent was asked to take a standardized test of academic reading, in a multiple-choice format, and simultaneously think aloud his or her response behaviors, either in English or their L1. Based on case studies of three individual respondents, the researcher found that the high and low scorers did not differ from each other in the kinds of strategies they used, but rather in how effectively they used these strategies individually or in conjunction with other strategies, as well as the ability to assess and monitor strategy use.

TOEFL test takers are commonly encouraged to employ certain test-taking strategies which have been shown to be useful. The strategies that can be used on the reading section of the test include familiarizing oneself with the test directions before taking it, reading cursorily and taking mental notes of the closest answers, proportioning time allotted among the number of question items, using any time left to mark the closest answers to question items whose key answers are not known for certain, and marking C or D options if guessing is not promising (Forster, Karn, Suzuki, & Tateyama, 1997, p. 90). Forster et al. (1997) also highlight other test-taking strategies when advising test takers to read the questions first and then the passage, postpone answering questions about the main idea or the title of the passage until they have answered the other questions, use elimination of alternatives with questions about excluded facts in order to best discern the key answer, rule out obscure and irrational alternatives, and consider more likely key answers those alternatives that are phrased synonymously with or using the same part of speech as that in the key information in the passage (pp. 120-136). While Forster et al. (1997) refer specifically to the paper-based format of the TOEFL in giving this account of strategies, their advice is also applicable to the computer- and the internet-based formats of the test. It is obviously the case that on standardized language tests, test takers ought to be familiar with the kinds of test-taking strategies that, as Yien (2001) suggests, can truly mediate between test takers' characteristics including proficiency and their performance on the test; otherwise, effective strategies would not be characterized as such.

In her study of strategy use on the reading section of the paper-based TOEFL, Tian (2000) worked with a sample of 43 Taiwanese students attending a coaching school. The participants were first asked to take a TOEFL reading practice test and at the same time think aloud their thought processes. Then, the participants were engaged in a recall task in which they had to write down whatever they could recall from their reading of the passage on the test. After that, the participants were interviewed as to how they went about preparing for the TOEFL and what they thought of the coaching school and the kind of training it offered. A taxonomy of the strategies used was developed, incorporating 42 strategies categorized as technical strategies, reasoning strategies, and self-adjustment strategies. The examinees were divided into three performance levels: high, middle, and low, as determined by their scores on the test. In general, the three performance groups reported in the interviews that they often resorted to certain test-management strategies, for example, starting with the question items first, answering the main idea question after going through the other items, and using word-based strategies and syntactic or semantic clues when confronted with challenging question items. The results of the interviews also revealed one major difference among the three performance levels, namely the awareness of how to use test-taking strategies effectively in terms of what strategies to use, when to use them, and how to use them. Such strategic awareness of effective strategy use increases with proficiency level, and so augments test performance, as pointed out earlier in other studies (e.g., Anderson, 1991).

Similarly, Nikolov (2006) found that the low scorers among her ESL respondents were more disposed than were high scorers to choose words they were not certain about to complete multiple-matching

219

tasks. This was found through an exploratory study in which Nikolov looked into test-taking strategy use, among 12- and 13-year-old EFL learners, on reading and writing tasks. A total of 52 participants were randomly sampled and divided into groups that represent three levels of proficiency. The data were collected using non-mediated think-aloud protocols on an individual basis as the respondent was taking a language test comprised of reading and writing tasks. Four case studies were carried out with two top scorers and two bottom scorers, selected on the basis of their scores on the language test. Overall, although the high performers did not show the use of quite as many test-taking strategies, they made more effective use of the strategies they used, when compared to those with low scores. Peculiarly, Assiri (2011) found that test takers sequenced strategies such that certain strategies derived from other strategies, endorsed or facilitated functions of other strategies, or acted in sync with other strategies. Such aspects of effective strategy use among test takers increased their chance of getting items right on the TOEFL-iBT reading tasks.

3. Conclusions

This paper has presented highlights of what scholarly thought and research have come to know and find in regard to test-taking strategy use on reading comprehension tests. The discussion throughout the paper explicates the major themes revolving around this topic, which can be summarized as follows.

Similar to language use strategies, test-taking strategies are compensatory in nature; however, whereas the former are used for communication goals, the latter serve test-taking purposes. As manifested throughout the discussion of test-taking strategy use on a variety of reading test formats, even highly proficient test takers cannot spare the use of strategies at all times, especially those strategies that can make a difference in test scores. The process of reading test-taking may call upon the use of both test-taking strategies as well as reading strategies; the former are used with the question items whereas the latter are used with the text. In this respect, the use of test-taking strategies far exceeds that of reading strategies since test takers' care for answering the question items surpasses their care to fully understand the text. It comes as no surprise that most of the comprehension that test takers attain when responding to a standardized reading test is brought about by their responses to the test questions and not an overall understanding based on the reading of the text.

Research has pointed out that formats or tasks used on reading tests play a crucial role in determining examinees' strategy choice and use. Therefore, those strategies that are typically used on cloze formats are different from those used on multiple-choice formats. Examples of the strategies used on the former include *look backward and forward to figure out the missing word* and *look at larger context after generating the word*, where those used on the latter include *read the questions then read the text to locate the critical information* and *select an option by eliminating other options*. For the purposes of validating language tests, research has confirmed that the study of test-taking strategy use on language tests can help us make sure that our tests measure what they are intended to measure, determine how various formats tap into different abilities underlying the main skill we seek to assess, and ensure that

220

our tests are not susceptible to test-wiseness.

There are varying degrees of test-taking strategy choice and use associated with varied levels of proficiency. The previous studies of how test-taking strategy use relates to proficiency have revealed that high- and low-proficient test takers approach reading tests in different ways and exhibit varied response behaviors reflective of their language competencies. The distinction between the two categories of test-taking strategies as test-management versus test-wiseness strategies is pertinent to the differential strategy use of two discrete levels of proficiency. To be specific, high-proficiency test takers make more use of test-management strategies than their low-proficiency counterparts whose choices are more limited to test-wiseness strategies. The choice and use of test-taking strategies on reading tests can either facilitate or debilitate test performance depending on whether or not strategies themselves are compatible with the test format, the level of knowledge and awareness of when and how to use a given strategy, and the level of skillfulness in using a given strategy in isolation or in association with other strategies. Finally yet importantly, adequate instruction of test-taking strategies is highly recommended in order to ensure that learners are well prepared for various language tests (cf., Cheng & Fox, 2017).

Acknowledgements

References

Abanomey, A. A. (2002). *The effect of texts' authenticity on reading-comprehension test-taking strategies used by adult Saudi learners of English as a foreign language* (Doctoral dissertation). ProQuest Dissertation and Theses database.

Alexander, P. A., & Jetton, T. L. (2000). Learning from text: A multidimensional and developmental perspective. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research: Vol. III* (pp. 285-310). Mahwah, NJ: Lawrence Erlbaum Associates.

Alfassi, M. (2004). Reading to learn: Effects of combined strategy instruction on high school students. *The Journal of Educational Research*, *97*(4), 171-184. https://doi.org/10.3200/JOER.97.4.171-185

Allan, A. (1992). Development and validation of a scale to measure test-wiseness in EFL/ESL reading test takers. *Language Testing*, *9*(2), 101-122. https://doi.org/10.1177/026553229200900201

Allan, A. (1995). Begging the questionnaire: Instrument effect on readers' responses to a self-report checklist. *Language Testing*, *12*, 133-153. https://doi.org/10.1177/026553229501200201

Anderson, N. J. (1991). Individual differences in strategy use in second language reading and testing. *Modern Language Journal*, *75*(4), 460-472. https://doi.org/10.1111/j.1540-4781.1991.tb05384.x

Anderson, N. J. (2005). L2 learning strategies. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 757-771). Mahwah, NJ: Lawrence Erlbaum.

Anderson, N. J., Bachman, L., Perkins, K., & Cohen, A. (1991). An exploratory study into the construct validity of a reading comprehension test: Triangulation of data sources. *Language Testing*, *8*(1), 41-66. https://doi.org/10.1177/026553229100800104

Assiri, M. S. (2011). *Test-taking strategy use on the reading section of the TOEFL iBT: A study of Arab ESL learners* (Doctoral dissertation). ProQuest dissertation and theses database.

Assiri, M. S. (2014). Metacognitive and cognitive strategy use and performance on a reading test with multiple-format tasks. *Arab World English Journal*, *5*(4), 187-202.

Assiri, M. S. (2016). Integration of stimulated recall, self-observation, and retrospective interview in the collection of strategy data in computer-assisted language testing. *Studies in English Language Teaching*, *4*(1), 104-122. https://doi.org/10.22158/selt.v4n1p104

Bachman, L. F. (1985). Performance on cloze tests with fixed-ratio and rational deletions. *TESOL Quarterly*, *19*(3), 535-556. https://doi.org/10.2307/3586277

Bachman, L. F. (1991). What does language testing have to offer? *TESOL Quarterly*, *25*, 671-704. https://doi.org/10.2307/3587082

Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Oxford, England: Oxford University Press.

Bachman, L. F., & Palmer, A. S. (2010). *Language testing in practice*. Oxford, England: Oxford University Press.

Baker, L. (2008). Metacognitive development in reading: Contributors and consequences. In K. Mokhtari, & R. Sheorey (Eds.), *Reading strategies of first- and second-language learners: See how they read* (pp. 25-42). Norwood, MA: Christopher-Gordon.

Baker, L., & Brown, A. (1984). Metacognitive skills and reading. In D. Pearson (Ed.), *Handbook of reading research* (pp. 353-394). New York: Longman.

Barnett, M. (1988). Reading through context: How real and perceived strategy use affects L2 Comprehension. *Modern Language Journal*, *72*, 150-162. https://doi.org/10.1111/j.1540-4781.1988.tb04177.x

Baumann, J. F., Jones, L. A., & Seifert-Kessell, N. (1993). Using think alouds to enhance children's comprehension monitoring abilities. *The Reading Teacher*, *47*(3), 184-193.

Birch, B. (2002). *English L2 reading: Getting to the bottom*. Mahwah, NJ: Erlbaum.

Cain, K., Oakhill, J., & Bryant, P. E. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology*, *96*, 31-42. https://doi.org/10.1037/0022-0663.96.1.31

Cheng, L., & Fox, J. (2017). *Assessment in the language classroom: Teachers supporting student learning*. London: Palgrave.

Cohen, A. D. (1984). On taking language tests: What the students report. *Language Testing*, *1*(1), 70-81. https://doi.org/10.1177/026553228400100106

Cohen, A. D. (1986). Mentalistic measures in reading strategy research: Some recent findings. *The ESP Journal*, *5*(2), 131-145. https://doi.org/10.1016/0889-4906(86)90019-0

Cohen, A. D. (1992). Test-taking strategies on language tests. *Journal of English and Foreign Languages*, *10-11*, 90-105.

Cohen, A. D. (1994). *Assessing language ability in the classroom* (2nd ed.). Boston, MA: Newbury House/Heinle & Heinle.

Cohen, A. D. (1998). Strategies and processes in test taking and SLA. In L. F. Bachman, & A. D. Cohen (Eds.), *Interfaces between second language acquisition and language testing research* (pp. 90-111). Cambridge: Cambridge University Press.

Cohen, A. D. (2006). The coming of age of research on test-taking strategies. *Language Assessment Quarterly*, *3*, 307-331. https://doi.org/10.1080/15434300701333129

Cohen, A. D. (2012). Test-taking strategies. In C. Coombe, P. Davidson, B. O'Sullivan, & S. Stoynoff (Eds.), *The Cambridge guide to second language assessment* (pp. 96-104). Cambridge University Press.

Cohen, A. D. (2014). *Strategies in learning and using a second language*. Routledge, NY: Longman Applied Linguistics.

Cohen, A. D., & Aphek, E. (1979). *Easifying second language learning*. Jerusalem, Israel: The Jacob Hiatt Institute.

Cohen, A. D., & Upton, T. A. (2006). *Strategies in responding to the New TOEFL reading tasks* (TOEFL Monograph Series Report No. 33). Princeton, NJ: Educational Testing Service. Retrieved from http://www.ets.org/Media/Research/pdf/RR-06-06.pdf

Cziko, G. A. (1978). Differences in first- and second-language reading: The use of syntactic, semantic and discourse constraints. *The Canadian Modern Language Review*, *34*, 473-489. https://doi.org/10.3138/cmlr.34.3.473

Cziko, G. A. (1980). Language competence and reading strategies: A Comparison of first- and second-language oral reading errors. *Language Learning*, *30*, 101-116. https://doi.org/10.1111/j.1467-1770.1980.tb00153.x

Dawadi, S., & Shrestha, P. (2018). Construct validity of the Nepalese school leaving English reading test. *Educational Assessment*, *23*(2), 102-120. https://doi.org/10.1080/10627197.2018.1430511

Dole, J. A., Duffy, G. G., Roehler, L. R., & Pearson, P. D. (1991). Moving from the old to the new: Research in reading comprehension instruction. *Review of Educational Research*, *61*, 239-264. https://doi.org/10.3102/00346543061002239

Dollerup, C., Glahn, E., & Hansen, C. (1982). Reading strategies and test-solving techniques in an EFL reading comprehension test: A preliminary report. *Journal of Applied Language Study*, *1*(1), 93-99.

Douglas, D. (1981). An exploratory study of bilingual reading proficiency. In S. Hudelson (Ed.), *Learning to read in different languages* (pp. 33-102). Washington, DC: Center for Applied Linguistics.

Farr, R., Pritchard, R., & Smitten, B. (1990). A description of what happens when an examinee takes a multiple-choice reading comprehension test. *Journal of Educational Measurement*, *27*(3), 209-226. https://doi.org/10.1111/j.1745-3984.1990.tb00744.x

Forster, D., Karn, R., Suzuki, S., & Tateyama, T. (1997). *Strategies 1: Building TOEIC/TOEFL test-taking skills*. Tokyo: Aratake Publishing Co., Ltd.

Gordon, C. (1987). *The effect of testing method on achievement in reading comprehension tests in English as a foreign language* (Master thesis). Tel Aviv, Israel: Tel Aviv University.

Grotjahn, R. (1987). How to construct and evaluate a C-Test: A discussion of some problems and some statistical analyses. In R. Grotjahn, C. Klein-Braley, & D. K. Stevenson (Eds.), *Taking their measure: The validity and validation of language tests* (pp. 219-254). Bochum, Germany: Brockmeyer.

Hauptman, P. C. (1979). A comparison of first and second language reading strategies among English-speaking university students. *Interlanguage Studies Bulletin*, *4*, 173-201.

Homburg, T. J., & Spaan, M. C. (1982). ESL reading proficiency assessment: Testing strategies. In M. Hines, & W. Rutherford (Eds.), *On TESOL '81* (pp. 25-33). TESOL, Washington.

Jung, J. (2017). *Effects of task complexity, glossing and working memory on L2 reading and L2 learning*. Doctoral thesis, UCL (University College London).

Kleiman, A. B., Cavalcanti, M. C., Terzi, S. B., & Ratto, I. (1986). *Perception of the lexicon and its discourse function: Some conditioning factors*. Universidade Estadual de Campinas, Campinas, Brazil.

Klein-Braley, C. (1985). A close-up on the C-Test: A study in the construct validation of authentic tests. *Language Testing*, *2*, 76-104. https://doi.org/10.1177/026553228500200108

Klein-Braley, C., & Raatz, U. (1984). A survey of research on the C-Test. *Language Testing*, *1*, 134-146. https://doi.org/10.1177/026553228400100202

Lee, Jia-Ying. (2015). Language learner strategy by Chinese-speaking EFL readers when comprehending familiar and unfamiliar texts. *Reading in a Foreign Language*, *27*(1), 71-95.

Logan, S., & Johnston, R. (2010). Investigating gender differences in reading. *Educational Review*, *62*(2), 175-187. https://doi.org/10.1080/00131911003637006

Lumley, T., & Brown, A. (2004). Test-taker and rater perspectives on integrated reading and writing tasks in the Next Generation TOEFL. *Language Testing Update*, *35*, 75-79.

Lumley, T., & Brown, A. (2006). *Test taker response to integrated reading/writing tasks in TOEFL: Evidence from writers, texts, and raters* (Final Report to ETS). Melbourne: Language Testing Research Centre, University of Melbourne.

Mangubhai, F. (1990). Towards a taxonomy of strategies used by ESL readers of varying proficiencies while doing cloze exercises. *Australian Journal of Reading*, *13*(2), 128-139.

Mangubhai, F. (1991). The processing behaviors of adult second language learners and their relationship to second language proficiency. *Applied Linguistics*, *12*, 268-298. https://doi.org/10.1093/applin/12.3.268

Matheson, I. (2018). *Unpacking reading comprehension by text type: An examination of strategy use and cognitive functioning in poor and typically-achieving comprehenders* (Doctoral dissertation). ProQuest dissertation and theses database.

McGeown, S., Goodwin, H., Henderson, N., & Wright, P. (2012). Gender differences in reading motivation: Does sex or gender identity provide a better account? *Journal of Research in Reading*, *35*(3), 328-336. https://doi.org/10.1111/j.1467-9817.2010.01481.x

Mokhtari, K. (2018). Prior knowledge fuels the deployment of reading comprehension strategies. In J. I. Liontas (Ed.), *The TESOL Encyclopedia of English Language Teaching* (pp. 1-10). Wiley-Blackwell.

Mokhtari, K., & Sheorey, R. (2008). *Reading strategies of first- and second-language learners: See how they read*. Norwood, MA: Christopher-Gordon.

Mokhtari, K., Sheorey, R., & Reichard, C. (2008). Measuring the reading strategies of first- and second-language readers. In K. Mokhtari, & R. Sheorey (Eds.), *Reading strategies of first- and second-language learners: See how they read* (pp. 43-65). Norwood, MA: Christopher-Gordon.

Nevo, N. (1989). Test-taking strategies on a multiple-choice test of reading comprehension. *Language Testing*, *6*(2), 199-215. https://doi.org/10.1177/026553228900600206

Nikolov, M. (2006). Test-taking strategies of 12-13-year-old Hungarian learners of EFL: Why whales have migraine. *Language Learning*, *57*(1), 1-51. https://doi.org/10.1111/j.0023-8333.2006.00341.x

Phakiti, A. (2003). A closer look at the relationship of cognitive and metacognitive strategy use to EFL reading achievement test performance. *Language Testing*, *20*(1), 26-56. https://doi.org/10.1191/0265532203lt243oa

Phakiti, A. (2008). Strategic competence as a third-order factor. *Language Assessment Quarterly*, *5*(1), 20-42. https://doi.org/10.1080/15434300701533596

Powers, D., & Wilson Leung, S. (1995). Answering the new SAT reading comprehension questions without the passages. *Journal of Educational Measurement*, *32*(2), 105-129. https://doi.org/10.1111/j.1745-3984.1995.tb00458.x

Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Pressley, M., & Gaskins, I. (2006). Metacognitively competent reading comprehension is constructively responsive reading: How can such reading be developed in students? *Metacognition and Learning*, *1*(1), 99-113. https://doi.org/10.1007/s11409-006-7263-7

Purpura, J. (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), 333-379. https://doi.org/10.1177/026553229801500303

Purpura, J. (1999). *Strategy use and second language test performance: A structural equation modeling approach*. Cambridge: Cambridge University Press.

Purpura, J. (2013). Cognition and language assessment. In A. J. Kunnan (Ed.), *The companion to language assessment* (pp. 1452-1476). Boston, MA: Wiley.

Rafi, M. F., & Islam, A. F. (2017). Test taking strategy on students' reading test. *Journal of English Language Teaching*, *4*(7), 42-51.

Roeschl-Heils, A., Schneider, W., & van Kraayenoord, C. (2003). Reading literacy, metacognition and motivation: A follow-up study of German students in Grades 7 and 8. *European Journal of Psychology of Education*, *18*, 75-86. https://doi.org/10.1007/BF03173605

Rogers, W. T., & Bateson, D. J. (1991). The influence of test-wiseness on performance of high school seniors on school leaving examinations. *Applied Measurement in Education*, *4*(2), 159-183. https://doi.org/10.1207/s15324818ame0402_5

Rupp, A., Ferne, T., & Choi, H. (2006). How assessing reading comprehension with multiple choice questions shapes the construct: A cognitive processing perspective. *Language Testing*, *23*(4), 441-474. https://doi.org/10.1191/0265532206lt337oa

Sheorey, R., & Mokhtari, K. (2001). Differences in the metacognitive awareness of reading strategies among native and non-native readers. *System*, *29*(4), 431-449. https://doi.org/10.1016/S0346-251X(01)00039-2

Singhal, M. (2001). Reading proficiency, reading strategies, metacognitive awareness and L2 readers. *The Reading Matrix*, *1*(1), 1-23.

Stemmer, B. (1991). *What's on a C-test taker's mind: Mental processes in C-test taking*. Bochum, Germany: Brockmeyer.

Storey, P. (1997). Examining the test-taking process: A cognitive perspective on the discourse cloze test. *Language Testing*, *14*(2), 214-231. https://doi.org/10.1177/026553229701400205

Szűcs, Á., & Kövér, Á. (2016). Reading skills involved in guided summary writing: A case study. *Working Papers in Language Pedagogy*, *10*, 56-72.

Tian, Shiau-ping. (2000). *TOEFL reading comprehension: Strategies used by Taiwanese students with coaching-school training* (Doctoral dissertation). ProQuest Dissertations and Theses database.

Tsagari, C. (1994). *Method effect on testing reading comprehension: How far can we go* (Master's thesis)? ERIC database.

Wijgh, I. F. (1995). A communicative test in analysis: Strategies in reading authentic texts. In A. Cumming, & R. Berwick (Eds.), *Validation in language testing* (pp. 154-170). Clevedon: Multilingual Matters Ltd.

Wu, A., & Stone, J. (2015). Validation through understanding test-taking strategies: An illustration with the CELPIP-General Reading Pilot Test using structural equation modeling. *Journal of Psychoeducational Assessment*, *34*(4), 362-379. https://doi.org/10.1177/0734282915608575

Wu, A., Chen, M., & Stone, J. (2017). Investigating how test takers change their strategies to handle difficulty in taking a reading comprehension test: Implications for score validation. *International Journal of Testing*, *17*(4), 1-23. https://doi.org/10.1080/15305058.2017.1396464

Yamashita, J. (2003). Processes of taking a gap-filling test: Comparison of skilled and less skilled EFL readers. *Language Testing*, *20*(3), 267-293. https://doi.org/10.1191/0265532203lt257oa

Yang, P. (2000). *Effects of test-wiseness upon performance on the Test of English as a Foreign Language* (Doctoral dissertation). ProQuest Dissertations and Theses database.

Yien, L. (2001). Effective test-taking strategies on English tests: Implications from Taiwanese students. *Hong Kong Journal of Applied Linguistics*, *6*(2), 22-43.

Yoshizawa, K. (2002). *Relationships among strategy use, foreign language aptitude, and second language proficiency: A structural equation modeling approach* (Doctoral dissertation). ProQuest Dissertations and Theses database.