Test-taking strategy use by Chinese young EFL learners in an authentic international standardized English test

Feifei Han
Griffith Institute for Educational Research, Griffith University

Received: 2 April 2022 / Accepted: 28 December 2022
DOI: https://doi.org/10.30827/portalin.vi39.24390
ISSN paper edition: 1697-7467, ISSN digital edition: 2695-8244

ABSTRACT: The current study examined language test-taking strategies used by Chinese young learners of English as a foreign language (EFLs) in an authentic international standardized English test–Cambridge Young Learners English Test: Flyers Test. It adopted a mixed-methods approach: the quantitative part was a survey administered to 138 participants, and the qualitative part consisted of think-aloud sessions and retrospective interviews with six participants (three were high-achievers and three were low-achievers). The quantitative results show that in general, children adopted metacognitive test-taking strategies more frequently than cognitive test-taking strategies irrespective of language skills; and the high-achievers used test-taking strategies more frequently than their low-achievers. The two-way interaction effects: levels of test performance-strategy type; and language skills-strategy type; were also significant. The qualitative results demonstrate that the high-achievers not only used a broader range of cognitive strategies than low-achievers, they also used the same metacognitive strategies at an in-depth level and in a more sophisticated way than low-achievers. The results of the study suggest that expanding Chinese young EFL learners' repertoire of test-taking strategies and modeling desirable ways of using strategies may help them improve their test-taking strategy use.
Key words: Chinese young EFL learners, test-taking strategy use, authentic international standardized English test, Cambridge Young Learners English Test, a mixed-methods approach.

Uso de estrategias para hacer exámenes por parte de jóvenes estudiantes chinos de inglés como lengua extranjera en un auténtico examen de inglés estandarizado internacional

RESUMEN: El estudio actual examinó las estrategias de realización de pruebas de idioma utilizadas por los jóvenes estudiantes chinos de inglés como lengua extranjera en una prueba de inglés estandarizada internacional auténtica: Cambridge Young Learners English Test: Flyers Test. Adoptó un enfoque de métodos mixtos: la parte cuantitativa fue una encuesta administrada a 138 participantes, y la parte cualitativa consistió en sesiones de pensamiento en voz alta y entrevistas retrospectivas con seis participantes. Los resultados cuantitativos muestran que, en general, los niños adoptaron estrategias de realización de pruebas metacognitivas con más frecuencia que estrategias de realización de pruebas cognitivas, independientemente de las habilidades lingüísticas; y los de alto rendimiento utilizaron estrategias para hacer exámenes con más frecuencia que los de bajo rendimiento. Los efectos de interacción bidireccional: niveles de desempeño de la prueba-tipo de estrategia; y habilidades lingüísticas-tipo de estrategia; también fueron significativos. Los resultados cualitativos de-

muestran que los de alto rendimiento no solo utilizaron una gama más amplia de estrategias cognitivas que los de bajo rendimiento, sino que también utilizaron las mismas estrategias metacognitivas a un nivel más profundo y de una manera más sofisticada que los de bajo rendimiento.
Palabras clave: Jóvenes estudiantes chinos de inglés como lengua extranjera, uso de la estrategia para tomar exámenes, auténtico examen de inglés estandarizado internacional, Prueba de inglés para jóvenes estudiantes de Cambridge, un enfoque de métodos mixtos.

1. Introduction

With an increasingly widespread use of English language in almost every domain in the world, English proficiency tests play an important role in many critical decision-making processes, such as acceptance of a student into a study program, recruitment of a worker, and a criterion for being a permanent resident in some English-speaking countries (Doe & Fox, 2011). These tests can motivate or discourage learners to learn English, in particular, in the context of EFL learning, where learners may not necessarily use English in their everyday life, but only learn English in classroom settings (Dawadi, 2020). As a result, it has long been vital for English language teachers, educational researchers, and English language training organizations to understand how various factors may impact on English language test performance.

Language assessment experts have proposed various theoretical models to explain language test performance (Bachman, 1990; 2000; Bachman & Palmer, 2010; McNamara et al., 2019). In these models, apart from test-takers' linguistic ability, test-taking strategy is regarded as one of the important non-linguistic factors accountable for variations in language test scores (Phakiti, 2016; Purpura, 2016). However, the existing literature on language test-taking strategy has largely focused on adult test-takers (Phakiti, 2003a, b; 2006, 2008a, b; Purpura, 1997, 1998, 1999;). There is a dearth of research on language test-taking strategy amongst young children (Gu & So, 2017; Han, 2018; Nikolov, 2006). With an increasing trend that English learning commences at an earlier age and the fact that more English proficiency tests are designed for young EFL learners (Nikolov & Timpe-Laughlin, 2021), research is needed to investigate language test-taking strategy use amongst the vast number of young children population. There is only limited literature on test-taking strategy use by young children. In particular, there is a lack of such investigation in authentic standardized testing situations. To fill this gap, the current study examined Chinese young EFL learners' test-taking strategy use in an international standardized English test–Cambridge Young Learners English Test: Flyers Test (hereafter Flyers Test). The following section reviews the relevant literature.

2. Literature Review

2.1. Operationalization of test-taking strategy use

In language assessment literature, test-taking strategy use belongs to strategic competence, which is one of the non-linguistic elements in Bachman and Palmer's (2010) communicative language ability. Bachman and Palmer define strategic competence as "the mental capacity

for implementing the components of language competence in contextualized communicative language use" (p. 106). Being is a higher-order cognitive mechanism, strategic competence executes through an individual's metacognition by orchestrating a set of strategies to regulate a person's online cognitive processing (e.g., coordinating strategies), linguistic processing (e.g., mental searching of linguistic elements), and psychological processing (e.g., regulating affect and anxiety), in order to accomplish a communicative goal (Phakiti, 2008a, b, 2016).

Strategic competence encompasses two components: strategic knowledge and strategic processing. While strategic knowledge is relatively stable and stored in the long-term memory (Han & Wang, 2017; Phakiti, 2008a, b); strategic processing, which is how an individual applies strategies to complete a specific task, heavily hinges upon the contexts and reflects an individual's concurrent information processing, and hence is relatively unstable (Cohen, 2007). In language testing, strategic knowledge concerns with test-takers' knowledge of test-taking strategies, whereas strategic competence is their actual use of the test-taking strategies in different testing situations. Having strategic knowledge is a necessary but not a sufficient condition for strategies to be used successfully, as concurrent information processing is also influenced by factors, such as task difficulty at hand, working memory capacity, and one's motivation of completing the tasks (Han, 2018; Jackson, 2020; Phakiti, 2008a). For example, the same child may use different test-taking strategies and/or use the same test-taking strategies qualitatively differently between when sitting an international standard language test and when completing a classroom language test.

Test-taking strategies have long been acknowledged to be difficult to define and to operationalize (Bachman & Cohen, 1998; Cohen, 2014; Nikolov, 2006; Phakiti, 2003a, 2016; Radwan, 2011). But researchers generally agree upon that for something to be meaningful as "strategies", the behaviors and processes have to take place within at least the peripheral attention if not within the focal attention of one's working memory (Cohen, 2007, 2014; Cohen & Macaro, 2007; Macaro, 2006; O'Malley & Chamot, 1990; Oxford, 2016).

Concurring with Cohen, Oxford (2016) states that "when strategy use is developed into an automatic operation (proceduralized) through repeated practice, it is no longer a strategy but an unconscious habit" (p. 51). Similarly, Grabe and Stroller (2013) agree that if strategies develop to a point of automaticity that learners are unaware of them or unable to describe them, they may not be called as strategies anymore. All these conceptualisations mean that *consciousness* and *intentionality* are essential characteristics of strategies, which can be used to distinguish strategies from other cognitive processing, such as skills, which are automatic and unconscious cognitive operations (Grabe & Stroller, 2013). Because of consciousness and intentionality qualities of strategies, assessment of strategy use can be achieved through learners' self-reports, such as responding to questionnaires, think-aloud methods, and/or retrospective reports.

2.2. Knowledge of test-taking strategies and language test performance

In language assessment research, the earlier studies on test-taking strategy did not make a clear distinction between knowledge of test-taking strategies and use of test-taking strategies. The design of these studies in reality examined the relations between knowledge of test-taking strategies and language test performance (Purpura, 1997, 1998, 1999; Song, 2005). Knowledge of test-taking strategies is often seen as consisting of knowledge of

cognitive and metacognitive strategies, which have been found to associate with and contribute to test performance differently; and the strength of association may be affected by test-takers' language proficiency.

For instance, Purpura (1997) found that adult learners' knowledge of cognitive strategies was directly and positively related to their performance in Cambridge First Certificate in English Anchor Test, whereas knowledge of metacognitive strategies was only indirectly related to the test performance through cognitive strategies. In a related study, Purpura (1998) used structural equation model (SEM) to model the relation between cognitive and metacognitive strategies as well as their relations with test performance amongst learners with two different levels of English proficiency. He found that cognitive strategies were more strongly associated with metacognitive strategies for the low-performers than for the high-performers. Amongst the high-performers, knowledge of metacognitive strategies had no direct effect on the test performance.

Song (2005) reduced 80 items in Purpura's (1997, 1998, 1999) questionnaire to 27 items and administered to adult test-takers of the Michigan English Language Assessment Battery. The exploratory factor analysis retained six cognitive strategies (i.e., repeating/confirming information strategies, writing strategies, practicing strategies, generating strategies, applying rules, and linking with prior knowledge) and three metacognitive strategies (i.e., evaluating, monitoring, and assessing). She found that knowledge of test-taking strategies contributed differently to test scores in different language skills, explaining 21.40% for writing, 17.20% for listening, and 12.50% for reading, grammar, and vocabulary section. Specifically, repeating/confirming information, linking with prior knowledge, writing and generating strategies were significant predictors of writing scores. Repeating/confirming information, linking with prior knowledge, and generating strategies were significant predictors of the listening performance. Repeating/confirming information, linking with prior knowledge, and monitoring were the three strategies, which significantly predicted the test-takers' performance on reading, grammar, and vocabulary section.

2.3. Test-taking strategy use and language test performance

Phakiti (2003a, b) is one of the first researchers who makes a clear distinction between test-taking strategy knowledge and use. To examine the actual use of test-taking strategies, Phakiti used simple past tense to construct questionnaire items and administered the questionnaire immediately after the tests. He instructed Thai university EFL learners to respond to the questionnaire by reflecting what strategies they had used and how frequently they had used them in completing a final-term English reading test. Different from the relational patterns between knowledge of test-taking strategies and test performance in Purpura (1997), Phakiti (2003b) found that both cognitive (r=.39) and metacognitive (r=.47) strategy use were significantly associated with the test performance. He also found that high-achievers used metacognitive strategies more frequently than moderate-achievers, who in turn, adopted metacognitive strategies more frequently than low-achievers. However, no significant difference was found between high- and moderate-achievers in terms of cognitive strategy use.

In another study, Phakiti (2006) used SEM to examine the relations between test-taking strategy use and English reading test performance. He found that out of the three types of cognitive strategies (i.e., comprehending, retrieving, and memorizing), only comprehending

had direct contribution to the reading test performance. The other two types of cognitive strategies functioned as mediators between metacognitive strategy use and the test performance. Retrieving mediated between evaluating strategy use and the test scores, whereas memorizing mediated between monitoring and the test scores. These results were similar to the results found in Purpura's (1998) participants with higher English proficiency, probably because the participants in Phakiti's study were also advanced EFL learners.

The direct and indirect contributions from cognitive and metacognitive strategy use to test performance were also replicated in longitudinal research by Phakiti (2007, 2008a, b). However, Phakiti (2007) mentioned that these relations should be interpreted by taking the characteristics of the test-takers (e.g., adult vs. child), the nature of the tests (e.g., standardized vs. non-standardized), and the skills of the tests (e.g., reading vs. speaking).

2.4. Children's test-taking strategy use

As mentioned, there is little research on children's test-taking strategy both in terms of knowledge and use of strategies. To the best of author's knowledge, only three studies have explored this area and two of them were small-scale qualitative studies. Gu and So (2017) explored 16 Chinese primary school students' (aged between 6 to 11) test-taking strategy use when they were undertaking a mock reading or listening section of *TOEFL Primary* test. Students were retrospectively interviewed immediately after they completed the task to gain an insight into their test-taking strategy use. Strategies were classified into language learner strategies, test-management strategies, and test-wiseness strategies according to Cohen's taxonomy (Cohen, 2012). Of the total 208 instances of strategy use, 12 types of language learner strategies were identified, accounting for approximately 40%. There were also 11 types test-management strategies and eight test-wiseness strategies, taking up 37 and 23% respectively.

Due to the small sample size, the relation between levels of test performance (high-performing, mid-performing, and low-performing) and test-taking strategy use were only examined using frequency counts to show some possible patterns. Of students with three levels of test performance, it was observed that high-performer used test-wiseness strategies least across reading and listening skills; but in reading they used language learner strategies the most. In listening, however, it was the mid-performers adopted language learner strategies the most.

These results seem to suggest an interaction between language skills, strategy type, and proficiency level, which was not able to be tested due to small sample size. Although Gu and So's study examined test-taking strategy use, it should be pointed out that the study only used the test items to elicit strategy use but was not conducted in a testing situation. It was likely that students were not motivated to use strategies as they would have in a real testing situation.

In another qualitative study conducted in a non-real testing context, Nikolov (2006) examined test-taking strategy use of 52 middle school students (aged between 12 to 13). She asked the participants to think aloud when they were undertaking reading and writing sections in a customarily designed English test. The study identified 16 types of metacognitive strategies, 13 types of cognitive strategies, and five types of social and affective strategies. Four case studies were also conducted with two high-achievers and two low-achievers. These case studies showed that using frequency of alone was difficult to reveal individual

differences of strategy use. Rather the ways how students combined different strategies qualitatively provided more insightful information (Nikolov). The two high-achievers were not only able to use their existing English knowledge but also related their world knowledge to the problems encountered in the test. In contrast, the two low-achievers frequently focused on unknown words.

Han (2018) was the only quantitative study which investigated children's language test-taking strategy use. This study was also conducted in a real testing scenario and involved four language skills. Three questionnaires were used to examine Chinese young EFL learners' test-taking strategy use in the three sections in the Flyer test (i.e., listening, reading and writing, and speaking). The confirmatory factor analyses found that Chinese young EFL learners' test-taking strategy use consisted of cognitive and metacognitive components across the four skills. An unexpected finding of this study was that cognitive and metacognitive strategy use in each skill did not significantly correlate with the test scores in that corresponding skill, except for a positive relation between metacognitive strategy use in reading and writing and the scores in reading and writing section.

3. Method

3.1. The current study and the research design

The current study is a follow-up of Han's (2008) study by addressing two gaps. First, as previous research suggested possible interactions between language skills, strategy type, and proficiency level (Gu & So, 2017), the current study will examine the extent to which language skills (listening, reading and writing, and speaking), strategy type (cognitive and metacognitive), levels of test performance (high-achieving, moderate-achieving, and low-achieving), and their interactions affect Chinese young EFL learners' test-taking strategy use. Second, to address the limitation of relying on frequency information alone (Nikolov, 2006), the current research also incorporated qualitative data from three high- and three low-achieving Chinese young test-takers to examine qualitatively different ways of using test-taking strategies by these learners.

The current research adopted mixed methods , which combined both quantitative and qualitative methodologies for data collection and analysis. Specifically, the current research adopted a convergent design, which involves concurrent and separate collection and analysis of quantitative and qualitative data (Creswell & Poth, 2017). The quantitative part used three questionnaires to collect test-takers' test-taking strategy use, whereas the qualitative part adopted think-aloud method and retrospective interview methods.

3.2. The research context

The research was conducted in an authentic international standardized English test for young EFL learners–Flyers Test. Flyers test is the third of a suite of three Cambridge Young Learners English Test and is designed to examine English proficiency of everyday written and spoken language for children in primary and lower-secondary school. It consists of three sections, namely listening, reading and writing, and speaking, and takes approximately 1

hour and 15 minutes to complete. The Flyers Test give a shield score between 1 and 5 to recognize learners' achievement in English learning rather than setting a cut-off score for pass or fail (see www.cambridgeenglish.org/exams-and-tests/flyers for details).

3.3. Participants

Altogether 144 Chinese young EFL learners (primary school students) from two test centres voluntarily participated in the study, because the number of the participants in one test centre was not sufficient for statistical analyses. Amongst them, 138 took part in the quantitative part and six were involved in the qualitative part. The 144 students all sat the Flyers Test held on the same day.

For the 138 participants in the quantitative part, the total shields were used to divide them into high-achievers (the top tertile: *N*=49), moderate-achievers (the middle tertile: *N*=43), and low-achievers (the bottom tertile: *N*=46). Table 1 presents the participants' total shields and the shields for different language skills.

Table 1. *Descriptive statistics of the participants' shields by proficiency*

Test scores (in shields)	High (N=49)		Moderate (N=43)		Low (N=46)	
	M	*SD*	*M*	*SD*	*M*	*SD*
Listening	3.80	0.64	3.00	0.62	2.35	0.95
Reading & writing	4.14	0.68	3.28	0.55	2.61	0.74
Speaking	4.98	0.14	4.98	0.15	4.91	0.28
Total	12.69	0.89	10.53	0.50	8.02	1.04

The six students who took part in the qualitative part of the study were selected from twenty students who indicated a willingness to participate in the qualitative part of the study. In order to achieve a balance between high-achievers and low-achievers, we asked students' Cambridge English training teacher to rate participants' English proficiency. We did not use students' shields of the Flyers Test because the shields results were not immediately available after the test, but the qualitative data collection had to be administered one week after the Flyers Test in order to minimize the lag between the Flyers Test and collection of the test-takers' retrospective reporting on their test-taking strategy use in the speaking. According to the English teacher's rating, three students who received highest ratings and three who received lowest ratings were selected as participants. One-week time was the shortest waiting period between the Flyers Test and the availability of the participants for the retrospective interviews. Of the six students, three were identified as high-achievers and three were low-achievers.

3.4. Instruments

3.4.1. Instruments for the quantitative data collection–Young EFL Learners' Strategic Processing Questionnaires

For the quantitative data collection, we used three questionnaires, namely Strategic Processing in Listening Questionnaire (13 items), Strategic Processing in Reading and Writing Questionnaire (17 items), and Strategic Processing in Speaking Questionnaire (14 items), to survey participants' test-taking strategy use immediately after they completed each test (see Han, 2008 for the details of the questionnaire development). Each questionnaire consisted of two scales and the reliability of the six scales were all acceptable: cognitive strategy use in listening (Cronbach's α = .81; e.g., "I used sound effects and tone of the speaker's voice to help me guess the meaning of words"), metacognitive strategy use in listening (Cronbach's α = .80; e.g., "When I had trouble understanding, I paid more attention and focused harder"), cognitive strategy use in reading and writing (Cronbach's α = .86; e.g., "I predicted what was going to happen next while I was reading the text"), metacognitive strategy use in reading and writing (Cronbach's α = .85; e.g., "When I lost concentration in reading I tried to pay more attention and focus harder"), cognitive strategy use in speaking (Cronbach's α = .81; e.g., "I made sure I used correct intonation when speaking"), and metacognitive strategy use in speaking (Cronbach's α = .75; e.g., "I planned what to say in my mind before I began speaking").

3.4.2. Instruments for the qualitative data collection–the sample Flyers Test and questions in the retrospective interviews

To elicit three high-achievers and three low-achievers' test-taking strategy use in listening, and reading and writing, students were asked to vocalize what was going on in their minds while they were completing the sample Flyers listening, and reading and writing tests (see https://www.cambridgeenglish.org/Images/young-learners-sample-papers-2018-vol1.pdf for the details of the sample tests). For the test-taking strategy use in speaking, the participants were asked to report what strategies they had used in the Flyers speaking test held one week earlier in the retrospective interviews (see Appendix 1 for the retrospective interview questions).

3.5. Data collection procedure

The responses to the Strategic Processing in Listening Questionnaire and Strategic Processing in Reading and Writing Questionnaire were collected in groups immediately after the participants completed the two tests. The collection of the Strategic Processing in Speaking Questionnaire was conducted individually following each test-taker's speaking test. The quantitative data collection was undertaken by the staff working in the Cambridge test centre. To minimize potential reading difficulties by students, the staff read each item in Chinese for the participants.

The qualitative data collection was carried out one week after the Flyers Test by an experienced research assistant. Both think-aloud sessions and retrospective interviews were conducted in Chinese on an individual basis in a quiet office. Before each think-aloud

session, the participant was trained in Chinese and was provided opportunities to practice thinking aloud until he/she felt comfortable to carry out the task. The retrospective interviews followed the think-aloud sessions.

3.6. Data analysis

To provide answers to the extent to which language skills, strategy type, levels of test performance, as well as their interactions affect Chinese young EFL learners' frequency of test-taking strategy use, a 3 - 2 - 3 mixed factorial MANOVA was conducted in SPSS 25. The within-subject independent variables were the three language skills and two strategy types. The between-subject independent variable was the three levels of test performance, and the dependent variables were frequencies of cognitive and metacognitive strategy use. To answer the question how high- and low-achieving Chinese young test-takers used test-taking strategies in qualitatively different ways, content analyses were adopted to analyse the think-aloud and retrospective interview data.

4. Results

4.1. The effects of language skills, strategy type, and levels of test performance on the frequency of cognitive and metacognitive test-taking strategy use

The descriptive statistics of cognitive and metacognitive test-taking strategy use in listening, reading and writing, and speaking by test-takers' proficiency are presented in Table 2.

Table 2. *Descriptive statistics by test-takers' proficiency*

Scales	High (N=46)		Moderate (N=49)		Low (N=43)	
	M	*SD*	*M*	*SD*	*M*	*SD*
Cognitive strategy use in listening	3.96	0.75	3.89	0.86	3.39	0.87
Metacognitive strategy use in listening	4.45	0.75	4.26	0.81	3.97	0.86
Cognitive strategy use in reading & writing	4.24	0.69	3.92	0.79	3.53	0.78
Metacognitive strategy use in reading & writing	4.37	0.74	4.04	0.79	3.65	0.83
Cognitive strategy use in speaking	4.02	0.81	3.79	0.74	3.19	0.79
Metacognitive strategy use in speaking	4.43	0.67	4.09	0.71	3.28	0.89

The results of the mixed MANOVA showed that all the three main effects were significant: language skills: $F(1, 270) = 9.47$, $p < .01$, partial $\eta^2 = .07$, strategy type: $F(1, 270) = 70.84$, $p < .01$, partial $\eta^2 = .34$, and levels of test performance: $F(2, 270) = 16.02$, $p < .01$, partial $\eta^2 = .19$.

In general, children used metacognitive strategies (*M*=4.06) significantly more frequently than cognitive strategies (*M*=3.77). The post-hoc analysis show that in the three sections of the test, children adopted test-taking strategies least frequently in speaking (*M*=3.80); but there was no significant difference between strategy use in listening (*M*=3.99), and reading and writing (*M*=3.96). High-achievers (*M*=4.21) used test-taking strategies significantly more frequently than moderate-achievers (*M*=3.97) and low-achievers (*M*=3.47), whose strategy use did not differ from each other.

In terms of the interaction effects, the three-way interaction effect was not significant: $F(4, 270)=2.17$, $p=.07$, partial $\eta^2=.03$, so was the two-way interaction between levels of test performance - strategy type: $F(2, 270)=0.65$, $p=.52$, partial $\eta^2=.01$. There was a significant interaction effect between language skills - strategy type: $F(2, 270)=12.97$, $p<.01$, partial $\eta^2=.09$. Separate paired-sample *t*-tests show that the mean differences between cognitive and metacognitive strategy use in the listening skill (-0.48) was more than that in the reading and writing skill (-0.12), and in the speaking skill (-0.27) (Figure 1).

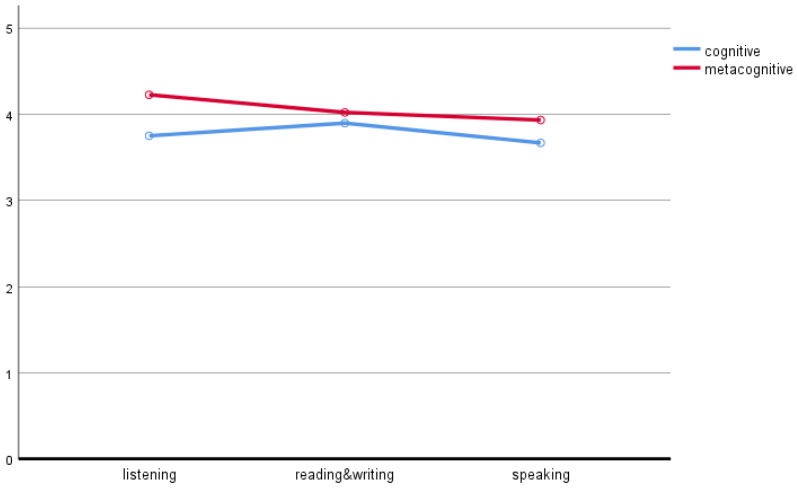

Figure 1. The interaction effect between language skills - strategy type

The interaction effect between language skills - levels of test performance was also significant: $F(4, 270)=5.10$, $p<.01$, partial $\eta^2=.07$. Post-hoc analyses demonstrate that the mean differences of test-taking strategy use between high- and low-achievers was more in speaking (0.99) than in reading and writing (0.72), and in listening (0.52). In speaking, and in reading and writing tests, the differences between high- and moderate-achievers, and between moderate- and low-achievers were significant. However, in listening test, the difference between high- and moderate-achievers was non-significant (Figure 2).

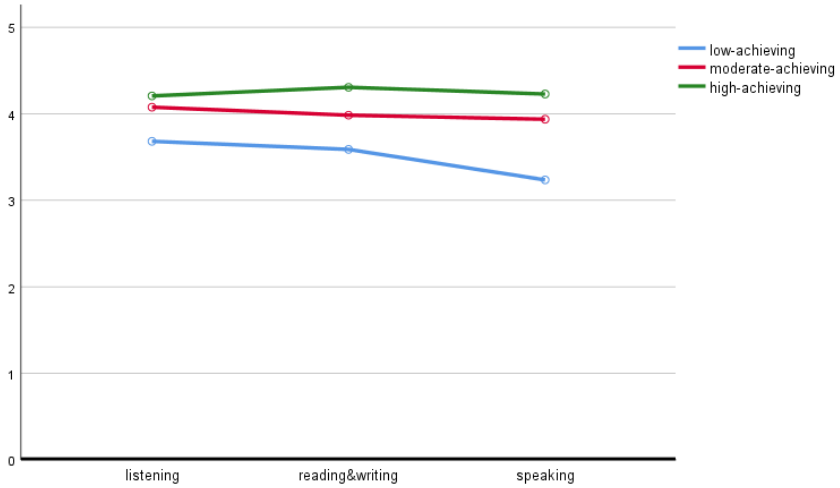

Figure 2. The interaction effect between language skills - levels of test performance

4.2. Qualitative differences between high- and low-achievers test-taking strategy use

The content analyses of the qualitative data revealed two main differences between the high-and low-achievers: one was concerned with cognitive strategy use and the other was about metacognitive strategy use. For cognitive strategy use, higher-achievers used a broader range of strategies than low-achievers. With regard to metacognitive strategy use, the high- and low-achievers used the same strategies qualitatively differently, with the high-achievers using the same strategies at an in-depth level and in a more sophisticated way.

While high-achievers were able to use grammatical knowledge (a cognitive strategy), this strategy was completely missing in low-achievers' reporting. For instance, in completing cloze test, both high- and low-achievers used comprehending strategy (a cognitive strategy) to assist them in selecting the appropriate item. In addition, all the three high-achievers used grammatical knowledge to help them filter out the items which would be grammatically incorrect if put into the sentence. Take the sentence of "Helen likes learning and ___ out about old things" as an example, a high-achiever reported that:

> I know that 'likes' should be followed by a noun or a gerund, like… the word 'learning' in the sentence is a gerund. But the word 'out' seems not be able to follow a noun, most likely it should be a gerund. So, when I look at the list of the items, I only need to focus on gerunds. Those items of other part of speech shall not be considered. So, after reading these words, of course I also need to consider if the meaning is appropriate or not. So, I now limited my selections to the word "seeing" and "finding" in the list. But you know… there is no such phrase like 'seeing out', 'seeing' is not supposed to be followed by a proposition, so, I think I will select 'finding'. 'Helen likes learning and finding out about old things", yes, I think the meaning is plausible too, I will stick to 'finding'.

Using grammatical knowledge was completely missing in all the three low-achievers' think-aloud data. Without using such strategy, they often made incorrect selection. For example, when a boy read the sentence "Dinosaurs have been extinct for ___", through comprehending strategy, he mentioned: "I think this blank required a word related to a word expressing time." But he did not think what grammatical category this required word should be. After reading all the items, he said: "I think none of the words which I know had a meaning related to time" (He misidentified the word "centuries" as "countries"). He then commented: "Now 'agreed' is the only unknown word left, so I had to select this one.

In addition, the high- and low-achievers also used metacognitive strategies qualitatively different from each other. Such differences were mostly manifested in the speaking. While both high- and low-achievers used monitoring strategy (a metacognitive strategy), the high-achievers reported that they used such strategy to check if both the content and grammar of their oral production was appropriate in responding to the examiner's question. However, the low-achievers used monitoring strategy to make sure that they imitated native speakers' intonation rather than focusing on meaning. Focusing on initiation in fact was avoided by high-achievers. For instance, a high-achiever mentioned:

> I tried not to pay much attention to the intonation, because I know the primary focus was to express the meaning correctly. So, I did not want to myself being distracted by paying too much intonation when I put my ideas into English words.

The high- and low-achievers also differed in using evaluation strategy (a metacognitive strategy) to make decision as to whether they would ask the examiner to repeat or clarify the question which they did not understand for the first time. A high-achiever said:

> Sometimes, when I did not understand a question raised by the examiner, I was thinking I should ask her to repeat because I knew that I wasn't listening, as I was too nervous and did not concentrate. But I know that if she repeated or explained in a more detailed way, I would understand. But I did not always ask her to repeat, because…I felt those new words in her speech, no way would I understand even if she said this sentence to me again. So, I did not bother to ask her to repeat. I just guessed what she might mean and then answered according to my guess.

The low-achievers, however, did not evaluate the scenario to make a distinction as to why they did not understand the examiner's question. They reported that whenever they did not understand the questions raised by the examiner, they never asked again and simply made a guess.

5. DISCUSSION

Different from Gu and So's (2017) results, our study showed that no matter what language skills, the high-achievers used test-taking strategies more frequently than low-achievers in general. It should be noted that even though Gu and So reported the reversed patterns of using language learner strategies between high- and mid-performers in reading and lis-

tening, the participants in reading and listening groups were different (a between-subjects design). Our study used a within-subjects design and examined the same participants' use of test-taking strategies across different language skills. Our findings are similar to those reported in Phakiti (2003a, b), which also used a within-subjects design and found similar patterns with adult EFL learners. This may suggest that child and adult EFL test-takers share similarity in using test-taking strategies.

An interesting finding was the interaction effect between language skills - levels of test performance. We found that while the differences between high- and low-achievers' strategy use was more pronounced in speaking than in other language skills, high- and low-achievers test- did not differ in their speaking test shields ($F(2, 137)=1.57$, $p=.21$, partial $\eta^2=.02$). There are two possible interpretations of this result. First, it could be that the speaking test was relatively easy than the listening test for all the participants. Hence, even though students differed much in using strategies during the speaking test, their test scores were not significantly different. Second, as indicated by Nikolov (2006), sometimes the qualitatively different ways of using a particular strategy rather than frequency per se were responsible for the individual differences between high- and low-achievers test-taking strategy use. Indeed, our qualitative results provide some evidence to the latter interpretation. We found that in speaking test, the high-achievers used both monitoring and evaluating strategies in a more deep and meaningful manner than low-achievers, who used them rather simplistically and at a surface level.

Another interaction effect–language skills - strategy type suggests that when the test-takers were involved in non-visual modality (listening and speaking tests), the frequency of using metacognitive and cognitive strategies were more different than when they were involved in visual modality (reading and writing test). A possible reason could be that in the non-visual modality, information could not be kept long and accessed later, hence, test-takers might focus more on planning and monitoring (metacognitive strategies) in the process. At the same time, because of the fast fading of non-visual modality, it might be difficult to apply some of the cognitive strategies which required longer processing time (e.g., translating, lexical inferencing) (Vandergrift, 2003).

6. Practical implications

As young EFL learners' repertoire of test-taking strategies is still in the development and not yet complete (Nikolov & Timpe-Laughlin, 2021), equipping young EFL learners with the knowledge of a wider range of strategies for them to draw on may help them become more strategic in language tests. Teachers may wish to teach test-taking strategies through explicit instruction, such as how grammatical knowledge can help in the tasks such as cloze test. Teachers can also invite high- and low-achievers to demonstrate how they use a certain strategy to show qualitatively different ways of using the same strategy. Through comparison and contrast, the desirable ways of applying strategies in language tests can be highlighted for students.

7. Acknowledgement

This paper reports on research funded through the Cambridge English Funded Research Program, Round Six, 2015, by Cambridge English Language Assessment.

8. References

Bachman, L. F. (1990). *Fundamental considerations in language testing.* Oxford University Press.

Bachman, L. F. (2000). Modern language testing at the turn of the century: Assuring that what we count counts. *Language Testing, 17*(1), 1-42. https://doi.org/10.1177/026553220001700101

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

Bachman, L., & Cohen, A. (1998). Language testing – SLA interfaces: An update. In L. Bachman & A. Cohen (Eds.), *Interfaces between second language acquisition and language testing research* (pp. 1-31). Cambridge University Press.

Cohen, A. D. (2007). The coming of age of research on test-taking strategies. In J. Fox, M. Weshe, D. Bayliss, L. Cheng, C. Turner, & C. Doe (Eds.), *Language testing reconsidered* (pp. 89-111). Ottawa University Press.

Cohen, A. D. (2014). *Strategies in learning and using a second language.* Routledge.

Cohen, A. D., & Macaro, E. (Eds.). (2007). *Language learner strategies: Thirty years of research and practice*. Oxford University Press.

Creswell, J. W., & Poth, C. N. (2017). *Qualitative inquiry and research design: Choosing among five approaches (4th ed.).* Sage.

Dawadi, S. (2020). Parental involvement in national EFL test preparation. *RELC Journal 51*(3), 427-439. https://doi.org/10.1177/0033688219848770

Doe, C., & Fox, J. (2011). Exploring the testing process: Three test takers' observed and reported strategy use over time and testing contexts. *Canadian Modern Language Review, 67*(1), 29-54. https://doi.org/doi:10.3138/cmlr.67.1.029

Grabe, W. P., & Stoller, F. L. (2013). *Teaching and researching: Reading.* Routledge.

Gu, L., & So, Y. (2017). Strategies used by young English learners in an assessment context. In M. Kim & Y. Butler (Eds.), *English language proficiency assessments for young learners* (pp. 118-135). Routledge.

Han, F. (2018). Strategic processing of Chinese young English language learners in an international standardized English language test. *Frontiers in Psychology.* https://doi.org/10.3389/fpsyg.2018.01020

Jackson, D. O. (2020). Working memory and second language development: A complex, dynamic future? *Studies in Second Language Learning and Teaching 10*(1), 89-109. https://doi.org/10.14746/ssllt.2020.10.1.5

Macaro, E. (2006). Strategies for language learning and for language use: Revising the theoretical framework. *The Modern Language Journal*, *90*(3), 320-337. https://doi.org/10.1111/j.1540-4781.2006.00425.x

McNamara, T., Knoch, U, Fan, J., & Rossner, R. (2019). *Fairness, justice & language assessment.* Oxford University Press.

Nikolov, M. (2006). Test-taking strategies of 12-and 13-year-old Hungarian learners of EFL: why whales have migraines. *Language Learning, 56*(1), 1-51. https://doi.org/10.1111/j.0023-8333.2006.00341.x

Nikolov, M., & Timpe-Laughlin, V. (2021). Assessing young learners' foreign language abilities. *Language Teaching, 54*(1), 1,37. https://doi.org/10.1017/S0261444820000294

O'Malley, J., & Chamot, U. (1990). *Learning strategies in second language acquisition.* Cambridge University Press.

Oxford, R. (2016). *Teaching and researching language learning strategies.* Longman.

Phakiti, A. (2003a). A closer look at gender and strategy use in L2 reading. *Language Learning, 53*(4), 649-702. https://doi.org/10.1046/j.1467-9922.2003.00239.x

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

Phakiti, A. (2006). Modeling cognitive and metacognitive strategies and their relationships to EFL reading test performance. *Melbourne Papers in Language Testing, 1*, 53-96.

Phakiti, A. (2008a). Construct validation of Bachman and Palmer's (1996) strategic competence model over time in EFL reading tests. *Language Testing, 25*, 237-272. https://doi.org/10.1177/0265532207086783

Phakiti, A. (2008b). Strategic competence as a fourth-order factor model: A structural equation modeling approach. *Language Assessment Quarterly, 5*(1), 20-42. ttps://doi.org/10.1080/15434300701533596

Phakiti, A. (2016). Test-takers' performance appraisals, appraisal calibration, and cognitive and metacognitive strategy use. *Language Assessment Quarterly, 13*(2), 75-108. https://doi.org/10.1080/15434300701533596

Purpura, J. E. (1997). An analysis of the relationships between test-takers' cognitive and metacognitive strategy use and second language test performance. *Language Learning, 47*, 289-325. https://doi.org/10.1111/0023-8333.91997009

Purpura, J. E. (1998). Investigating the effects of strategy use and second language test performance with high- and low-achieving test-takers: A structural equation modeling approach. *Language Testing, 15*, 333-379.

Purpura, J. E. (1999). *Learner strategy use and performance on language tests: A structural equation modeling approach.* Cambridge University Press.

Purpura, J. E. (2016). Second and foreign language assessment. *The Modern Language Journal, 100*(S1), 190-208. https://doi.org/10.1111/modl.12308

Radwan, A. A. (2011). Effects of L2 proficiency and gender on choice of language learning strategies by university students majoring in English. *Asian EFL Journal*, *13*(1), 115-163.

Song, X. (2005). Language learner strategy use and English proficiency on the Michigan English Language Assessment Battery. *Spaan Fellow Working Papers in Second or Foreign Language Assessment, 3*, 1-26.

Vandergrift, L. (2003). Relationships among motivation orientations, metacognitive awareness and proficiency in L2 listening. *Applied Linguistics, 26*, 70-89. https://doi.org/10.1093/applin/amh039

Wang, Z. & Han, F. (2017). Metacognitive knowledge and metacognitive control of writing strategy between high- and low-performing Chinese EFL writers. *Theory and Practice in Language Studies, 7*(7), 523-532. http://doi.org/10.17507/tpls.0707.04

9. Appendix

Retrospective Interview Questions

You had a speaking test last Sunday, didn't you? Could you please reflect on that experience and answered my questions in relation to that experience?

1. In the speaking test, when you could not understand the examiner's question, what did you do?
2. In the speaking test, when you realized that you did not make it clear of the meaning you would like to express, what did you do?
3. In the speaking test, what element(s) did you pay attention to and how you monitor the element(s) (e.g., pronunciation, intonation, use of words, the meaning of the sentences)?