CrossMark

Behind a spoken performance: test takers' strategic reactions in a simulated part 3 of the IELTS speaking test

Christina Judy Fernandez

Correspondence: cjud2115@uni.
sydney.edu.au
Sydney School of Education and
Social Work, University of Sydney,
Sydney, New South Wales 2006,
Australia

Abstract

Making inferences about a test taker's language ability has been a concern for language testers. Defining language ability in a speaking test is made difficult because the construct involves factors within and beyond the test taker. One underlying factor is the use of strategies. This study probes test takers' strategic processes and their use in completing the simulated part 3 of the International English Language Testing System (IELTS) Speaking Test. Data based on stimulated recall from 12 international university students in Sydney, Australia, upon the completion of the speaking task reveal that strategies are mobilised from the moment an input question is presented to the end of a response. Overall, a total of 18 individual strategies from three categories were used by participants in this study: seven cognitive, five communication and six metacognitive strategies. Findings suggest that strategies were deployed in clusters to produce a response. Commonly mobilised strategies were not always useful nor did they positively impact participants' test response quality.

Keywords: IELTS speaking test, Stimulated recall, Test-taking strategies, Strategic competence, Cognitive processes, Cognitive validity

Introduction

With the growing international mobilisation of individuals, especially for migration and higher education purposes comes the demand for individuals with elevated language proficiency or ability levels (Bachman 2014; Purpura 2016). Governments and educational institutions alike require these individuals to prove their proficiency levels by taking a language proficiency test.

For Australian-bound international students, the International English Language Testing System (IELTS) is one of the four English proficiency test evidence shown to support their visa and university applications (Australian Government Department of Home Affairs n.d.). The IELTS consists of reading, writing, listening and speaking sub-tests. Test takers are awarded an aggregated score based on the four sub-tests where a Band 1 is deemed a 'Non-user' while a Band 9 is an 'Expert user' (IELTS 2017). Decision to admit or reject a prospective student into a university program is made on that single test score by a university staff member.

Making such life-changing decisions for an individual solely on a single test score continues to be a concern among language testing researchers. Language knowledge

alone is not a determinant of a test taker's language ability. In fact, test takers' individual cognitive or mental processes invoked by a task are crucial as they result in test score variability (Purpura 2014). These processes may be automatically or deliberately mobilized to complete a task. This conscious, intentional choice is what differentiates strategies from automatic cognitive processes (Cohen 2014; Phakiti 2007). Understanding the relationship between strategy use and the processes, researchers argue, would provide insights into test takers' language ability and how they could perform in real world situations (e.g. O'Sullivan and Weir 2011). Research studies on how these processes work and impact test takers' performance are lacking especially in the speaking test domain. The present study addresses this gap and investigates test takers' strategic processes in a two-way speaking discussion test task.

Literature review
Strategic competence in speaking test construct
Efforts to explain and define the complexity of an individual's language ability have been made through the proposal of various multi-componential models (e.g. Bachman 1990; Bachman and Palmer 2010; Canale and Swain 1980). For instance, both Canale and Swain (1980) and Bachman and Palmer (2010) in their models included linguistic (i.e. grammatical, sociolinguistic and pragmatic knowledge) and non-linguistic components (i.e. ability to appropriately use the language) known as *strategic competence*.

Earlier models described strategic competence as serving two roles in performance: (i) to remedy imminent communication problems due to test takers' grammatical and pragmatic deficiencies and (ii) to enhance communication (Canale 1983; Canale and Swain 1980). Drawing from research developments in cognitive psychology and speech production, Bachman (1990) explains that strategic competence involves cognitive processes that manage the use of language. Later, Bachman and Palmer (2010) extended that role to include metacognitive strategies that both manage cognitive processes and interact with test takers' internal factors such as language and topical knowledge, attributes and affective schemata. Although these models conceptualize strategic competence in the speaking skill, empirical evidence is lacking on how it works in a speaking test.

Given its role in test performance, a growing number of researchers believe that the inclusion of strategic competence in test constructs (and scoring rubrics) strengthens their validity (e.g. Weir 2005). Weir (2005) and later O'Sullivan and Weir (2011) in their socio-cognitive framework proposed *cognitive validity* or evidence of test takers' cognitive processes as a component of test validation across the four language skills. Cognitive validity deals with the degree to which test tasks in any skill prompting cognitive processes when the skill is applied in a real language use situation (e.g. O'Sullivan and Weir 2011; Weir 2005). For a test to be valid, O'Sullivan and Weir (2011) emphasize that a test must find evidence of test takers' characteristics, cognitive processes, the test context (i.e. response format, time constraints) and rater and rating process. These are also factors that work together to impact test takers' performance. Cognitive processes invoked should be aligned with those in real-world situations in order for tests to be valid (Weir 2005). However, to include such processes in a speaking test scoring rubric, more studies are called for to investigate how the processes

could be (i) observed in a spoken performance and by the rater(s), ii) used differently across score levels and tasks and iii) useful to test score users (e.g. Barkaoui et al. 2013; Swain et al. 2009). This study responds to the first area of research needed.

IELTS speaking test

The IELTS speaking test is a three-part interview administered by an IELTS certified examiner lasting 11 to 14 min. In part 1 (Introduction), after a scripted introduction and checking of a candidate's identification, the examiner presents the candidate with a series questions about themselves, their interests and other familiar topics (Seedhouse and Nakatsuhara 2018). This part lasts 4 to 5 min. In part 2 (Long turn), the candidate speaks at length for 1 to 2 min on a topic chosen by the examiner. One minute of preparation time is given prior to the monologue. The examiner ends this part with one or two rounding-off questions. This part lasts 3 to 4 min. To ensure similar test experiences and controlled input for all candidates, examiners adhere to a strict script in both parts 1 and 2 (Seedhouse and Nakatsuhara 2018). In part 3 (Two-way Discussion), candidates discuss more abstract issues and concepts related to the topic from part 2. This final part also lasts 4 to 5 min. Unlike previous parts, here, the examiner may rephrase the graded question prompts to fit the level of a candidate. Furthermore, in part 3, candidates are expected to use a range of language functions from explaining and describing to making speculations (Taylor 2007). Candidates are marked on four criteria: fluency and coherence, lexical resource, grammatical range and accuracy, and pronunciation (see Seedhouse and Nakatsuhara 2018 for further details).

Validation studies on the IELTS speaking test have investigated various factors affecting test takers' performance. These include test task or format (Brown 2006; Iwashita and Vasquez 2015), timing (Wigglesworth and Elder 2010) and topic (Khabbazbashi 2017). Others have investigated examiner and examiner rating processes (Carey and Mannell 2009). Thus far, only one study by Huang (2013) has investigated how strategic processes impact performance. Most of these studies investigated student performance in all three parts or only part 2 of the speaking test (Iwashita and Vasquez 2015; Wigglesworth and Elder 2010). None of the aforementioned studies focused on the final or part 3 of the IELTS speaking test which, Quaid (2018) asserts by its less structured nature provides 'examiners with a better indication of true performance and score' compared to the previous parts (p. 7).

Test taking strategies

During a test, test takers mobilise three types of strategies: (1) *language learner strategies*; (2) *test management strategies* and (3) *test wiseness* (Cohen 2014). Language learner strategies assist test takers in prompting necessary language for a task. The latter two strategies, test management and test wiseness are test-taking strategies. Test management strategies are conscious actions and thoughts that help test takers respond to a test item or task meaningfully. They are mobilised to prompt test takers' task-related language knowledge or skills and are, therefore, construct-relevant. Test wiseness strategies involve using knowledge about the test format and its peripherals rather than linguistic ability in completing a test task.

These strategies are construct-irrelevant. The degree to which construct relevant and irrelevant strategies are used, Cohen (2014) argues, determines the validity of a test.

Researchers have employed different methods to elicit test takers' strategy use. The most commonly used method is stimulated recall (e.g. Huang 2013; Huang 2016; Phaiboonnugulkij and Prapphal 2013; Swain et al. 2009). Swain et al. (2009) also used semi-structured interviews to clarify details from stimulated recall sessions conducted earlier. Huang (2013) gathered data from participants' video-recorded speaking test or 'oral language production' to verify observable strategies that were not revealed in their stimulated recall sessions (p. 11). A self-designed strategy inventory was used in addition to drawing from data obtained through stimulated recall in the Huang (2016) study. Shohamy (1994) and Yoshida-Morise (1998) elicited strategy use through analysis of recorded interviews. Both studies investigated the use of communication strategies only.

Research on test taking strategies in speaking tests

Swain et al. (2009) (and Barkaoui et al., (Barkaoui et al. 2013) essentially based on the same study) explored test takers' strategies use in two independent (speaking only) and four integrated (listening, reading and speaking or listening and speaking) tasks of the TOEFL iBT. All 30 participants underwent stimulated recall after each task and strategies from five categories were found: affective, approach, cognitive, communication and metacognitive. Metacognitive, communication and cognitive strategies were most frequently used. Using the Spearman's *rho* non-parametric test, it was found that some individual strategies had negative correlation with test scores (i.e. *self-correcting, attending*). This was attributed to such strategies requiring more use of attentional resources than others while speaking.

Huang (2016) examined how strategies impacted test takers' performance in six tasks of the TOEIC speaking test, also a computer-based test. Of the 215 Taiwanese students tested, eight were chosen for the stimulated recall. A 41-item self-designed inventory was drawn up from this data to collect quantitative data. The inventory was based on the taxonomy from previous studies (e.g. Barkaoui et al. 2013; Huang 2013; Swain et al. 2009). Focusing on three broad strategy categories, namely, cognitive, communication and affective, Huang (2016) employed exploratory factor analysis and structural equalling model (SEM) to analyse the data. He reported that cognitive and communication strategies positively affected performance. Communication strategies impacted performance more directly than cognitive strategies.

Huang (2013), unlike the aforementioned studies, probed test takers' strategy use in the IELTS speaking test, a direct test. Purported to investigate strategy use in testing and non-testing situations, her 40 Chinese-speaking university students were divided equally into two respective groups. Through the use of MANOVA, Huang found that test takers used strategies differently in testing and non-testing situations. Findings also showed that different strategies were invoked across the three tasks. Participants in the testing group employed more approach and metacognitive strategies in task 3 compared to those in the non-testing group.

Implications for the present study

The literature reviewed has informed this study in terms of theory and methodology. Various studies reviewed found that different tasks prompted the use of different strategies. Huang (2013) investigated strategy use in the IELTS speaking test and found that test takers used different individual strategies in the three parts. Her study, however, did not address the reasons and ways strategies were used in completing the tasks.

The aforementioned speaking test studies have mostly involved semi-direct tests where test takers are devoid of interaction with another language user (e.g. Barkaoui et al. 2013; Huang 2016; Phaiboonnugulkij and Prapphal 2013; Swain et al. 2009). This 'non-reciprocal' language situation is limited in terms of meaning and or message negotiation which test takers would otherwise experience in direct tests (Bachman and Palmer 2010, p. 34). Therefore, what is lacking in test-taking strategies research is how strategies are used when language use is reciprocal or involves an unfamiliar interlocutor. With just the Huang (2013) study having focused on this issue, further studies are necessary.

Quantitative approaches employed to analyse data in the studies reviewed have undoubtedly been useful in providing frequency counts that inform patterns of strategy use. However, they do not describe patterns of such phenomenon (Creswell 2015). Swain et al. (2009) admitted that using quantitative analysis limited them from exploring issues such as whether strategies worked in a particular sequence, how test takers choose strategies, and how each individual strategy works in a task.

In light of the research gaps discussed above, this study aims to address the following research questions:

1. What strategies do test takers report using in the simulated part 3 of the IELTS speaking test?
2. How useful are these strategies in helping test takers succeed in the speaking test task?

Methodology

Research participants

The participants were 12 international students attending a university in Sydney, Australia. They were seven females and five males whose ages ranged from 19 to 36 years. At the time of data collection, eight participants were pursuing a Master's degree while four others were completing a Bachelor's degree. They represented seven nationalities (four from China; three from Malaysia; one each from Brazil, India, Mexico, Saudi Arabia and Vietnam) and seven different language backgrounds. Their official IELTS speaking test scores ranged from 5.5 to 8.5.

Purposive sampling was applied in that only those who have taken the IELTS examination and, at the time of study, were enrolled in a degree or postgraduate studies at a university were chosen. Participants' familiarity with the IELTS speaking test format was emphasized (as they might have taken other proficiency tests before arriving in Australia) since no test preparation training for it was going to be provided prior to or during the study. Ethics approval was obtained through the researcher's university

ethics committee (2015/974). Prior to the study, participants were emailed about the overview of the study and the methodology involving audio- and video-recordings of themselves. On the day of the study, participants were reminded of this again and they signed consent forms as well as a background questionnaire. A pseudonym was assigned as per participants' gender to protect their identity.

Data collection
Data collection began with each participant experiencing a 4–5-min simulated part 3 of the IELTS speaking test (hereafter, speaking test). A 1–2-min warm-up session was conducted asking them three informal questions prior to introducing the discussion topic. All participants were then administered the same topic about meetings at school or work to minimize topic effect. The test topic with two prompt frames for part 3 of the IELTS speaking test was adapted from the Cambridge IELTS 8 text (see Additional file 1). This session was audio- and video-recorded.

Participants were asked to leave the room for a 10-min break after the speaking test. During this time, the researcher previewed the video recording to determine when and where to pause, and what possible follow-up questions to ask (if any). This ensured that the researcher remained an unobtrusive, active listener during the stimulated recall (McKay 2009).

Instructions for the stimulated recall, adapted from Swain et al. (2009) were then presented to the participants (see Additional file 2). The exact time(s) for pausing the video recording was noted to ease locating the different times when strategies were used.

Data coding and reliability
The coding scheme was derived from existing classifications found in speaking test-taking studies (e.g. Huang 2013; Swain et al. 2009) and other strategy use literature (e.g. Nakatani 2005; Phakiti 2003; Purpura 1997). Huang (2013) and Swain et al. (2009) shared similar classifications of five categories: *affective, approach, cognitive, communication and metacognitive*. Huang (2013) included *social*, a sixth category, since her study investigated strategy use in the IELTS speaking test where interaction with an examiner was crucial. The present study focused on *cognitive, communication* and *metacognitive* strategies (see Additional file 3).

Both Huang (2013) and Swain et al. (2009) found the use of 90 and 49 individual strategies by their test takers encompassing the aforementioned six and five categories, respectively. In order to increase coding consistency and reliability, the individual strategies from their taxonomies were collapsed at two levels: (1) between categories and (2) within categories (Révész 2012). In the first level, individual strategies from different categories that shared similar meaning or use were combined into one individual strategy. For example, four individual strategies from different categories reportedly used by test takers in Swain et al. (2009) and Huang (2013) were combined into *planning*, a metacognitive strategy for the present study:

1) *Thinking ahead* (communication)

2) *Organizing thoughts* (communication)
3) *Setting goal* (metacognitive)
4) *Outlining* (cognitive)

In the second level, where separate strategies were listed within a category in the aforementioned taxonomies, these were treated as a single strategy with subdivision or types. For instance, Huang (2013) listed *evaluating* (metacognitive) as seven separate individual strategies but was treated as a single strategy subdivided into four types in this study (see Additional file 3). These evidence-based adjustments made on the coding schemes from previous studies were necessary not only to ensure coding reliability but also to fit the purpose of this study in exploring the use of strategies in performance (Gu 2014; Révész 2012).

Data coding entailed several stages. The first stage involved electronic coding of individual transcripts (see Table 1). Only complete utterances or sentences that indicated strategy use were coded. For instance, an excerpt of a stimulated recall section from Holly, a graduate student test taker was coded as below.

The first sentence 'I stopped…meetings.' was coded *PAS (b)* (*Pausing/Slowing*—b type). Next, 'the first action…tongue' was coded as *UFL (a)* (*Using First Language*—a type). The sentence 'I think although it… spend some time.' was coded as *EVA (d)* (*Evaluating*—d type)**.** The sentences were highlighted in pink, green and blue to represent strategies in communication, cognitive and metacognitive categories, respectively.

Next, four transcripts were randomly chosen for the second stage of coding by an expert whose doctoral study was in test-taking strategies. The inter-rater reliability percentage was calculated based on the number of agreements on codes between the researcher and the expert coder divided by the number of initial codes given by the researcher on all four scripts (Goh 2002). The reliability percentage was 70%. All differences were discussed and resolved by either eliminating the item coded or compromising.

In the final stage, the researcher recoded the remaining eight transcripts in a random order. The intra-coder reliability percentage was calculated using a similar formula to the one for inter-rater reliability. The intra-coder reliability recorded 82% similarity with the researcher's total number of codes from the first stage.

Results
Speaking task strategies
In order to answer the first research question on strategy use in part 3 of the IELTS, a frequency count of strategies used was conducted. Overall, test takers reported using seven cognitive, five communicative and six metacognitive individual strategies. To further understand the usage pattern(s) of each strategy, all protocols of a particular strategy from the stimulated recall transcripts were collated.

Table 1 Coding procedure and codes used

Protocol	Code
I stopped there coz I am thinking about different meetings. You know when you ask me a question, the first action is to translate your question into my mother tongue. I think although it [translating the question] spend some time.	PAS (b) UFL (a) EVA (d)

Cognitive strategies

Table 2 provides protocols from test takers' stimulated recall and the usage patterns for each individual cognitive strategy. Although listed as isolated individual strategies, they appeared to be used in conjunction with other strategies to produce a response (e.g. *recalling* followed by *abandoning*, a communicative strategy). There was also a tendency for two strategies to be used for the same purpose (e.g. *making up an answer* and *linking with previous knowledge/experience*). Another tendency was for strategies to be used both before and while speaking as in the case of *linking with previous knowledge/experience* and *using first language.*

Communication strategies

Table 3 highlights protocols and the usage patterns of communication strategies based on the stimulated recall. Similar to cognitive strategies, individual strategies here were also deployed in conjunction with others to complete the task (e.g. *asking for examiner's help* utilized after *analyzing the input*, a cognitive strategy).

Table 2 Individual cognitive strategies used and usage patterns

	Stimulated recall	Usage patterns
1. Analyzing input	When you say meetings, -- more than one, it's not just one meeting or two meetings… why do people meet there… (Andy)	- Used upon examining keywords from input question(s) to then plan content of response
2. Anticipating grade or examiner's reaction/impression	I want to use some good words not just 'They are nervous' to get a high score. (Dolly) I gave you [the examiner] some personal experience so you can think 'this candidate is not memorizing a whole sample' (Faye)	- Mobilised when adjustments made post-monitoring or evaluating of performance - Driven by goal to obtain higher score or impress the examiner
3. Linking previous knowledge/ experience	I remember my first class here… so I took as an example for answer. (Diane)	- Triggered by input questions; language and content for response drawn from mental images of experience or knowledge - Idea development hindered when topic knowledge is limited; tendency to then to make-up details
4. Making-up answer	The example, I ask my friend but he is not willing to attend because sheer waste of my time… did not happen. (Harry)	- Used when no relevant experience or knowledge comes to mind - Combine reality and fictional facts to make up; examiner cannot detect truth or lie
5. Recalling	To explain diplomatic, I could not find the word …like a good relationship between two countries. (Alice)	- Used while producing utterances to express or clarify ideas - Time-consuming lead to topic avoidance or reformulating the response
6. Summarizing thoughts verbally	During a past skype meeting I was checking my facebook or on the mobile phone…. That's why I said people are much more 'present' at face-to-face meetings. (Bill)	- Used while speaking to control rush of ideas and to avoid lengthy, irrelevant explanation
7. Using first language	I was thinking in Malay how to categorize them [meetings]…what comes to mind was company meeting…and with friends (Gina)	- Used to understand meaning while listening to input questions or to locate synonyms while responding - Applied when having difficulty explaining concepts in L2 (e.g. categorizing formal and informal meeting) or those learned in L1

Table 3 Individual communication strategies used and usage patterns

	Stimulated recall	Usage patterns
1. Abandoning the message	I did not know how to say it because I do not have the information…So I could not include it. (Alice)	- Applied when thoughts are irrelevant to question - Also when ideas or words inaccessible while speaking
2. Asking examiner's help	I wasn't clear about the question so I ask to repeat. (Don)	- Used when input question or expressions not understood, heard clearly or as time-buying measure when mind is drawing blank
3. Elaborating	I was trying to say something about my neighbour… to talk for a longer period. (Diane)	- Applied to keep response fluid, meet speaking test requirements or clarify a message - Often resort to personal experiences or aspects not covered earlier
4. Pausing/slowing down	I spent time thinking about other meetings. At first, I did not think of… (Holly)	- Used to generate ideas, formulate speech or filter through a string of thoughts
5. Reformulating output	I was trying to structure the sentence -- a thought came and I was thinking how to express that also. So I started with 'sort of' and changed. (Bill)	- Used when a different idea had presented itself or useful language was inaccessible while responding

Abandoning the message or *reformulating output* were both mobilised for the same purpose when words or thoughts were inaccessible while speaking.

Metacognitive strategies

Table 4 presents the protocols and usage patterns for metacognitive strategies. Individual metacognitive strategies also tended to be used in conjunction with other strategies

Table 4 Individual metacognitive strategies used and usage patterns

	Stimulated recall	Usage patterns
1. Assessing the situation	Why it's important to go to meetings? I paused. I get a little lost…do not think seriously about this topic and never encounter reasons… (Gwen) Your [examiner] response is important to me – you smile or nod to me I think may be this answer is right. (Dolly)	- Applied when making judgements while or after listening to input question(s) - Judgments about question difficulty, familiarity with topic, examiner's purpose for asking a question - Judgments about examiner's reaction/response while speaking
2. Attending	Different pictures came to mind. I chose education and medicine…(Andy)	- Used when deciding from several word- or content-related choices; - Occurs after analyzing input or assessing the question
3. Evaluating	I was thinking I did not say everything there was. (Alfred)	- Used when judging own: language production, response quality, thought processes and strategies used; - Often followed by planning
4. Monitoring	I thought meetings and interviews are different. Should I or not talk about my interview? So I stopped talking about it. (Gina)	- Applied when a doubt or a problem is identified while speaking that leads to an immediate action
5. Planning	The next point in my mind is ok I have to give an example to prove my point. (Harry)	- Used after examining input question or identifying a content-related problem while speaking
6. Self-correcting	'Occurred to me'…but I was saying 'occurred to mind', Oh it's bad grammar. So I changed.(Faye)	- Used upon realizing a grammatical error - Corrections made immediately in mid-sentence with no pauses or hesitations but not always correct

(e.g. *attending* used after *analyzing the input*, a cognitive strategy or *assessing the situation*, a metacognitive). Similar to cognitive strategies, there was no fixed point as to when some strategies were used (e.g. *planning*).

Individual strategy use in performance

In order to answer the second research question on usefulness of individual strategies in completing the speaking task, data from the stimulated recall (*SR*) were compared to the speaking test (*ST*) transcripts. More, specifically, each strategy use was matched up to the corresponding segment of the speaking test. Speaking test responses were then analysed in relation to the three IELTS speaking test criteria adapted from Brown (2006):

1) Fluency and coherence: response length, speech rate (pauses/slowing down) and topic coherence
2) Lexical resource: range, accuracy and effectiveness of paraphrasing and circumlocution
3) Grammatical range and accuracy: sentence structure variety, accuracy and complexity

Detailed analysis of the three most reported individual strategies in each group are discussed below.

Individual cognitive strategies

These help test takers understand the examiner's input to then formulate a response and manipulate their use of language to complete the speaking task. Of the seven individual strategies, test takers reported mostly using *linking previous knowledge or experience*, *analyzing input* and *anticipating grade or examiner impression.* The three strategies are analysed in detail below.

Linking previous knowledge or experience Used mostly to clarify or elaborate on an idea, this strategy tended to have positive or negative impacts on test takers' response quality. Test takers attributed their depth of knowledge on the topic as the reason for such differences. Excerpts 1 and 2 highlight the different impacts of using this strategy on two test takers' responses on why international leaders meet. In Excerpt 1, Bill produced a generally fluid 21-s response with minimal hesitations and pauses at the beginning. The reformulation 'to take part uhm' to 'to help out' did not affect the coherence of the message. His inclusion of 'war', 'poverty' and 'natural disasters' not only strengthened the topic development but also showcased his knowledge of relevant vocabulary.

Excerpt 1

ST: (LAUGHTER) (.) That's an interesting topic uhm (.) I think international leaders meet to discuss about their economic relationships or issues in the world such as war or poverty or even natural disasters to take part uhm to help out or give aid to other countries.

SR: The point you mention international meetings uhm I heard about meetings
…..also remembered some United Nations meetings to solve problems…(Bill).

Excerpt 2, however, shows the negative impact using this strategy had on Gina's
response. The lack of knowledge depth on a certain topic seem to contribute to
the ineffectiveness of this strategy use. Gina's response here lasted 45 s but was
filled with pauses, fillers and other time-buying measures such as elongation of
words (i.e. may: be:). Thus, the fluency of her message suffered. Furthermore, a
'textbook' or generic answer based on her previous knowledge left her response
without much elaboration, and limited her use of language to convey the basic
meaning of cultural differences.

Excerpt 2

ST: (3) uhm (.) because (.) may: be: there are certain issues that's (3) kinda (.) uhm
(5) probably because you know when you are talking to people from certain regions
they have their own culture, you know (.) their own perspectives, the way they say it
is different and (.) most probably leaders from – they have all these meetings
because that certain issue are really related so they have to come to like a conclusion
to uhm (.) that's going to benefit both of them.

SR: I know that country leaders meet and they want to find benefits – that's all…I
am not confident to talk about it…. so I just gave a 'textbook answer'. (Gina).

Analyzing input Test takers mobilised this strategy to understand the meaning of the ex-
aminer's input question by examining its keyword(s) which consequently helped them de-
cide on the content scope of their response. This strategy use tends to impact test takers'
response positively. Andy, in Excerpt 3, produced a response lasting for 1 min and 14 s.
Although there was about a 12-s (including the hesitation) pause initially, the processing
time Andy spent examining the question proved fruitful. By considering an individual's
reasons for attending meetings, Andy included expressions such as 'say or present some-
thing', 'questions to ask you' and 'problems are solved', resulting in a topically coherent and
appropriately worded response to the question.

Excerpt 3

ST: Examiner: Why is it important to attend meetings?

Andy: (6) uhm (.) I think (4) I think it's -- it depends on why are you going there? For
example, like (.) if you want to say or present something, you should be there…
Sometimes after you present, they [your audience or colleagues] will have questions to
ask you, talk to you…you also listen to other people…problems are solved eventually.

SR: When you [the examiner] say meetings, means more than one. It's not just one
meeting or two meetings or just meetings – the importance is why do people meet
there. (Andy).

Anticipating grade or examiner impression Mobilised mainly to obtain a better grade or to impress the examiner, this strategy involved test takers' making sudden content, vocabulary or grammar-related adjustments. In the study, test takers gravitated to this strategy after evaluating or monitoring their performance. In Excerpt 4, Dolly was asked to clarify her idea of 'social skill' in a follow-up question from the examiner. To avoid repetition, Dolly used expressions such as 'trustworthy', 'talk to others appropriately' and 'good practice' which did not merely showcase her vocabulary range but also coherently developed the topic. The presence of hesitations, however, minimally affected the fluidity of her message.

Excerpt 4

ST: Examiner: What do you mean by social skills?

Dolly: You know uhm some people are not good at uhm (.) I mean some people does not know who is trustworthy or how to uhm uh (.) talk to other appropriately...have more and more meetings maybe it's good practice to improve your social skills.

SR: I want to use some good words not just 'They are nervous' to get a high score. (Dolly)

Individual communication strategies

Generally, test takers mobilised these strategies to solve content or linguistic-related problems to produce a message while maintaining interaction with the examiner. The three most frequently used individual communication strategies were: *pausing/slowing down*, *elaborating* and *reformulating the output*.

Pausing/slowing down Test takers used this strategy to generate ideas, formulate speech or filter through a string of thoughts. The tendency was for test takers to use this strategy at the beginning and midpoint of their responses. Excerpt 5 illustrates Andy's 23-s response in which he paused at several points (about 8 s in total). Pauses and hesitations resulted in a disfluent response with numerous ungrammatical structures (e,g., 'if there's a disconnected between ....'). The abrupt ending indicating his preference for a face-to-face meeting was also irrelevant to the question, thereby affecting the coherence of his response.

Excerpt 5

ST: Examiner: What causes misunderstanding?

Andy: uh sometimes depends on the language you are communicating in? Hmm and also (2) sometimes it's like, not sure (.) but it's like it happens if (.) if there's a disconnected between each other like (.) depends on the call and (2) I think it's (.) actually, I prefer the face to face one.

SR: I was thinking in this moment… what I am going to say or something real that happened before. (Andy).

Elaborating Used generally to clarify a message and fulfil the question or speaking test requirement(s), this strategy helped test takers to speak in great length later. Test takers accomplished this by either providing an example or introducing other ideas not presented earlier. Don, in Excerpt 6 below, expressed his attitudinal change in approaching strangers by narrating his initial experience in Australia. By so doing, he overcame a brief period of silence (about 3 s) and continued speaking for another 21 s.

Excerpt 6

ST: At that time, it was hard for me to talk to strangers. At the moment it's better because I have some uhm (.) yeah it's a lot – I remember it's the first time I come, I came to Australia,…((narrates his experience)).

SR: I want to explain more for you. I compare my feeling at the moment and the past time. I think may be more clear for you. (Don).

Reformulating the output Test takers deployed this strategy when changes to their original message were necessary while speaking. These adjustments were often necessary because a different idea had presented itself or useful language was inaccessible while speaking. In their responses, these were apparent when test takers pause and paraphrase or re-start the message. In Excerpt 7, Bill paused before re-starting to describe the feeling of 'invisible'. Bill's use of this strategy appeared to have helped him showcase his ability to manipulate the language necessary to express the downside of Skype interviews.

Excerpt 7

ST:… If it's over skype, then sometimes you sort of – I mean (.) there will be more place for you to feel that 'I am invisible' there? Even though I am on the screen …

SR: I was trying to structure the sentence -- a thought came and I was thinking how to express that also. So I started with 'sort of' and changed. (Bill)

Individual metacognitive strategies
These strategies help test takers plan, monitor and evaluate the process of formulating a response and overall performance. Detailed analysis of the three most commonly reported strategies *evaluating*, *planning* and *assessing the situation* are presented below.

Evaluating This strategy was mobilised when test takers evaluated their response or performance, thought processes and strategy use. Given that judgments occur in test

takers' minds, the impact of using this strategy is not apparent in their responses. Therefore, Excerpt 8, only depicts Faye's SR data. Faye became aware that her discussion was limited to similarities and not the differences in formal meetings. Although the content of her response could not be changed, mental notes were made on how she could have changed the situation.

Excerpt 8

SR: So it's kinda narrow, my way to answer this question. I realized that I only mentioned the similarities but not the differences. So I am thinking that maybe if I mention more kinds of the meeting in the beginning, that would really help. (Faye)

Planning Test takers deployed this strategy after being presented with an input question or identifying a problem or confusion while speaking. In Excerpt 9, Alfred responded to why world leaders meet. This 36-s response began with a brief pause when Alfred was possibly reviewing topics he had covered earlier and thereafter, made plans for his response. Overall, his response was well-worded and fluid. Drawing on his knowledge about the Paris meeting, other relevant ideas such as 'to partner', 'listen to their opinions' and 'vote' were included to support reasons for international meetings.

Excerpt 9

ST: Uhm (.) The last world meeting that I have heard which was in Paris, the world leaders met to discuss environmental issues… really important because they have to be there to listen, to partner maybe with one another. And also to listen to their opinions and to take a decision together so they can vote in some different issues or aspects.

SR: I had already said the importance positive, negative aspects and now she [the Examiner]'s asking world, the global. So I thought about I will talk about the environmental issues and environmental aspects… I did not say anything about that before. (Alfred)

Assessing the situation Test takers employed this strategy to assess two situational factors: the examiner's response or feedback and the input question. Judgments were made on the examiner's verbal and non-verbal responses. A smile and nod from the examiner may indicate that the response is acceptable, and thus motivate test takers to speak more. In fact, in Excerpt 10, Harry shared how assessing the examiner's non-verbal responses spared him from re-explaining an idea.

Excerpt 10

SR: The examiner's expression is very important here so I know my point has been made clear. Otherwise I'll need to go back and re-explain. (Harry).

The examiner's intentions for input questions, however, were not always assessed positively. Excerpt 11 lists test takers' evaluative statements about follow-up questions. Diane was left confused while Gina found the follow-up question repetitive.

Excerpt 11

SR: The next question you [the examiner] re-direct me… I thought I answered the question wrongly. (Diane)

I said about miscommunication … and then you asked me about communication efficiency again. (Gina)

Excerpt 12 highlights Gwen's response upon assessing the follow-up question on importance of attending meetings. Other than a 2-s pause at the beginning, her response lasted for 19 s and was generally fluid and coherent. She explains further that accessing her personal reasons for attending meetings helped her respond.

Excerpt 12

ST: Uhm (.) I think it's the face-to-face value because (.) when you confront with real people….questions solved completely.

SR: Why it's important to go to meetings? I paused. I get a little lost…do not think seriously about this topic and never encounter reasons… meetings ok for me so I think some reasons. (Gwen)

Discussion

The objective of this study was to explore test takers' strategic processes and how these work in a speaking test task of dialogic nature. The following discussion begins with the overall use of strategies in the speaking test followed by the patterns of their use in successfully completing the speaking test task.

Findings revealed that test takers used strategies from the moment an input or follow-up question was presented until the end of their response to that question and thereafter. These findings suggest that strategies were not merely used when test takers had a language or content-related problem, but also to sustain or manage and complete the task. Two possible explanations emerge for the use of strategies in this study. Firstly, this could be the case of the cognitive processes invoked by the task itself (Weir 2005). The speaking task required constant listening and comprehending of input, planning for a response not just to a question but also to fulfil language function requirements. These continuous requirements may have triggered test takers' different cognitive processes and strategy use.

Secondly, this phenomenon could be what Fulcher (2003) described as learners' 'cognitive capacity to manage communication' (p. 31). In other words, test takers' use of strategies here may have been automatic. Such automaticity would only be possible with learners' increased knowledge about the second language and use in communicative situations (Field 2011; Fulcher 2003; Pawlak and Waniek-Klimczak 2015). This

may be applicable to the test takers in the current study as they were already in the midst of doing an undergraduate or graduate degree in an English-speaking country and were familiar with the IELTS test format. The need for this group of test takers to use strategies, however, could be attributed to possibly the different language acquisition stages that they may be in, given they come from different language backgrounds to acquire English (Ortega 2009). With different groups of first language learners' tendency to stay longer than others in acquisition stage, completing a second language communicative task such as the IELTS task may require them to rely on strategies to maintain communication.

In the present study, test takers tended to rely considerably more on metacognitive and cognitive strategies, rather than communication strategies, to complete the speaking task. This finding only partially supports that of Swain et al. (2009) and Huang (2013) where their test takers used more metacognitive and communication strategies than cognitive strategies. In fact, in Huang (2013), of the six categories investigated, cognitive strategies was reported as the fifth-most used strategy group. A possible explanation for this could be attributed to the fact that this study only focused on the two-way discussion of the IELTS Speaking Test. In this task, test takers are faced with not only a discussion in a second language on an unfamiliar topic and frequent shifts in topic sequence, but also the need to showcase their range of language functions from explaining and describing to making speculations (Taylor 2007). Complexities in producing a response may have caused test takers to be consistently planning, monitoring and evaluating their performance. Thus, the high use of metacognitive strategies. These challenges coupled with the speed to comprehend input and produce a response might also have resulted in their frequent use of cognitive strategies. Similar findings (metacognitive, cognitive and communication strategy categories) were, however, reported in Phaiboonnugulkij and Prapphal (2013) where test takers were subjected to producing appropriate responses related to complex work-related language functions. This seems to suggest that the use of metacognitive and cognitive strategies might be associated with task difficulty.

Another probable explanation for this finding could be attributed to how participants in this study were prompted during the recall sessions. Although questions by the researcher were kept to a minimum, the emphasis of the prompting questions on 'thinking' or 'thoughts' may have led participants to highlight more metacognitive and cognitive rather than communicative strategies during the recall process.

The strategies used were not always beneficial to the test takers. In the current study, there were numerous instances where the cognitive strategy, *linking to previous knowledge/experience*, mobilized by test takers to retrieve language or content for response production had negatively affected test takers response quality. In Song (2005), this strategy was reported as having positive impact on test takers' performance. Swain et al. (2009) found that metacognitive strategies had negative effect on test takers' performance. In both studies, the individual strategy use were compared to test takers' score to arrive at such conclusion. The present study compared the coded transcripts of test takers' stimulated recall and the oral data from their speaking test qualitatively. The difference in these findings may be attributed to the analyses methods used.

The overall use patterns of strategies seem to support the theory that strategies work in clusters and not in isolation as discussed by Cohen (2011; 2014). Cohen (2011) writes that in a conversation, speakers may employ strategies in sequence and then sporadically in clusters to complete the task at hand. The qualitative analysis of this study suggests that each strategy has a different role in the process of producing an utterance. These individual strategy roles fit into the successful characteristics of completing the IELTS speaking test as outlined by Seedhouse and Harris (2011). Individual strategies were likely to assist test takers by: 1) identifying and understanding the topic of the input question; 2) providing an answer to that question; and 3) developing or expanding the topic or idea. Test takers, in the current study, were likely to use a set of strategies such as *analyzing input* (cognitive) or *assessing the situation* (metacognitive) to identify key aspect(s) in formulating the probable content of their conversation, followed by *planning* (metacognitive) or *recall* (cognitive) to generate ideas to develop the topic and *elaborating* (communication) or *making up an answer* (cognitive) to expand on or develop an idea.

Limitations and directions for future research

Interpreting the findings of this study should be considered in light of a number of limitations. Firstly, this study involved only 12 international university students who have taken the IELTS and whose scores ranged between 5.5 and 8.5. The findings, therefore, might be reflective only of test takers who are well-versed with the test format and mid to high scoring. Strategies reported by this group of participants, therefore, might not encompass those used by test takers of lower proficiency level and without previous test experience. Another limitation is that the investigation focused on the strategy types and their usefulness in task completion of the part 3 of the IELTS speaking test. As mentioned earlier, this part of the test was chosen for its nature to draw out test taker' performance that could be predictive of their performance in the real-world (Quaid 2018). The findings of this study, therefore, may not be representative of those used in this entire test. Thirdly, because this study examined test takers' strategy use in a simulated test, strategies reported here may not be representative of ones used in a real IELTS test as anxiety and other factors may play a role in the nature of strategies used. Therefore, the reported findings may not be conclusive.

Future studies could be done involving a larger group of participants with a wider range of proficiency levels. These studies could also compare test takers' strategy use with examiners' thought process while marking the speaking test. By doing this, researchers might be able to identify strategies that are observable to the examiner and also evaluate the successful use of each strategy with regards to test takers from different proficiency levels.

Consideration should also be given to questions that were not addressed in this study. These include questions pertaining to relationship between strategy use and test performance, and also individual and contextual factors that attribute to test takers' use of strategies. Future studies investigating these questions may consider the use of mixed methods design. Data from the current study could be used to develop a questionnaire to obtain data from more test takers and possibly administered on test takers after a real speaking test.

Implications of the study

Although a small-scale study, insights obtained from this study can nonetheless be invaluable. The range of useful and helpful strategies presented here goes beyond a testing context. Teachers preparing students for IELTS speaking test and even general communication contexts might benefit from incorporating strategies with multiple purposes to best help enhance their students' performance. One strategy that teachers could focus on could be *elaborating* (communication). This could be done by first helping students build their topical knowledge and gradually moving into techniques of how to discuss a topic from general to specific. Techniques such as providing examples, explaining causes and effects as well as using sources to support their argument could help them continue speaking in a logical, linear manner in both testing and communication contexts.

For test writers and examiners, these findings may guide them to write and present purposeful input and follow-up questions that would invoke strategies to successfully complete the speaking task. Perhaps, better guidelines of what constitutes a follow-up question could be drawn up and included in examiner training materials. These guidelines would also help examiners decipher the way part 3 is administered, so as to go beyond the sequence of 'topic-based question answer adjacency pairs' as is commonly the case in part 1 of the IELTS speaking test (Seedhouse and Harris 2011, p. 69).

Conclusion

The present study was motivated by the lack of strategy use studies in the speaking test domain, especially involving direct or face-to-face tests. Although the use of strategies or strategic competence has been acknowledged as a key component affecting individual test takers' performance in various models of speaking ability, it has yet to be included in assessment rubrics or scales. This study responds to the call for studies to examine how strategies work and can be observed in test performance so that their use could eventually be included in scoring scales.

In line with this call, test takers' use of strategies and how individual strategies were used in in completing the third and final part of the IELTS speaking test was explored. By employing stimulated recall, insights on the inner workings of strategy use were obtained from 12 international student participants who underwent a simulated part 3 of the test. It was revealed that strategies were used continuously further supporting that strategy use is integral to performance. The predominant use of metacognitive and cognitive strategies further suggests that such speaking test tasks require test takers to think about their thoughts and produce a response in real time. Examiners and test designers alike, therefore, need to be cognizant and empathetic of test takers' response process when designing and or administering test tasks.

A comparison between the stimulated recall and participants' oral test data revealed that individual strategies have specific roles that contribute to successfully completing of the speaking test task. Strategies tended to be used in clusters to achieve better quality responses. This means that incorporating strategy use into the teaching of speaking is of importance to enhance students' performance in not just a testing but also in this ever-changing communicative context.

Additional files

Additional file 1: Appendix 1. (DOCX 17 kb)
Additional file 2: Appendix 2. (DOCX 15 kb)
Additional file 3: Appendix 3. (DOCX 26 kb)

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References
Australian Government Department of Home Affairs (n.d.a). Immigration and citizenship: English language proficient English. Retrieved from November 21, 2018, from https://immi.homeaffairs.gov.au/help-support/meeting-our-requirements/english-language/proficient-english
Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
Bachman, L. F. (2014). Ongoing challenges in language assessment. In A. J. Kunnan (Ed.), *The Companion to Language Assessment* (Vol. III, pp. 1538–1603). Hoboken: Wiley/Blackwell.
Bachman, L. F., & Palmer, A. (2010). *Language assessment in practice*. Oxford: Oxford University Press.
Barkaoui, K., Brooks, L., Swain, M., & Lapkin, S. (2013). Test-takers' strategic behaviors in independent and integrated speaking tasks. *Applied Linguistics, 34*(3), 304–324.
Brown, A. (2006). Candidate discourse in the revised IELTS speaking test. *IELTS Research Report, 6*, 71–90.
Canale, M. (1983). On some dimensions of language proficiency. In J. Oller (Ed.), *Issues in language testing research* (pp. 333–342). Rowley: Newsbury House.
Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics, 1*(1), 1–47.
Carey, M. D., & Mannell, R. H. (2009). *The contribution of interlanguage phonology accommodation to inter-examiner variation in the rating of pronunciation in oral proficiency interviews, IELTS Research Reports, 9* (pp. 217–236). Canberra: British Council and IELTS Australia.
Cohen, A. D. (2011). Strategies in learning and Using a Second Language. (2nd ed.) New York: Pearson Education Limited.
Cohen, A. D. (2014). Using test-wiseness strategy research in task development. In A. J. Kunnan (Ed.), *The Companion to Language Assessment* (Vol. II, pp. 893–905). Hoboken: Wiley/Blackwell.
Creswell, J. W. (2015). *Educational research: Planning, conducting and evaluating quantitative and qualitative research* (5th ed.). Upper Saddle River: Pearson Education Inc..
Field, J. (2011). Cognitive validity. In L. Taylor (Ed.), *Studies in language testing 30: Examining speaking: research and practice in assessing second language speaking* (pp. 65–111). Cambridge: Cambridge University Press.
Fulcher, G. (2003). *Applied Linguistics and Language Study: Testing Second Language Speaking*. Harlow: Pearson Education Limited.
Goh, C. M. (2002). Exploring listening comprehension tactics and their interaction patterns. *System, 30*(2), 185–206.
Gu, Y. (2014). To code or not to code: dilemmas in analyzing think-aloud protocols in learning strategies research. *System, 43*, 74–81.
Huang, H. D. (2016). Test taking strategies in L2 assessment: the test of English for international communication speaking test. *Perceptual and Motor Skills, 123*(1), 64–90.
Huang, L. S. (2013). Cognitive processes involved in performing the IELTS speaking test: respondents' strategic behaviours in simulated testing and non-testing contexts. *IELTS Research Reports Online Series, 1*, 1–51.
IELTS, (2017). How IELTS is scored. Retrieved April 6, 2018, from https://www.ielts.org/about-the-test/how-ielts-is-scored

Iwashita, N., & Vasquez, C. (2015). *Examination of discourse competence at different proficiency levels in IELTS speaking part 2, IELTS Research Reports, 5* (pp. 1–44). Melbourne: IELTS Partners: British Council, Cambridge English Language Assessment and IDP: IELTS Australia.

Khabbazbashi, N. (2017). Topic and background knowledge effects on performance in speaking assessment. *Language Testing, 34*(1), 23–48.

McKay, S. L. (2009). Introspective techniques. In J. Heigham & R. A. Croker (Eds.), *Qualitative research in applied linguistics: a practical introduction* (pp. 220–241). Basingstoke: Palgrave Macmillan.

Nakatani, Y. (2005). The effects of awareness-raising training on oral communication strategy use. *The Modern Language Journal, 89*(1), 76–91.

O'Sullivan, B., & Weir, C. J. (2011). Test development and validation. In B. O'Sullivan (Ed.), *Language Testing: Theories and Practices* (pp. 13–32). Hampshire: MacMillan Palgrave.

Ortega, L. (2009). *Understanding second language acquisition*. London: Hodder Education.

Pawlak, M., & Waniek-Klimczak, E. (2015). *Issues in teaching, learning and speaking and in a second language*. Berlin: Springer.

Phaiboonnugulkij, M., & Prapphal, K. (2013). Online speaking strategy assessment for improving speaking ability in the area of language for specific purposes: the care of tourism. *English Language Teaching, 6*(9), 19–29.

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.

Phakiti, A. (2007). *Strategic competence and EFL reading test performance: a structural equation modeling approach*. Oxford; Frankfurt: Peter Lang.

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*(2), 289–325.

Purpura, J. E. (2014). Cognition and language assessment. In A. J. Kunnan (Ed.), *The Companion to Language Assessment* (Vol. III, pp. 1452–1476). Hoboken: Wiley/Blackwell.

Purpura, J. E. (2016). Second and foreign language assessment. *The Modern Language Journal*, *100* (supplement). In *190–208*.

Quaid, E. D. (2018). Reviewing the IELTS speaking test in East Asia: theoretical and practice-based insights. *Language Testing in Asia, 8*(2), 1–9.

Révész, A. (2012). Coding second language data validly and reliably. In A. Mackey & S. M. Gass (Eds.), *Research methods in second language acquisition: a practical guide* (pp. 203–221). Malden, MA: Wiley-Blackwell.

Seedhouse, P., & Harris, A. (2011). Topic development in the IELTS Speaking Test. *IELTS Research Reports, 12*, 55–110.

Seedhouse, P., & Nakatsuhara, F. (2018). *The discourse of the IELTS speaking test: interactional design and practice*. Cambridge: Cambridge University Press.

Shohamy, E. (1994). The validity of direct versus semi- direct oral tests. *Language Testing, 11*(2), 99–123.

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.

Swain, M., Huang, L. S., Barkaoui, K., Brooks, L., & Lapkin, S. (2009). *The speaking section of the TOEFL iBT (SSTiBT): test-takers' strategic behaviours, TOEFLiBT Research Report, 10*. Princeton: Educational Testing Service.

Taylor, L. (2007). The impact of joint-funded research studies on IELTS speaking module. In L. Taylor & P. Falvey (Eds.), *Studies in language testing 19: IELTS collected papers*. Cambridge; New York: Cambridge University Press.

Weir, C. J. (2005). *Language testing and validation: an evidence-based approach*. Hampshire: MacMillan Palgrave.

Wigglesworth, G., & Elder, C. (2010). An investigation of the effectiveness and validity of planning time in speaking test tasks. *Language Assessment Quarterly, 7*(1), 1–24.

Yoshida-Morise, Y. (1998). The use of communication strategies in LPIs. In R. Young & W. He (Eds.), *Talking and testing: discourse approaches to the assessment of oral proficiency* (pp. 205–238). Amsterdam: John Benjamins.