

The Impact of the IELTS Speaking Test on Teaching Methods and Test Preparation Strategies in China

LIU YUAN

[0000-0001-5899-0201]

School of Languages, Literacies and Translation

Universiti Sains Malaysia

Penang, Malaysia

Yuan.liu 7@student.usm.my

ALLA KHAN (Corresponding author)

[0000-0003-4435-4583]

School of Languages, Literacies and Translation

Universiti Sains Malaysia

Penang, Malaysia

allabaksh@usm.my

SALASIAH CHE LAH

[0000-0001-6814-6346]

School of Languages, Literacies and Translation

Universiti Sains Malaysia

Penang, Malaysia

salasiah@usm.my

ABSTRACT

The increasing popularity of the IELTS test preparation course in China has sparked a discussion about the level of knowledge among instructors in terms of language assessment literacy and their selected teaching methods approach. The primary objective of this study is to gain insights into the intended test evaluation goals by Chinese IELTS speaking test teachers and how it influences their selection of teaching methods. For this study, 32 Chinese IELTS speaking test teachers were invited to participate in the research. The findings indicate that teachers demonstrated a relatively satisfactory understanding of the IELTS speaking test design. However, there were differing opinions among teachers regarding the test criteria, particularly in regard to the absence of explicit mention of skills in the test materials. Moreover, the findings of the present study revealed a link between how teachers perceived the design of the target test and their support for two instructional strategies that they believed could facilitate students' learning of the abilities that will be evaluated in the actual test. The findings of this study offer valuable insights to educators, curriculum developers, and policymakers involved in IELTS, particularly regarding the significance of gathering teachers' perspectives on the goals of the target test. This study specifically highlights how vital it is to prepare teachers in language assessment literacy. Providing teachers with knowledge about testing allows them to make informed decisions and utilise more effective pedagogical practices. Through a thorough examination of test preparation in the Chinese context, the study could not only provide practical recommendations for enhancing IELTS instructional practices, but also contribute valuable insights to broader discussions surrounding the relationship between high stakes testing and teachers' pedagogical approaches.

Keywords: Language Assessment Literacy; IELTS speaking test design test preparation course; teaching methods

Received: 12 October 2023

Accepted: 4 April 2024

Published: 17 October 2025

To cite this article: Yuan, L., Khan, A., & Lah, S. C. (2025). The Impact of the IELTS Speaking Test on Teaching Methods and Test Preparation Strategies in China. *International Journal of Language, Literacy and Translation* 8(2), 1-27. <https://doi.org/10.36777/ijollt2025.8.2.104>

To link to this article: <https://doi.org/10.36777/ijollt2025.8.2.104>

BACKGROUND

Notably, there has been an increase in the number of students who have chosen to pursue their studies in different English-speaking countries, including the United Kingdom, Australia, and New Zealand. One of the requirement to enrol in the course is by achieving the required language proficiency test score. As a result, it is quite common for students from China to enrol in a preparation course in order to adequately prepare for the actual test (Hu & Trenkic, 2021, Ma & Chong, 2022). However, compared to regular classroom teachers, test-preparation classroom instructors' teaching curriculum and activities can be influenced by the target test (Watanabe, 2004). The main aim of cramming school teachers is to assist students in attaining the test scores they desire. However, there have been limited studies in the literature that specifically examine the influence of contextual factors on teachers' perception of language assessment literacy and the potential implications for their instructional practices. For instance, a study conducted by Inbar-Lourie (2017) delved into the relationship being discussed. Taking this into account, the present study specifically examined the test preparation context to explore whether instructors have adequate language assessment literacy. The study aimed to analyse how their ability to conduct assessment-related pedagogical activities might be impacted by this knowledge.

INTRODUCTION

This study aims to examine how Chinese IELTS teachers interpret the test's evaluation goals and how these interpretations influence their teaching methods, thereby reflecting their level of language assessment literacy. In the last decade, many studies have been conducted to explore the importance of language assessment literacy and its role in teachers' pedagogical activities (Jin,2010; Fulcher, 2012; Taylor, 2013). These language assessment literacy studies mostly concentrated on the general feature of language assessment literacy rather than examining such concerns within particular situations, particularly in the context of test preparation. The major reason is language teachers are continually engaged in assessment activities and develop classroom activities based on interpreting the target test design (Berry, 2019). Moreover, the findings of earlier studies have hinted that teachers' interpretation of the target test content can play a larger role in determining their choice of teaching methods in comparison to regular classroom teachers (Wang,2004; Jin, 2010). In addition, many previous studies have explored the importance of language teachers equipping themselves with the appropriate level of assessment literacy (Brindley, 2001; Inbar-Lourie,2008; Taylor, 2013). However, a handful of number of studies focused on Chinese EFL test preparation course context. The major reason to focus on test-preparation course context could be seen as instructors are directly involved in assisting students to get prepared for the tests, especially in high-stakes tests, in which their teaching methods could be affected by the test (Alderson & Wall, 1993). Alderson and Wall (1993) seminal study hypothesised the impact of the target test on teaching, namely the washback on teaching, such as 'A test will influence what teachers' teach' and 'A test will influence how teachers teach. Furthermore, in the specific context of Chinese English as a Foreign Language (EFL), it has been discovered by scholars that a

significant number of Chinese English teachers lack the necessary language assessment literacy skills required for effective marking activities (Qian,2014). Additionally, some teachers faced challenges in successfully integrating assessment knowledge into their assessment practices, as identified by Koh et al. (2018). By considering these factors, the exploration of this issue within this context can offer more data to demonstrate to what extent test-preparation school teachers are equipped with the necessary assessment-related knowledge and how this knowledge is connected with their teaching methods selection.

The following research questions are proposed:

1. To what extent do Chinese IELTS teachers adequately interpret the intended evaluation goals of the IELTS speaking test?
2. What is the relationship between teachers' interpretation of the intended evaluation goals and teaching methods selection?

LITERATURE REVIEW

DEFINITIONS OF LANGUAGE ASSESSMENT LITERACY

The term language assessment literacy has been frequently mentioned in recent years, as this knowledge can be important for different stakeholders. One of the earliest accepted definitions of assessment literacy was made by Stiggins (1991), which defines this term as a person who has enough assessment knowledge to identify the appropriate assessment for students and uses the result to make inferences about students' achievement. However, for language assessment, this definition is rather general, in which the language components have not been included. Furthermore, term is based on general education instead of specific fields, such as language teaching and learning. Consequently, some language assessment experts argued that the linguistic part should be considered in assessment literacy and separate from general assessment literacy. The reasons can be seen as the 'unique complexities that are entailed in the testing and assessment of linguistic skills, knowledge, and communicative competence' (Harding & Kremmel, 2016, p.414; Taylor, 2013). Due to this issue, a collective agreement on the definition of this term has not been reached in this field. As previously mentioned, some key issues are still under discussion, including the variations and extent of knowledge as well as skills for different stakeholders that need to be acquired and how such knowledge could affect their pedagogical activities. However, even though a collective definition of this term has not been reached, many researchers have agreed on some parts of LAL that should be considered by different stakeholders. For example, Davies (2008) argued that language assessment literacy should include three essential parts: skills to develop a test, assessment and language knowledge, and testing principles, such as validity and reliability. Besides, Taylor (2009) held a similar idea that These are the three essential components of language assessment literacy that professionals must fully comprehend. The reason is that this knowledge could help them develop and use the appropriate tests for different purposes.

In the perspective of Fulcher (2012), the researcher argued that language assessment literacy should be organised into various classifications. This includes practical knowledge;

theoretical and procedural knowledge, and socio-historical understanding. Fulcher (2012) argued that practical knowledge is the base and more important than all other aspects of language assessment literacy. In Taylor's (2013) review paper, the author put forward the idea that language assessment literacy can only be achieved through the acquisition of specific levels of knowledge. The author proposed a comprehensive framework comprising eight levels, which are: knowledge of theory, technical skills, principles and concepts, language pedagogy, sociocultural values, local practices, personal beliefs/attitudes, and scores and decision making. Despite Taylor's (2013) caution in labelling it as a model, her suggestion provided a valuable starting point, which ultimately facilitated additional research into the conceptualisation of language assessment literacy. Hence, there is a potentially lack of comprehension regarding the stakeholders' perceived LAL requirements, and the potential variations across various roles, professions, as well as social contexts. To put it differently, the scope of language assessment literacy training ought to be diverse. according to the role of teachers and the context.

TEACHERS' ACQUISITION OF LANGUAGE ASSESSMENT LITERACY

As previously mentioned, the necessity for language teachers to gain language assessment-related knowledge has been emphasised by many researchers, as language teachers are one of the most direct users of different assessment tools. However, some previous studies have found that some teachers did not have adequate assessment knowledge when they were conducting assessment-related activities (Crusana et al., 2016; Nemati et al., 2017). While efforts have been made to incorporate measurement and assessment-related courses into teachers' learning programmes, it is evident that there were still some teachers who struggled with conducting assessment activities due to their limited skills and knowledge in this area (Qian, 2014; DeLuca & Johnson, 2017). Melone (2013) conducted a study within the context of foreign languages in the USA. The researcher found that Language teachers were eager to learn how to use assessment tools, unlike their counterparts, who were language testers, focused on accurately understanding the theoretical aspects of assessment. In addition, Vogt and Tsagari (2014) conducted a study to gauge European foreign language teachers' language assessment literacy and identify their training needs for acquiring such knowledge. The study found that teachers believed they did not have adequate preparation to organise assessment-related activities because of the lack of support or training from their educational program. As a result, teachers had to adopt the assessment approaches of their mentors or colleagues.

However, some researchers may argue that teachers' job is teaching and not assessing and therefore, they are not expected to have adequate knowledge in writing or developing test items (Popham, 2001). The primary job of teachers is indeed teaching students' knowledge or skills. Nevertheless, in many EFL contexts, teachers also play an important role in both summative and formative assessments of students' learning in regular language teaching classrooms (Jin, 2018). Moreover, language teachers are expected to utilise the results of these assessments to inform their teaching effectiveness and monitor students' learning progress. Teachers' assessment integrity and instructional practices may suffer from inadequate assessment knowledge and abilities. For example, teachers may score students' academic performance unfairly and unintentionally because of inadequate assessment literacy (Leirhaug et al., 2016). For teachers, therefore, it is essential to equip adequate language assessment

literacy. Scarino (2013) argued that the discussion of LAL should include teachers' interpretation of the assessment they use. In other words, language teachers' LAL can be shaped by many factors, including the teaching contexts, practices, beliefs and attitudes. For example, Yan et al. (2018) conducted an interview study with three secondary-level experienced Chinese EFL teachers to better understand how contextual and experiential factors could affect their assessment-related training needs. The study found that both factors could affect teachers' motivation for assessment practices. In addition, teachers expressed a great desire for assessment-related training because they felt it would enable them to embrace and modify any additional modifications to their teaching and assessment methods in the local setting. Additionally, the research conducted by Baker and Riches (2018) aimed to examine the characterisation of language assessment literacy among 120 Haitian language teachers. An alternative language assessment literacy aspect was proposed by them as a requirement for language teachers and assessors. During their discussion, they provided further details on the distinctions between language assessment literacy for these two groups. However, their expertise in this area can be seen as complementary when it comes to achieving collaborative tasks.

The previous studies have argued the importance of practitioners getting equipped with adequate language assessment literacy in terms of the professional practices. However, most of these studies are conducted in the top-down approach. The description of the LAL is based on textbook resources (Davies, 2008), language testing and assessment courses (Jeong, 2013) and developed by the researcher (Fulcher, 2012). The components of LAL can range from specific skills, like item writing, to complex knowledge and second language acquisition. However, the focus of LAL for teachers in high-stakes test preparation courses might differ from teachers in other contexts. The reason could be explained as instructors are primarily teaching for the test, and the target test can strongly and directly influence their teaching compared to teachers in regular classrooms (Alderson & Hamp-Lyons, 1996). This means more assessment-related training should be provided to teachers, especially to teachers who directly engaged in teaching activities. This is because they need to be "conversant and competent in the principles and practice of language assessment" (Harding & Kremmel, 2016).

The current study narrows the scope of language assessment literacy as teachers' general assessment knowledge and their interpretation of the intended goals of the IELTS speaking test, including test construct and test criteria. Fulcher and Davidson (2009) argued, 'Test developers normally state target domains of language use, and the range of knowledge, skills or abilities that underpin the test.' As a consequence, it seems very important for instructors to clearly understand what skills the target test aims to assess as well as the standard the examiners used to assess students' performance.'

LAL IN TEST PREPARATION CONTEXT

After the discussion of the different models of language assessment literacy, this study aimed to respond to the call from Taylor (2013) as well as the findings from Kremmel & Harding (2020) to conduct the research by focusing on the contextual factors and personal beliefs of test users. In their respective works, Taylor (2013) and Fulcher (2012) put forward the argument that stakeholders involved in the entire assessment process do not necessarily require an equal

level of language assessment literacy. It is recommended that individuals shift their focus away from unrelated matters and dedicate more attention to the specific aspects of the special activities they are engaged in. Furthermore, a large-scale survey study was conducted by Kremmel and Harding (2020) in which they examined the components that comprise language assessment literacy using a sample of 1086 respondents. The study's findings have shown nine possible separate components of LAL. The components described in this study largely correspond to the components theorised by Taylor (2013). It is important to note that this study presents several significant distinctions and expansions. Taylor (2013) identified a dimension called "Sociocultural values" and another dimension called "Local practices," but these two dimensions should be considered as one combined dimension. The reason behind the intuitive appeal of this idea lies because local contexts are often influenced by sociocultural values, which are known to be context-dependent, and therefore, these values tend to guide policy-making and practices. Furthermore, the study highlighted the concept of "washback and preparation" as a separate dimension, implying that the concerns regarding washback may have relevance for various groups involved. Thus, these results have the potential to serve as a valuable tool for conducting further research, specifically in investigating the variations in LAL needs within different local contexts.

In the Chinese EFL context, it is a common requirement for teachers to hold an undergraduate degree in English or a related field, as well as acquire a teaching certificate, prior to being able to teach English in schools or universities. However, some scholars found that many of them did not have sufficient language assessment literacy skills to do marking activities (Qian, 2014) and some of them had difficulties in incorporating assessment knowledge into their assessment activities (Koh et al., 2018). The reason was that there was a huge variation in many Chinese universities in terms of the requirement for pre-service teachers to take language testing and assessment course as part of the programme (Wang, 2004; Jin, 2010). In view of the test preparation course context in China, even though some have been conducted by collecting teachers' views in terms of how the target test could affect teaching and learning (Yu et al., 2017), teachers' assessment-related knowledge of high-stakes language test in test preparation schools, especially in the speaking test is still under exploration. The focus of this study, therefore, aimed to explore Chinese IELTS teachers' language assessment literacy, where the influence of the test on teachers' teaching actions was expected to be significantly more influential compared to teachers in other educational contexts. In order to provide a more detailed analysis, this study focused on exploring how instructors perceive key principles of the target test, such as the test construct and criteria. Besides, this study planned to examine the potential implications of these interpretations on their instructional methods. The reason for exploring teachers' interpretation of the target test design and this relationship could be explained as teachers' teaching behaviours can be strongly affected in terms of how they teach and the contents they teach (Alderson & Wall, 1993).

METHODOLOGY

The current study used mixed methods, especially a sequential explanatory approach, to answer two research questions. This study aims to explore Chinese IELTS teachers' interpretation of

the intended evaluation goals of the test and its impact on their teaching methods selection. Therefore, the researchers included quantitative data questions (Multiple-choice questions, 5-Likert scale) and qualitative data questions (Open-ended questions) in the questionnaire (Appendix 1) to gather data to answer two research questions.

SAMPLING

The current study used a purposive sampling method to recruit the participants. To be more specific, the researchers aimed to invite experienced IELTS teachers from various institutions across China. This approach ensured that participants had extensive exposure to the teaching of IELTS test preparation course. Thus, the researchers asked the colleagues to send the invitation letter and questionnaire link to the target population. Secondly, the researchers sent the invitation letter to an IELTS teacher WeChat group. There were over 150 IELTS teachers or trainers in this group. The researchers ensured that each participant received the invitation letter and consent form, underscoring the importance of voluntary involvement in this study.

INSTRUMENTS

QUESTIONNAIRE

This study utilised a questionnaire with three parts (see Appendix 1). In the first section, participants' demographic information, teaching experiences, and the size of the classroom they teach in were gathered. The second part was designed to understand teachers' interpretation of the target test construct and criteria. The researchers designed different items based on the IELTS official materials. In this part, multiple-choice questions were designed to check instructors' interpretation of the speaking test construct and a 5-point Likert scale to check their interpretation of test criteria. For multiple-choice questions, if teachers chose the answer matched with the skill that the test aims to assess got two marks, otherwise one mark. The total score of their interpretation of the target test construct was 10 (got correct answers to 5 multiple-choice questions). Besides this, one open-ended question is asked after each multiple-choice question. The final part aimed to understand what teaching methods that teachers often use in their classrooms. The form of questions in this part is based on the observation data. This part used a 5-point-Likert scale to gather the teaching methods that IELTS speaking test teachers frequently use in their classrooms.

PILOT STUDY

The questionnaire was piloted before conducting the study in terms of its validity and reliability. Regarding the validity, the researchers checked the face validity and content validity of the instrument. In terms of the face validity, the researchers asked 5 IELTS teachers who have background knowledge of Applied linguistics and language testing and assessment to ask their thoughts in terms of the validity of the questionnaire. In order to provide more detailed information, the researchers have extended invitations to five potential participants. These participants were asked to use the think-aloud approach to share their thoughts and reflections

regarding the confusions and issues they encountered while reading each item of the questionnaire. Along with assessing face validity, the content validity was ensured by inviting one Ph.D. candidate and one Ph.D. holder in language testing and assessment. Their role was to review the survey and confirm that the questions adequately address all aspects of the construct being measured. The researchers revised some wording of the questions to make them more concise and clearer to be read and understood. For the reliability, the researchers checked the Cronbach alpha level of the questionnaire in SPSS. The result showed the overall Cronbach alpha level of the instrument was $\alpha=0.75$ and this was acceptable, according to the literature (Brown, 2001). Besides, the researchers administrated the survey with 5 participants twice in two consecutive weeks and checked the agreement of answers by checking the correlation by using the Analysis ToolPak add-in within Excel. The reason for employing this ToolPak add-in within Excel stems from its ability to offer a straightforward and user-friendly interface in Excel and offers a range of other statistical functions, such as correlation analysis, regression analysis, descriptive statistics, and hypothesis testing. This ToolPak add-in is equipped with the functionality to swiftly compute correlation coefficients and provide immediate results. The correlation analysis results showed that $p<0.05$, $r= 0.38$, in which these two sets of answers were correlated, and this proved that the questionnaire was reliable. After the evaluation of the validity and reliability of the instrument, the questionnaire was employed for the main study.

NON-PARTICIPANT CLASSROOM OBSERVATION

Given the ongoing pandemic, the study had to be conducted with specific limitations, which led to researchers having limited opportunities to observe classes in person. Subsequently, the researcher opted to observe the online IELTS speaking classes of five teachers, doing so three times over the course of two weeks. The classroom size was considered as small, as there were only 5 - 7 students in each class. It is worth mentioning that the participants had a language proficiency level ranging from 5 to 5.5. They expressed a strong desire to enrol in the test preparation course to achieve a Band 6 or higher in the speaking test. In order to be eligible for degree programs in English-speaking countries like the UK, Australia, and New Zealand, it was mandatory for them to achieve a minimum Band 6 in each component of the IELTS test, which was one of the admission requirements (Who accepts IELTS, 2024).

The observation made by the researchers showed that every teacher had covered all three parts of the IELTS speaking test in their instruction. The researchers conducted nonparticipant classroom observation in order to maintain objectivity. During these observations, the researcher intentionally did not participate in any teaching or learning activities and solely observed the entire class. The researchers noted the teaching methods instructors chose to teach their students in their classes. After collecting the observation data, the researcher used this data to form questions in the third part (teaching methods selection) of the questionnaire. The researcher has summarised five major teaching methods that teachers employed in the classes, including (1) asking students to practise making points and provide support; (2) teaching students to use a wide range of vocabulary, and (3) asking students to recite some high mark answers, (4) using question cards to help students to prepare for the test; (5) practising pronunciation and giving feedback in the class.

PARTICIPANTS

The researchers distributed the information of the current study in a WeChat group of over 100 IELTS speaking test teachers who worked in various language education institutions across China which also included an invitation letter and a consent form. Consequently, the researchers could collect data from a considerable number of teachers, specifically 34 in total. It is important to mention, however, one teacher chose not to participate in the study, while another initially agreed to be part of it but later decided to withdraw. As a result, the final sample size consisted of 32 valid cases, with 10 male teachers and 22 female teachers being part of it. Table 1 displayed the demographic information of the participants.

The researchers identified three types of classroom sizes for IELTS speaking test preparation courses. The first one was a small-sized classroom, in which teachers were doing one-to-one personal tutoring or teaching a classroom with no more than 10 students. According to the data, 15 teachers were teaching in small-sized classrooms and 9 teachers were doing personal tutoring. The second type of classroom size was medium size, in which teachers taught students below 25 students in a classroom and there were 4 teachers informed they taught medium-sized classrooms. The final type was large-sized classrooms, in which there were over 25 students in a classroom and 4 teachers were teaching large-size classrooms. In addition, teachers' teaching experiences varied from less than a year to more than three years. The demographic data showed that only five teachers had taught IELTS speaking for less than a year, 13 teachers had at least 1-year teaching experience and 14 teachers have taught IELTS speaking for over three years. In terms of each IELTS speaking class's length, the majority teach each IELTS speaking class for at least 1.5 hours. Specifically, 14 teachers taught each class 2 hours and 6 teachers taught each class 1.5 hours 6 teachers taught each class of 2 hours and six teachers taught each class 1 hour in total.

Table 1
Demographic Information

Classroom Size	
	Number of teachers
Personal tutoring	9
Small	15
Medium	4
Large	4
Total	32
Years of teaching experience	
Less than a year	5
More than a year	13
More than 3 years	14
Total	32
Classroom time	
1 hour	6
1.5 hour	6

2 hours	14
More than 2 hours	6
Total	32

DATA ANALYSIS

TEACHERS' INTERPRETATION OF IELTS SPEAKING TEST DESIGN

To address RQ1 (1. To what extent do Chinese IELTS teachers adequately interpret the intended evaluation goals of the IELTS speaking test?), the quantitative data of the second part of the questionnaire was first analysed descriptively, using SPSS 24. In terms of the open-ended questions data, the researchers manually coded the data for content analysis (Cabrera & Reiner, 2018).

THE IMPACT ON TEACHING METHODS SELECTION

The third part of the questionnaire's questions was analysed by using descriptive and inferential statistics to explore the second research question (RQ2: What is the relationship between teachers' interpretation of the intended evaluation goals of the test and teaching methods selection?). The researchers used Spearman Correlation to explore the connection between IELTS teachers' teaching methods selection with their interpretation of the target test design.

FINDINGS

TEACHERS' INTERPRETATION OF IELTS SPEAKING TEST DESIGN

QUANTITATIVE DATA

After analysing the quantitative data, these findings were identified in terms of Chinese IELTS teachers' interpretation of the IELTS speaking test.

Q1: To what extent do Chinese IELTS teachers adequately interpret the intended evaluation goals of the IELTS speaking test?

Table 2

Teachers' interpretation of intended goals of the IELTS speaking test

	Mean	SD	N
Test construct	7.78	1.408	32
Test criteria	17.03	3.277	32
intended evaluation goals	24.81	3.496	32

In the current study, the intended evaluation goals of the IELTS speaking test include two parts, test construct and test criteria. The researchers decided to use the parcelling strategy to combine

the data of the test construct and test criteria as the intended evaluation goals. The reasons could be explained as these two parts constitute the intended evaluation goals of the test and the current study used teachers' interpretation of the intended evaluation goals as the independent variable to explore its impact on teaching methods. The combination of the items can reduce the impact of measurement error associated with individual items (Little et al., 2002). The researchers used SPSS 24 to analyse the data and demonstrated descriptive statistics of IELTS teachers' interpretation of test construct and test criteria in table 2. The data showed that from the sole interpretation of the mean score, the participants may have an adequate interpretation of the target test design, including test construct, Mean=7.78/10 and test criteria, Mean=17.03/20. However, the standard deviation of test criteria, SD=3.277 was huge. The presence of a large standard deviation suggested that there was a significant amount of variability in the observed data, with values spread out widely around the mean. This may indicate that teachers may hold their own understanding in terms of test criteria. In other words, some teachers did not recognise the standard the examiners used in the test that was similar to what they believed.

QUALITATIVE DATA

One open-ended question in the questionnaire followed the multiple-choice questions to ask teachers what other ability they think the IELTS speaking test assesses, and some teachers offered some ideas. The researchers used thematic analysis to decode teachers' responses regarding their thoughts on the IELTS speaking test's intended evaluation goals. In addition, all the open-ended responses were related to each part of the IELTS speaking test and then this information was taken as an approach to develop the code to analyse the data. To put it another way, the code scheme utilised in this study was developed by drawing from the structure of the IELTS speaking test. Participants were also requested to provide answers they believed would showcase the abilities in the different areas assessed in each part of the test.

Table 3

What other ability do you think IELTS speaking part 1 assesses?

Code	Examples
Communication	daily communication; basic communication skills; to express feelings of certain situations
Authenticity	Real talk; authentic language use; authentic language
Listening	Listening skills, understand questions, understanding of different questions and relevance of your answers.

IELTS speaking test part 1 claims to focus on 'the ability to communicate opinions and information on everyday topics and common experiences or situations by answering a range of questions (IELTS, 2020).' After coding participants' responses, there are overlaps between the language ability test developers plan to assess and teaching practitioners' beliefs. For example, one of the mentioned language abilities is communication skill. Teachers believe IELTS speaking part 1 evaluates test-takers' ability to understand different everyday topics and daily communication. In addition, teachers have mentioned listening skill and vocabulary. Although

the IELTS speaking test does not assess test-takers' listening skill and vocabulary explicitly, test-takers still need these skills to complete different communicative tasks successfully in the speaking test (IELTS guide for teachers, 2019).

Table 4

What other ability do you think IELTS speaking part 2 assesses?

Code	Examples
Logic and Organisation	Telling a 2-minute story with proper signposting to guide and attract your audience's attention; Organisational and logical speaking; the ability to give a detailed story
Coherence	To identify the main point of a given topic and concentrate on it; relevance; sorting of information; the sequence of information delivery

IELTS speaking test part 2 claims 'This part of the test focuses on the ability to speak at length on a given topic (without further prompts from the examiner), using appropriate language and organising ideas coherently. The test-takers will probably need to draw on their own experience to complete the long turn (IELTS, 2020).' After coding the responses, the participants' comments could be categorised as two main ideas, 'logic and organisation'; 'Coherence'. These responses, in general, match what the IELTS speaking test aims to assess in this part.

Table 5

What other ability do you think IELTS speaking part 3 assesses?

Code	Examples
Discussion	Debating; the ability to debate; the ability to defend and support your opinion; discussion abilities; discussion skills; critical thinking (*3)
Communication	Deep communication; the ability to give longish answers towards complex questions; deepen ideas about the abstract and complex questions.

According to the information on the IELTS website, part 3 test focuses on test-takers 'ability to express and justify opinions and to analyse, discuss and speculate about the issue (IELTS, 2020).' The participants' entries mentioned discussion ability, which they refer it as debating, defending and supporting one's opinion. This ability has been explicitly addressed in the IELTS test construct. Additionally, teachers believed that communication skills will be assessed in this part. However, this part's communication skill is different from the one in part 1. They believed that this part assesses test-takers' ability to answer different complex and abstract questions and give more extended answers.

Besides these two skills, in this part, the participants mentioned critical thinking skill. According to the test design information, this term has not been explicitly written in the test construct. Nevertheless, the definition of this term may overlap with some parts of the test

construct, such as expressing and justifying opinions. In addition, a popular definition of critical thinking refers to individuals' ability to be responsible for one's thinking and understand and evaluate different perspectives and solve problems (Maiorana,1992; Elder & Paul,1994). In the current study, this could be viewed as some IELTS speaking test teachers believed the target test assesses test-takers' critical thinking skill, in particular to focus on how test-takers' understand the target language and think critically and creatively of different questions (Kabilan, 2000; Mahyuddin et al.,2004).

In summary, after decoding participants' responses to quantitative data and three open-ended questions, it has been demonstrated that the skills the IELTS speaking exam seeks to evaluate are in line with how teachers interpret the target test design. However, some differences have also been noticed, especially in the use of authentic language in part 1 and having critical thinking in part 3. According to the IELTS test design, these skills are not explicitly written in the construct as these skills are not language skills. However, in the real test, test-takers may need to acquire these skills to finish different tasks.

THE IMPACT ON TEACHING METHODS

Q2: What is the relationship between teachers' interpretation of the intended evaluation goals and teaching methods selection?

There are five teaching methods that teachers employed in their classroom, including Method 1: asking students to practise making points and provide support; Method 2: teaching students to use a wide range of vocabulary; Method 3: asking students to recite some high mark answers; Method 4: using questions cards to help students to prepare for the test; Method 5: practising pronunciation and giving feedback in the class.

Table 6

Teaching methods selection

	N	Mean	SD	Sig.
Method 1	32	4.25	1.016	.000*
Method 2	32	4.09	.743	.175
Method 3	32	2.84	1.194	.964
Method 4	32	4.06	1.076	.002*
Method 5	32	3.97	.822	.089

The Spearman's rho correlation coefficient was used to explore the relationship between teachers' interpretation of the intended evaluation goals of the test and teachers' choice of teaching methods. The researchers found a medium-strong correlation between teachers' interpretation of the IELTS speaking test design and the two teaching methods. The first one is method 1, asking students to practise making points and provide supporting, $R_s(32) = .500, p < 0.05$. The second one is Method 4, using questions cards to help students to prepare for the test, $R_s(32) = .584, p < 0.001$.

It is not surprising that there was a correlation between teachers' interpretation of test design and method 1, teaching students to practise making points, and providing supporting

details and method 4, the use of question cards to help students prepare for the test. Firstly, this result could be explained by the findings from the qualitative data of teachers' interpretation of test design. It has indicated that teachers believed that the IELTS speaking test assesses students' abilities to use critical thinking skills to organise sentences and ideas to answer different questions in the test. Moreover, they also believed that the IELTS speaking test aims to assess students' ability to make a logical speech and defend one's opinion in part 3. In addition, the data from teachers' choice of teaching methods also support this idea as teachers stated they often chose to use question cards to help students get familiar with different topics and help students develop ideas of different questions. Therefore, the correlation between the two methods teachers often used in the classroom and their interpretation of what abilities the test aims to assess in part 1 and part 3 could be explained by the qualitative and quantitative data.

In spite of this, it was surprising that teachers did not favour method 3 (where students are asked to recite high-scoring answers) as one of the commonly employed strategies in their classrooms. In practice, it is not uncommon for Chinese IELTS test-takers to rely on memorised materials in order to complete the speaking tasks effectively, with part 2 being the area of focus. In addition, during the researchers' observation of classrooms, teachers presented model answers of different speaking test questions during the teaching, especially in part 2, which required test-takers to deliver a 1–2-minute monologue in the test. Additionally, some teachers asked students to memorise these answers after class by themselves. One reason to explain this mismatch could be that teachers did not think this should be considered as a particular teaching method, in which passive learning is still very popular in the Chinese EFL context (Wang & Ryan, 2020)

In terms of method 2 and method 5, these two instructions primarily aimed to help students develop their linguistic competence. The researchers discovered that some teachers taught vocabulary based on themes and paid attention to students' pronunciation during the observation period in the classroom. However, in the questionnaire, teachers did not think they have frequently used these methods in their classes. One explanation would be that they didn't employ these teaching strategies independently in their classes since they might take more time and be less successful than other approaches at both educating and assisting students in achieving their desired grades. To rephrase, teachers would prioritise helping students get ready for the test by employing the most effective approaches and strategies, rather than solely focusing on their language development, which may require a longer timeframe.

In conclusion, the examination of the quantitative data led to the identification of a correlation between teachers' understanding of the target test design and their use of two specific teaching approaches. The main objective of these two teaching methods was to enhance students' test-taking strategies, rather than focusing on their linguistic competence, including vocabulary and pronunciation.

DISCUSSION

This study employed a mixed mode questionnaire to collect Chinese IELTS teachers' language assessment literacy. To be more specific, the current study narrows the focus to the test-

preparation context, especially to teachers' acquisition of some key principles of designing a test, such as test construct and test criteria, and how this could affect their teaching methods. According to the quantitative data of teachers' interpretation of intended evaluation goals, the current study found that Chinese IELTS speaking test teachers had a relatively adequate interpretation of the intended evaluation goals of the test, including test construct and test criteria. However, the standard deviation of teachers' interpretation of test criteria was relatively huge, which was 3.277. This data could be interpreted as some teachers may hold a very different idea from the standard that the examiners used to assess students' performance in the speaking test. Teachers may hold their own beliefs in terms of the test criteria would be employed to assess test-takers' performance. This could be interpreted as a sign that IELTS instructors require additional materials and instruction regarding the structure and criteria of the IELTS speaking test. According to several researchers in the literature, teaching practitioners do not need to acquire the same level of language assessment literacy as other test takers, and instructors' needs for this literacy vary depending on the environment. (Taylor, 2013; Davidson, 2004; Vogt & Tsagari, 2014; Kremmel & Harding, 2020). Their argument centred on the idea that content instruction should be localised, focused on specific subject areas, and provide teachers with the autonomy to make their own choices (Lam, 2015b; Leahy & Wiliam, 2012). Despite the extensive research conducted on training, teachers continue to face challenges in acquiring assessment knowledge as they perceive it to be theoretical and disconnected from their everyday classroom assessment practices (Popham, 2009; Yan et al., 2018). Leung (2014) argued that the absence of contextualisation in the learning process is a contributing factor, as learners are often exposed to related assessment knowledge in a one-size-fits-all manner. Therefore, this could be argued that more localised and context specific assessment-related activities in-service training should be provided to test-preparation school teachers to understand the fundamental principles of language testing and assessment as well as the specific information of the target test that they are teaching.

Regarding the qualitative comments in terms of the abilities that they believed the test evaluates, the researcher found that teachers mentioned some skills that are not explicitly written in the test construct. The first ability was the usage of authentic language to answer different questions. According to the IELTS official website, the test claims it is a communicative test, but the authenticity of language use has not been explicitly written as a part of the test construct. However, in the public version of test criteria, the use of idiomatic expressions has been mentioned in Band 7 and above in the Lexical resource section. For band 6 and lower, there is no requirement for using idiomatic language. However, from raters' perspective, the definition of authentic language use does limit to not only the use of vocabulary but also the produced discourses in the real test, for example, the good task response of different questions, the approaches to deal with breakdowns and self-correction (Burton, 2020). Some teachers mentioned that the part 1 evaluates students' ability to use authentic English to answer different questions, such as having 'real talk'; 'authentic language use'. In the part 3, they believed that students should give longer answers to each question and use critical thinking skills to discuss with examiners. In other words, teachers believed that students should be able to have the ability to use the language resources to get the meaning across in different contexts. These phrases could be seen as teachers believed that students need to use authentic language to complete different tasks. Some studies argued that test-preparation school classes are more

score oriented and narrow the width and depth of test-preparation activities, which the course primarily led to negative washback on learners as it narrowed down the scope of learning by focusing on cramming for the test and memorising materials and then affecting the authenticity of the produced discourses in the test (Luk, 2010; Lam, 2015a; Wink & Lim, 2017). Burton (2020) conducted a study to explore the authenticity of eight IELTS speaking tests samples from 58 raters' perspective. According to their value systems and perceptions of improper preparation techniques, examiners were able to identify instances of inauthentic language use, according to the study. Furthermore, this study discovered that certain raters were able to determine if test-takers had been exposed to test prompts and materials by evaluating the genuineness of their spoken discourses. Additionally, Ellis (2005) and Lam (2015a) found that the test preparation activities could affect test-takers' test spoken utterances in the real test, which raters could easily observe. These studies found test-takers' inauthentic language use in the test had a connection with the test preparation course and course contents they used to prepare for the test. However, this claim is not conclusive, which the impact of test preparation activities on test-takers' use of inauthentic language in the real IELTS speaking test is still not clear. This is because the contents of test preparation classes and the teaching methods teachers choose may vary according to the assessment-related training and knowledge they have had and thus affecting their teaching methods selection (DeLuca & Johnson, 2017; Melone, 2013).

This means for test preparation contexts, even though there are some issues, such as narrowing the teaching curriculum, this does not mean all cramming schools should be closed entirely as it will not change anything. Instead, different institutions need to provide in-service training of language assessment-related knowledge to teachers who are directly using and teaching to the test, such as key principles of assessment and the target test design, thus promoting positive washback on teaching and learning.

Besides the usage of authentic languages, some teachers argued that the evaluation of critical thinking abilities could be found from the test. In terms of critical thinking, some researchers argued that critical thinking skills closely connect with one's listening and speaking skills (Davidson & Dunham, 1997; Yang et al., 2013). However, a handful of studies have been conducted to explore how critical thinking could connect with language testing. Even though the IELTS speaking test has included some features of critical thinking in the test criteria, which students should be able to justify their own ideas. Nonetheless, for critical thinking this ability, it has not been written as an explicit construct in any particular high-stakes test. For example, in the IELTS speaking test, only the 'Fluency and Coherence' section mentioned some skills related to critical thinking, such as using cohesive markers in the speech and developing topics coherently and appropriately. Therefore, some teachers may interpret test criteria differently than others, which could be partially explained by the official materials' poor display of critical thinking.

All in all, the current study discussed the necessity for IELTS language teachers to get equipped with assessment-related knowledge or language assessment literacy by paying more attention to the test they are teaching. There is no doubt that the goal of the test preparation course is to support students to get higher score. Helping instructors acquire strong assessment-related knowledge, however, could be one method to accomplish this goal. This information could help teachers select instructional activities that are appropriate for a certain test or situation. By doing so, teachers could be able to capture the aims of the target test and then

select or modify their pedagogical practices based on such knowledge. Besides Chinese EFL context, the findings of this study could be applicable to other EFL contexts, which test preparation culture is also very popular, such as Japan and South Korea.

CONCLUSION AND RECOMMENDATION

There is no doubt that this study has some issues that need to be resolved by further studies in order to get better insight into teachers' language assessment literacy. Firstly, more methods that should be employed to explore the relationship between instructors' language assessment literacy and its impact on their teaching methods selection. The questionnaire was the only tool utilised in this study to gather information for the research topics. Despite included open-ended questions in the questionnaire, the researchers were unable to adequately explain how teachers interpreted the target test design. For example, some teachers believed the IELTS speaking test aims to evaluate students' abilities to use authentic language and have critical thinking to form answers. The inclusion of more qualitative explanatory studies, such as semi-structured interviews or focus groups, could offer a more comprehensive understanding of this issue.

In addition, the scope of language assessment literacy in this study was limited to the specific skills of the test and the context. There is no doubt this exploration could contribute a better insight into how the equipment of language assessment literacy could be varied according to the role of teachers in the test preparation course context. However, the narrow focus in this study may not provide a broader view in terms of how the needs of language assessment literacy would be the same or not for all teachers across different courses or contexts and how this knowledge could have a connection with their teaching practices. Therefore, several quantitative techniques with high sample sizes and multi-variant analysis, including factor analysis and structural equation modelling, can be used in future research to investigate this topic.. By conducting studies in this manner, the results may have a higher level of generalizability to different contexts or could serve as a basis for future research on similar issues in other tests.

In conclusion, the results revealed that the importance for teaching practitioners in test preparation schools to get equipped with assessment related knowledge and how this knowledge might affect their teaching. The findings of the present study suggested, in line with previous LAL studies, that the in-service training of language assessment should be provided to teachers, especially teachers who are directly engaged in teaching (Taylor, 2013; Vogt & Tsagari, 2020). By enhancing teachers' assessment-related knowledge, there is a potential for a significant impact on their teaching practices. This, in turn, can be directly related to the target test's intended evaluation goal, which eventually causes their teaching methods to concentrate on the abilities the test seeks to evaluate—specifically, the positive washback effect on instruction.

ACKNOWLEDGEMENT

The researchers would like to express our appreciation to all the Chinese IELTS teachers who participated in this study.

REFERENCE

- Alderson, J. C., & Hamp-Lyons, L. (1996). TOEFL preparation courses: A study of washback. *Language Testing*, 13(3), 280–297. <https://doi.org/10.1177/026553229601300304>
- Berry, V., Sheehan, S., & Munro, S. (2019). What does language assessment literacy mean to teachers? *ELT Journal*, 73(2), 113–123. <https://doi.org/10.1093/elt/ccy055>
- Burton, J. D. (2020). "How scripted is this going to be?" Raters' views of authenticity in speaking-performance tests. *Language Assessment Quarterly*, 17(3), 244–261. <https://doi.org/10.1080/15434303.2020.1754829>
- Brown, J. D. (2001). *Using surveys in language programs*. Cambridge University Press.
- Cabrera, L. Y., & Reiner, P. B. (2018). A novel sequential mixed-method technique for contrastive analysis of unscripted qualitative data: Contrastive quantitized content analysis. *Sociological Methods & Research*, 47(3), 532–548. <https://doi.org/10.1177/0049124116661575>
- Davidson, B. W., & Dunham, R. A. (1997). Assessing EFL student progress in critical thinking with the Ennis-Weir Critical Thinking Essay Test. *JALT Journal*, 19(1), 43–57.
- Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing*, 25(3), 327–347. <https://doi.org/10.1177/0265532208090156>
- DeLuca, C., & Johnson, S. (2017). Developing assessment capable teachers in this age of accountability. *Assessment in Education: Principles, Policy and Practice*, 24(2), 121–126. <https://doi.org/10.1080/0969594X.2017.1297010>
- Elder, L., & Paul, R. (1994). Critical thinking: Why we must transform our teaching. *Journal of Developmental Education*, 18(1), 34–35.
- Ellis, R. (Ed.). (2005). *Planning and task performance in a second language* (Vol. 11). John Benjamins Publishing. <https://doi.org/10.1075/llt.11>
- Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9(2), 113–132. <https://doi.org/10.1080/15434303.2011.642041>
- Fulcher, G., & Davidson, F. (2009). Test architecture, test retrofit. *Language Testing*, 26(1), 123–144. <https://doi.org/10.1177/0265532208097339>
- Harding, L., & Kremmel, B. (2016). Teacher assessment literacy and professional development. In D. Tsagari & J. Banerjee (Eds.), *Handbook of second language assessment* (pp. 413–428). De Gruyter Mouton. <https://doi.org/10.1515/9781614513827-027>
- Hu, R., & Trenkic, D. (2021). The effects of coaching and repeated test-taking on Chinese candidates' IELTS scores, their English proficiency, and subsequent academic achievement. *International Journal of Bilingual Education and Bilingualism*, 24(10), 1486–1501. <https://doi.org/10.1080/13670050.2018.1535620>
- IELTS guide for teachers. (2019). *IELTS*. <https://www.ielts.org/-/media/publications/guide-for-teachers/ielts-guide-for-teachers-uk.ashx>
- Ingram, E. (1968). Attainment and diagnostic testing. In A. Davies (Ed.), *Language Testing Symposium: A Psycholinguistic Approach* (pp. 147–159). Oxford University Press.

- <https://doi.org/10.1111/j.1467-1770.1968.tb00228.x>
- Jin, Y. (2010). The place of language testing and assessment in the professional preparation of foreign language teachers in China. *Language Testing*, 27(4), 555–584. <https://doi.org/10.1177/0265532209351431>
- Jin, Y. (2018). Defining and developing the assessment literacy of foreign language teachers. *Foreign Language Education in China*, 1(2), 65–72.
- Kabilan, K. M. (2000). Creative and critical thinking in language classrooms. *The Internet TESL Journal*, 6(6). <http://iteslj.org/Techniques/Kabilan-CriticalThinking.html>
- Kremmel, B., & Harding, L. (2020). Towards a comprehensive, empirical model of language assessment literacy across stakeholder groups: Developing the language assessment literacy survey. *Language Assessment Quarterly*, 17(1), 100–120. <https://doi.org/10.1080/15434303.2019.1674855>
- Lam, D. (2015a). Contriving authentic interaction: Task implementation and engagement in school-based speaking assessment in Hong Kong. In G. Yu & Y. Jin (Eds.), *Assessing Chinese learners of English: Language constructs, consequences and conundrums* (pp. 38–60). Palgrave Macmillan. https://doi.org/10.1057/9781137449788_3
- Lam, R. (2015b). Language assessment training in Hong Kong: Implications for language assessment literacy. *Language Testing*, 32(2), 169–197. <https://doi.org/10.1177/0265532214554321>
- Leirhaug, P. E., MacPhail, A., & Annerstedt, C. (2016). "The grade alone provides no learning": Investigating assessment literacy among Norwegian physical education teachers. *Asia-Pacific Journal of Health, Sport and Physical Education*, 7(1), 21–36. <https://doi.org/10.1080/18377122.2016.1145429>
- Leung, C. (2013). Classroom-based assessment issues for language teacher education. In A. J. Kunnan (Ed.), *The Companion to Language Assessment* (Vol. 3, pp. 1510–1519). Wiley-Blackwell. <https://doi.org/10.1002/9781118411360.wbcla116>
- Levy-Vered, A., & Alhija, F. N.-A. (2015). Modeling beginning teachers' assessment literacy: The contribution of training, self-efficacy, and conceptions of assessment. *Educational Research and Evaluation*, 21 (5–6), 378–406. <https://doi.org/10.1080/13803611.2015.1117980>
- Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling*, 9(2), 151–173. https://doi.org/10.1207/S15328007SEM0902_1
- Luk, J. (2010). Talking to score: Impression management in L2 oral assessment and the co-construction of a test discourse genre. *Language Assessment Quarterly*, 7(1), 25–53. <https://doi.org/10.1080/15434300903473997>
- Mahyuddin, R., Lope Pihie, Z. A., Elias, H., & Konting, M. M. (2004). The incorporation of thinking skills in the school curriculum. *Kajian Malaysia*, 22(2), 23–33.
- Maiorana, V. P. (1992). *Critical thinking across the curriculum: Building the analytical classroom*. ERIC Clearinghouse on Reading and Communication Skills.
- Malone, M. (2013). The essentials of assessment literacy: Contrasts between testers and users. *Language Testing*, 30(3), 329–344. <https://doi.org/10.1177/0265532213480129>
- Ma, H., & Chong, S. W. (2022). Predictability of IELTS in a high-stakes context: A mixed methods study of Chinese students' perspectives on test preparation. *Language Testing*

- in Asia, 12, 1–18. <https://doi.org/10.1186/s40468-021-00152-3>
- Nemati, M., Alavi, S. M., Mohebbi, H., & Masjedlou, A. P. (2017). Teachers' writing proficiency and assessment ability: The missing link in teachers' written corrective feedback practice in an Iranian EFL context. *Language Testing in Asia*, 7, 1–18. <https://doi.org/10.1186/s40468-017-0053-0>
- Qian, D. D. (2014). School-based English language assessment as a high-stakes examination component in Hong Kong: Insights of frontline assessors. *Assessment in Education: Principles, Policy & Practice*, 21(3), 251–270. <https://doi.org/10.1080/0969594X.2014.915207>
- Scarino, A. (2013). Language assessment literacy as self-awareness: Understanding the role of interpretation in assessment and in teacher learning. *Language Testing*, 30(3), 309–327. <https://doi.org/10.1177/0265532213480128>
- Stiggins, R. (1995). Assessment literacy for the 21st century. *Phi Delta Kappan*, 77(3), 238–245.
- Taylor, L. (2013). Communicating the theory, practice and principles of language testing to test stakeholders: Some reflections. *Language Testing*, 30(3), 403–412. <https://doi.org/10.1177/0265532213480338>
- Vogt, K., & Tsigari, D. (2014). Assessment literacy of foreign language teachers: Findings of a European study. *Language Assessment Quarterly*, 11(4), 374–402. <https://doi.org/10.1080/15434303.2014.960046>
- Vogt, K., Tsigari, D., & Spanoudis, G. (2020). What do teachers think they want? A comparative study of in-service language teachers' beliefs on LAL training needs. *Language Assessment Quarterly*, 17(4), 386–409. <https://doi.org/10.1080/15434303.2020.1781128>
- Wang, Y., & Ryan, J. (2020). The complexity of control shift for learner autonomy: A mixed-method case study of Chinese EFL teachers' practice and cognition. *Language Teaching Research*. <https://doi.org/10.1177/1362168820957922>
- Winke, P., & Lim, H. (2017). The effects of test preparation on second-language listening test performance. *Language Assessment Quarterly*, 14(4), 380–397. <https://doi.org/10.1080/15434303.2017.1399396>
- Who accepts IELTS? (2024, February 9). *IELTS*. <https://ielts.org/take-a-test/why-choose-ielts/who-accepts-ielts>
- Yang, Y. T. C., Chuang, Y. C., Li, L. Y., & Tseng, S. S. (2013). A blended learning environment for individualized English listening and speaking integrating critical thinking. *Computers & Education*, 63, 285–305. <https://doi.org/10.1016/j.compedu.2012.12.012>
- Yan, X., Zhang, C., & Fan, J. J. (2018). "Assessment knowledge is important, but...": How contextual and experiential factors mediate assessment practice and training needs of language teachers. *System*, 74, 158–168. <https://doi.org/10.1016/j.system.2018.03.003>
- Yu, G., He, L., Rae-Dickens, P., Kiely, R., Lu, Y., Zhang, J., & Zhang, Y. (2017). Preparing for the speaking tasks of the TOEFL iBT® test: An investigation of the journeys of Chinese test takers. *TOEFL iBT Research Report*, 28. <https://doi.org/10.1002/ets2.12145>

APPENDIX

Appendix 1 Consent form

Consent form

Dear Participant,

This current project is an investigation of the impact of IELTS on Chinese IELTS speaking test teachers' teaching methods selection. The aim of the project is to gain an understanding of IELTS teachers' knowledge about IELTS speaking test, and the relationship this knowledge has with their classroom teaching. Your responses to this questionnaire will be treated in confidence and only used for the stated purposes of the study.

If you have any questions about this project, please email me: yuan.liu7@outlook.com

Thank you very much for your time and cooperation. I should also be grateful if you would tick the consent option below.

Kind Regards,

Liu Yuan (Quentin)

1. Your consent to participate in the Impact Study

I understand that:

The purpose of the study is to collect and analyse information from those familiar with international English language tests;

My name will not appear in any project publication;

The information I give, but not my name, may be quoted;

I am free to refuse to participate in the study and may withdraw at any time;

My completed questionnaire is for the study team only; it will not be shown to anyone not connected with the study

☐ Yes, I understand it.

☐ No, I do not understand.

Appendix 2 Questionnaire

1. General Background

1. What is your gender?

- ☐A. Male
- ☐B. Female
- ☐C. Prefer not to say

2. What size of IELTS speaking classroom are you currently or recently teaching?

- ☐A. Personal tutoring (1 to 1)
- ☐B. small size classroom (below 10 people)
- ☐C. medium size classroom (below 25 people)
- ☐D. large size classroom (above 25 people)

3. How long have you been teaching IELTS speaking test?

- ☐A. Less than a year
- ☐B. 1-3 years
- ☐C. Over 3 years

4. How long is each IELTS speaking class?

- ☐A. 1 hour in total
- ☐B. 1.5 hours in total
- ☐C. 2 hours in total
- ☐D. More than 2 hours

2. Interpretation of the IELTS speaking test construct

5. What language ability do you think IELTS speaking part 1 assesses?

Speaking sample task – Part 1

Part 1 Introduction and interview

[This part of the test begins with the examiner introducing himself or herself and checking the candidate's identification. It then continues as an interview.]

Let's talk about your home town or village.

- What kind of place is it?
- What's the most interesting part of your town/village?
- What kind of jobs do the people in your town/village do?
- Would you say it's a good place to live? (Why?)

Let's move on to talk about accommodation.

- Tell me about the kind of accommodation you live in?
- How long have you lived there?
- What do you like about living there?
- What sort of accommodation would you most like to live in?

- A. to give a long and detailed answer to each question on everyday topics and common experiences or situations.
- B. to give a short and quick response to each question on everyday topics and common experiences or situations.
- C. to communicate opinions and information of each question on everyday topics and common experiences or situations.

6. What other ability do you think IELTS speaking test part 1 assesses?

7. What language ability do you think IELTS speaking part 2 assesses?

Part 2 – Individual long turn

Candidate Task Card

Describe something you own which is very important to you.

You should say:

**where you got it from
how long you have had it
what you use it for**

and explain why it is important to you.

You will have to talk about the topic for 1 to 2 minutes.
You have one minute to think about what you're going to say.
You can make some notes to help you if you wish.

- A. to integrate the memorised materials to give a clear and fluent speech for 1-2 minutes.
- B. to organise ideas coherently and give a fluent speech on the given topic for 1-2 minutes.
- C. to give a fluent and clear 1-2 minutes speech that is related to the theme of the given topic.

8. What other ability do you think IELTS speaking test part 2 assesses?

9. What language ability do you think IELTS speaking part 3 assesses?

Speaking sample task – Part 3

Part 3 – Two-way discussion

Let's consider first of all how people's values have changed.

- What kind of things give status to people in your country?
- Have things changed since your parents' time?

Finally, let's talk about the role of advertising.

- Do you think advertising influences what people buy?

- A. to analyse, discuss and speculate about issues of each question.
- B. to have a discussion with the examiner about issues of each question.
- C. to give innovative ideas or good solutions about issues of each question.

10. What other ability do you think IELTS speaking test part 3 assesses?

11. What vocabulary and grammar knowledge do you think you should teach to help students to answer different types of questions in the whole speaking test?

- A. to teach them to use accurate, appropriate and range of vocabulary and grammar to form all answers.
- B. to teach some complicated vocabulary and complex grammar structures that they can use to form all answers.
- C. to teach some high-frequency vocabulary and simple grammar structures that all levels of students can form all answers.

12. Which way would you think students should choose to form their responses to answer different types of questions in the speaking test?

- A. recall some memorised materials to answer all the questions.
- B. organise my own words to fully answer all the questions.
- C. combine memorised materials with the topic to answer all the questions.

3. Interpretation of test criteria

13. In the speaking test, I know the examiner checks students' ability to...

	No Understand	Somewhat no understand	Neutral	Somewhat understand	Fully Understan d
speak with normal levels of continuity, rate and effort.	☐	☐	☐	☐	☐
link ideas and language together to form coherent and connected speech to answer all types of questions.	☐	☐	☐	☐	☐
use accurate and range of vocabulary and grammar and the precision with which meanings and attitudes can be expressed.	☐	☐	☐	☐	☐
produce an intelligible speech to fulfil the speaking test requirements.	☐	☐	☐	☐	☐

4. Teaching Methods Selection

14. Do you think the IELTS test influences your choice of methodology (i.e. the way you teach) for IELTS preparation lessons? For example, you might do more or less group work, or you might spend more time explaining grammar rules, etc.

☐ Yes. If yes, please note here how the IELTS influences the way you teach

☐ No. If no, please note here why you have the same methodology for IELTS and non-IELTS

lessons: _____

15. How often would you use these methods to teach your IELTS speaking test classes?

	Never	Rarely	Sometimes	Often	Always
Practising making a point and providing supporting examples.	☐	☐	☐	☐	☐
Teaching students to use wide range of vocabulary and grammar structures.	☐	☐	☐	☐	☐
Asking students to recite some high mark answers to prepare for the test.	☐	☐	☐	☐	☐
Using question cards to help students to prepare for the test.	☐	☐	☐	☐	☐
Practicing pronunciation and giving feedback in the class.	☐	☐	☐	☐	☐

This is the end of the study. Thank you for your contribution. ☺.