Article

Metacognitive Instruction for Sustainable Learning: Learners' Perceptions of Task Difficulty and Use of Metacognitive Strategies in Completing Integrated Speaking Tasks

Weiwei Zhang, Donglan Zhang and Lawrence Jun Zhang *

Faculty of Education and Social Work, The University of Auckland, Auckland 1023, New Zealand;
weiwei.zhang@auckland.ac.nz (W.Z.); donglan.zhang@auckland.ac.nz (D.Z.)
* Correspondence: lj.zhang@auckland.ac.nz; Tel.: +64-9-6238899 (ext. 487550)

Abstract: This mixed-methods study investigated English-as-a-foreign-language (EFL) learners' perceptions of task difficulty and their use of metacognitive strategies in completing integrated speaking tasks as empirical evidence for the effects of metacognitive instruction. A total of 130 university students were invited to complete four integrated speaking tasks and answer a metacognitive strategy inventory and a self-rating scale. A sub-sample of eight students participated in the subsequent interviews. One-way repeated measures MANOVA and structure coding with content analysis led to two main findings: (a) EFL learners' use of metacognitive strategies, in particular, problem-solving, was considerably affected by their perceptions of task difficulty in completing the integrated speaking tasks; (b) EFL learners were not active users of metacognitive strategies in performing these tasks. These findings not only support the necessity of taking into account learners' perceptions of task difficulty in designing lesson plans for metacognitive instruction, but also support a metacognitive instruction model. In addition, the findings provide empirical support for the utility of Kormos' Bilingual Speech Production Model. As the integrated speaking tasks came from a high-stakes test, these findings also offer validity evidence for test development in language assessment to ascertain sustainable EFL learning for nurturing learner autonomy as an ultimate goal.

Keywords: metacognitive strategies; metacognitive instruction; learners' perceptions of task difficulty; Kormos' Bilingual Speech Production Model; integrated speaking tasks

1. Introduction

As one of the pivotal metacognitive elements, metacognitive strategies play a seminal role in learning English as a foreign language (EFL), and the mastery of these strategies is crucial to EFL learners' sustainable development for learner autonomy [1–5]. Acknowledging the importance of metacognitive strategies, from the perspective of pedagogy, foregrounds the significance of research on metacognitive instruction (e.g., [5–7]). A recent review by Sato and Loewen [8] shows that the current literature on metacognitive instruction is mainly devoted to how to implement such a teaching practice and whether metacognitive instruction is effective in improving EFL learners' performance. Evidently, two core variables, EFL learners, and tasks used in teaching them, are severely underexplored. In EFL classroom instructions, how to select, grade, and sequence tasks suitable for learners with different language proficiency levels so that the tasks can stimulate assumed performance in learners for a specific pedagogic purpose is a major challenge to teachers [9,10]. To address this issue, researchers interested in task-based language teaching have proposed and accordingly produced empirical evidence that obtaining information about learners' perceptions of task difficulty is an effective solution [11,12]. In light of this, in metacognitive instruction, understanding EFL learners' perceptions of task difficulty and their metacognitive strategies for completing the tasks can inform teachers of whether the tasks they adopt for their teaching activate their students' metacognitive strategies so

as to achieve their pedagogical purposes [9,10]. This thinking provides a rationale for why we decided to revisit the sustainable benefit of metacognitive instruction for EFL learners through the lens of their perceptions of task difficulty.

Also, prior empirical studies on this topic primarily focus on the language skills of listening, reading, vocabulary, and writing. Consequently, speaking has been greatly neglected [13], although proficient speaking skills contribute to learners' academic success, and speaking is a skill, which requires individuals' good mastery and deployment of metacognitive strategies [14–16]. Among various speaking tasks, integrated speaking tasks "broaden the scope of strategies called upon" [17] (p. 16), and are immediately related to individuals' metacognitive strategy use [18]. Therefore, we used this type of speaking tasks to elicit EFL learners' perceptions of task difficulty and the metacognitive strategies they intended to use. As researchers commonly used the speaking tasks from the computer-assisted TOEFL iBT (Test of English as a Foreign Language) to examine learners' performance in integrated speaking tasks (e.g., [19,20]), we also adopted this test in the present study.

In the available literature on metacognitive instruction, empirical evidence mostly comes from daily classrooms, and so studies in assessment contexts are exceptionally insufficient [13]. Given that the washback effect of foreign and/or second language (L2) assessment on student learning in classroom settings is positive (e.g., [19,20]) and that the purpose of engaging in L2 learning activities for many learners is to achieve academic success through passing high-stakes tests in one form or another (e.g., [16,20]), seeking empirical evidence from testing conditions warrants additional research efforts. Our employment of the TOEFL iBT integrated speaking tasks, to some extent, serves one such effort. Moreover, as EFL speakers' metacognitive strategy use is well illustrated in Kormos' [14] Bilingual Speech Production Model [21,22], our examination of EFL learners' metacognitive strategy use was framed within this model.

Taken together, an investigation into EFL learners' perceptions of task difficulty involved in the TOEFL iBT integrated speaking tasks, their use of metacognitive strategies within Kormos' [14] model, and the relationships between the two constructs is poised to fill the above-mentioned research gaps, which also formulated the scope of our research. We hope that the findings of our research will offer some insight into teachers' design and use of appropriate speaking tasks for improving the effectiveness of their metacognitive instruction. The effectiveness is evidence of the pivotal role of metacognitive strategies that teachers might want to harness to benefit their students by enhancing their strategic competence for sustainable EFL learning and becoming autonomous learners. Concomitantly, the study will provide additional validity evidence for test development; namely, in designing speaking tests, test-designers need to take into consideration test-takers' perceptions of task difficulty. In this way, test tasks can truly serve to assess test-takers' language ability for meeting the assumptions of test validity and reliability [23].

2. Literature Review

2.1. Metacognitive Strategies and Metacognitive Instruction

As a cross-disciplinary construct, metacognitive strategies are believed to comprise planning, monitoring, evaluating, and problem-solving in the research fields of metacognition, language learning strategies, and language assessment that exert considerable influence on learners' performance (e.g., [1–7,23,24]). Some scholars (e.g., [5,6,25]) have postulated that acquiring metacognitive strategies can help learners ease their sustainable life-long learning. They have also pointed out that such acquisition is challenging for learners if external support or scaffolding (e.g., a teachers' help) is not available. Therefore, metacognitive scaffolding, the most often used scaffolding type in education, is regarded as valuable and plays a critical part in EFL learning (e.g., [5,6,25–27]).

Metacognitive scaffolding, derived from metacognition, refers to instructional guidance that facilitates learners' metacognitive thinking to support their use of metacognitive strategies in the learning process. The underlying goal of metacognitive scaffolding is

to assist learners in solving problems [25–27], and it empowers learners to increase their self-efficacy for sustainable learning [5–7]. Amongst various means through which metacognitive scaffolding can be achieved, metacognitive instruction is ascertained to be the most direct and most popular [25–27]. In tackling an actual task, learners' learning strategies including metacognitive strategies, are not operating in isolation. Rather, they are acting both independently and interactively to influence learners' cognitive activities. Hence, a learner often uses more than one strategy at a time to complete a task [28–32]. Based on this recognition, some researchers [33–35] advocated two strategy instruction approaches: single-strategy instruction (teaching students how to use a single strategy for aiding their learning), and multiple-strategy instruction (teaching students how to choose strategies from a set of strategies for successful learning).

However, regarding how to effectively implement the two approaches, research efforts have generated diverse and inconclusive findings [25–27]. Oxford [33] (p. 312), for instance, commented that "no strategy instruction guidelines were consistently applied and that the strategy instruction took place in many different situations and conditions and with disparate learners". This comment indicates that the implementation of metacognitive instruction varies across learners. Goh [36] and Sato and Loewen [8] held a similar opinion, claiming that metacognitive instruction is often built upon learner factors, particularly learners' perceptions of tasks. As noted earlier, in pedagogy, learners' perceptions of task difficulty, to a great degree, determines if a pedagogical purpose can be achieved. With reference to this, the above scholars' views on learner factors in metacognitive instruction can be interpreted in relation to how learners' perceptions of task difficulty affect this pedagogical practice. Indeed, some researchers [36–38] suggest that learners' perceptions of task difficulty directly influence their motivation which further impacts learners' metacognitive engagement or metacognitive strategy use in classroom instructions, which we review in some detail next.

2.2. Learners' Perceptions of Task Difficulty and Metacognitive Strategy Use

Learners' perceptions of task difficulty denote the interaction between task characteristics and task-takers [38], and it is suggested that such perceptions influence learners' task motivation (e.g., [39–41]). Task motivation is a composite of trait motivation and state motivation [39,41]. The former refers to L2 learners' general motivation, one of the most important individual factors that impact their metacognitive strategy use [36], whereas the latter indicates how learners perform in a task [39]. According to Kormos and Wilby [41], several elements account for EFL learners' task motivation, including self-efficacy, expectancy-value, intrinsic motivation, and interest. Self-efficacy is learners' belief that they are competent enough to complete a given task or accomplish a specific goal. It is a critical determinant of learners' motivation. Expectancy-value theory regards learners' value of tasks and their expectations of successful task completion as significant. Intrinsic motivation and interest are interchangeable terms with both concerning learners' internal feelings in performing tasks [41–43].

Many researchers hold that task motivation elements are immediately connected to learners' perceptions of task difficulty. In other words, an increase in task difficulty perceived by learners typically attributes to a decrease in their self-efficacy, expectancy-value, intrinsic motivation, and interest. The combined decrease then leads to the decrease in learners' overall task motivation (e.g., [37,41–44]). As task motivation is the macro aspect of learners' motivation that determines metacognitive strategy use as noted above, the increase in learner's perceptions of task difficulty will in the end result in the decrease in their metacognitive strategy use.

Additionally, in learning activities, learners engage in performing tasks emotionally and cognitively [45]. Emotional engagement refers to learners' affective reactions (e.g., interest, confidence, motivation, and anxiety) to learning tasks and learning environments. In contrast, cognitive engagement refers to the approaches taken by learners to perform tasks, which highlights learners' efforts in conducting cognitive and metacognitive activities to

regulate their learning process. The two types of engagement commonly work interactively, influencing each other, and as a result, how learners perceive task difficulty affects their emotional engagement [46]. Therefore, when learners' emotional engagement is negatively affected by their perceptions of task difficulty, their cognitive engagement, including their use of metacognitive strategies in engaging in the tasks given, will be impacted in an adverse manner. This will further influence their emotional engagement [45–47]. Drawing on the above exposition, a model that reflects the effects of task difficulty on metacognitive strategy use from the perspective of learners can be established in Figure 1.
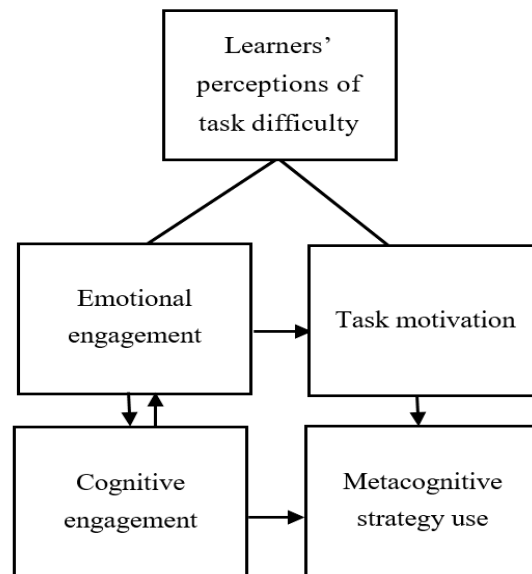

Figure 1. Task difficulty and metacognitive strategy use model.

In research, learners' perceptions of task difficulty are commonly elicited through their rating of task difficulty on a scale originally developed by Robinson [48]. For instance, Révész et al. [12] administered this scale to 48 native English speakers and 48 EFL learners to collect information on their perceptions of task difficulty involved in the simple and complex versions of three oral tasks. Likewise, Sasayama [49] investigated 53 Japanese EFL learners' perceptions of task difficulty in several oral tasks via this scale. In accordance with this, we also used the self-rating scale to gain insights into how Chinese EFL learners perceive task difficulty in completing the TOEFL iBT integrated speaking tasks.

2.3. Metacognitive Strategy Use in L2 Speaking

In the existing literature on L2 speaking, models generated in the research field of psycholinguistics are widely recognised and applied in understanding the phenomenon of L2 speaking (e.g., [14,50–52]). Among them, Kormos' [14] Bilingual Speech Production Model is "more elaborate and more targeted" [52] (p. 397), and has been regarded as the major L2 speech production model that illustrates the working mode of L2 speakers' speaking ability or their metacognitive strategy use [21,22]. As a result, this model has been employed in many empirical studies on L2 speaking (e.g., [14,50,51]). Considering its solid theoretical grounding and strong empirical support [50–53], our review of how metacognitive strategies work in the context of integrated speaking tasks was conducted within Kormos' model, as stated previously.

According to Kormos [14], L2 speech production is divided into four stages: conceptualization, in which speakers plan what they are going to speak according to task demands; formulation, where the speakers encode linguistically the intended message generated in the conceptualization; articulation, through which speakers execute their speech sounds by controlling the articulatory muscles, converting the phonetic plan generated in the formulation to overt speech; monitoring, with which speakers check and notice errors for

possible modifications and corrections to make their utterance consistent with given tasks. As speakers' L2 knowledge is typically incomplete, they will encounter various problems in speech production. In this situation, speakers will resort to problem-solving strategies to tackle these problems.

Kormos further claimed that three loops of monitoring take place in L2 speech production to inspect the outcome of the speaking process. The first loop of monitoring is to examine whether task demands are met via speakers' planning of the content and the language used in their intended discourse. The second loop of monitoring inspects errors in the phonetic plan or the internal speech produced in the stage of formulation before articulation. In the final loop of monitoring, once errors are perceived, the monitoring system will issue a signal which will initiate a new round of speech production [15,16,50,53]. During the three loops of monitoring, evaluation works in concert with this strategy in that, without evaluation, speakers are unlikely to execute their comparison between the preverbal plan generated in the conceptualization and the intended messages to be encoded. Similarly, when speakers use the monitoring to check the internal speech and the overt speech, they have to use evaluation; otherwise, they are not able to judge whether or not their actual utterances are consistent with task demands [15,16,33]. Echoing Bygate [15], Kormos [14] also proposed that metacognitive strategies involved in L2 speech production operate both covertly and overtly.

Framed in Kormos' model, the working mode of metacognitive strategy use in our study is illustrated in Figure 2, which additionally reveals the significance of metacognitive instruction in EFL speaking classrooms. Of note, as metacognitive strategies work independently, and simultaneously interactively, we define the working mode of the construct in the same way. Yet, due to the complexity of such an interactive characteristic, Figure 2 does not display the interactive working mode of the four metacognitive strategies.
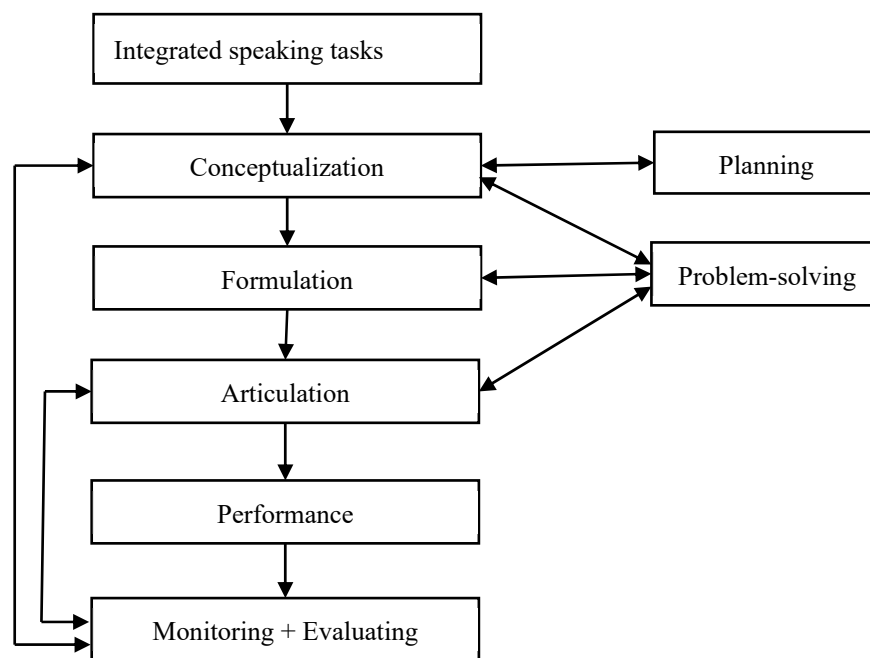
Figure 2. Metacognitive strategy use in L2 speaking.

2.4. Integrated Speaking Tasks

Integrated speaking tasks are so called in that they integrate textual (reading) and/or aural (listening) input into speaking as foundation knowledge for learners to prepare their oral responses [19,20]. Compared with independent speaking tasks that involve only speaking, this type of tasks better duplicates authentic language use. This is because, in a real-world context, language skills (reading, listening, writing, and speaking) are

typically used in an integrated form, and it is impossible to break the use of a language into isolated language skills since language users must receive input in either a written or spoken form before a real communication begins [15,19,20]. The authenticity characteristic of integrated speaking tasks enables them to illustrate well the features of real language use tasks, and they have been recognised by many as an essential means for developing learners' overall metacognitive strategy use (e.g., [19,20]). Therefore, some scholars have advocated for inclusion of this type of tasks as an important pedagogical component in EFL classroom instructions [11,20]. Despite this, integrated speaking tasks have not been widely employed in real EFL classrooms worldwide due to a lack of empirical evidence for their effectiveness [11,20].

It is also because of their authenticity, integrated speaking tasks have been applied in many popular high-stakes tests such as the TOEFL iBT speaking section [16,19,20]. As noted earlier, we used this test in our investigation into learners' perceptions of task difficulty and their tackling of metacognitive strategies. In addition, to further study the relationships between the two constructs, we selected a full set of the TOEFL iBT speaking section composed of four tasks: Task 1, Task 2, Task 3, and Task 4 (see Section 3.3.4). Such a situation as described above further augments our rationale for the adoption of integrated speaking tasks.

3. Methods

3.1. Research Design

Based on the literature review and the research scope delineated above, we addressed the following four research questions (RQ) in a mixed-methods research design [54]. A quantitative investigation was executed to answer RQ1, RQ2 and RQ3, and a qualitative exploration for answering RQ4.

RQ1. *How do EFL learners perceive the task difficulty involved in the four TOEFL iBT integrated speaking tasks?*

RQ2. *What are the metacognitive strategies used by EFL learners in completing the four TOEFL iBT integrated speaking tasks?*

RQ3. *Do EFL learners' perceptions of task difficulty affect their use of metacognitive strategies in completing the four TOEFL iBT integrated speaking tasks?*

RQ4. *Why do EFL learners use certain metacognitive strategies to tackle the task difficulty they perceive in completing the four TOEFL iBT integrated speaking tasks?*

3.2. Participants

The study included 130 EFL learners from two universities in mainland China (People's Republic of China). They were recruited via convenience sampling on a voluntary basis. Male students (*n* = 45) and female students (*n* = 85) accounted for 34.62% and 65.38%, respectively, and they were aged from 18 to 20. On average, these EFL learners reported 10 years of formal English language learning (*M* = 10.36, *SD* = 1.95).

The score range of the participants on the College English Test—Band 4 (CET-4), an authoritative English language proficiency test administered nation-wide in China [55], was from 425 points to 500 points. The CET-4 has been reported to have high reliability through tests and retests in several iterations across the country [56]. According to the official scoring interpretation of the test published by the National Education Examinations Authorities [56], such a score range suggests that the participants' language proficiency was at an upper-intermediate level, which enabled them to distinguish tasks with varying degrees of difficulty [57]. Consequently, the validity of their perceptions of task difficulty involved in the four integrated speaking test tasks that require rather higher language proficiency was established [58].

After data cleaning and assumption testing, the valid sample size was 95, meeting the threshold of the statistical procedures involved [59]. A subset of eight students participated in the subsequent semi-structured interviews.

3.3. Instruments

3.3.1. The Strategic Competence Inventory for Computer-Assisted Speaking Assessment

In empirical studies, learners' metacognitive strategy use elicited by a certain task is typically examined via self-report inventories [60,61]. In accordance with this line of research practices, we used an existing metacognitive strategy use inventory: The Strategic Competence Inventory for Computer-assisted Speaking Assessment (SCICASA), which measured EFL learners' use of metacognitive strategies or their strategic competence [14–16,23] in performing computer-assisted L2 speaking assessment. The SCICASA is a published inventory, whose reliability and validity are documented in Zhang et al. [62]. The SCICASA has two versions: The English version and the Chinese version (i.e., written Mandarin Chinese). To facilitate better understanding, we would like to provide information on the validation process in the development of the inventory. In anticipation of possible ambiguity in the Chinese version of the inventory, Zhang et al. [62], in the process of developing their inventory or questionnaire, followed the procedures that Dörnyei and Taguchi [63] recommend. Dörnyei and Taguchi recommend forward translation, retranslation, and backward translation as strategies for ensuring cultural transferability. Item consistency or reliability was established accordingly and statistics indicate the inventory has a strong reliability ($\alpha = 0.941$).

To investigate Chinese EFL learners' metacognitive strategy use in response to the TOEFL iBT integrated speaking tasks, we used the standard Chinese version of the SCICASA which has four constructs: planning, problem-solving, monitoring and evaluating. It is composed of 23 items on a 6-point Likert scale: 0 (never), 1 (rarely), 2 (sometimes), 3 (often), 4 (usually), and 5 (always). Five questions on EFL learners' background information such as age and EFL learning experience are also included. The purpose of the closing question of the inventory is to recruit possible interviewees. In order to help the international readership understand the content of the inventory, we decided to provide an English version of the SCICASA (see Appendix A).

3.3.2. Task Difficulty Self-Rating Scale

We employed Révész et al.'s [12] widely used task difficulty self-rating scale to measure learners' perceptions of task difficulty. The scale has one item, and it is rated on a 9-point Likert scale with 1 suggesting that the task is not difficult, whilst 9 indicating that the task is extremely difficult.

We translated the scale from the original language of English to standard Chinese (Mandarin Chinese) after consulting with two EFL linguistics professors whose native language is the standard Chinese for back translation [63]. The scale was then piloted on two students [54]. The scale has only one item, and therefore, we included it in the SCICASA as the second section, immediately after the participants' background information (see Appendix A).

3.3.3. The Semi-Structured Interview Guide (SSIG)

The SSIG developed by Zhang [64] was employed for an in-depth probe into Chinese EFL learner's metacognitive strategy use in tackling tasks with varying degrees of difficulty. The SSIG is comprised of five prompts or questions, and a sample question is "Could you tell me what you did in the planning time for each task?" Like the SCICASA, the SSIG has two versions (English and Mandarin Chinese). We administered the Chinese version to the Chinese EFL participants in our study. Nonetheless, considering the international readership, we presented the English version of the SSIG in Appendix B.

3.3.4. TOEFL iBT Integrated Speaking Tasks

We adopted four integrated speaking tasks from the TOEFL iBT practice online software package, TPO, to ensure authenticity. TPO features real and past test questions and aims at allowing learners to experience the real TOEFL iBT test [65]. Because of the established high validity and reliability of the test [19,20], we did not make any modifications to the test tasks, and our task selection was in light of "cultural neutrality, religious neutrality, and low controversy-provoking possibility" [66] (p. 250).

In the four tasks, Task 1 and Task 2 are reading–listening–speaking tasks with 30 s for preparations and 60 s for speaking, while Task 3 and Task 4 are listening–speaking test tasks with 20 s for preparations and one minute for speaking. Furthermore, Task 1 and Task 3 are on campus life situations whilst Task 2 and Task 4 relate to academic lectures. In terms of task type, Task 1 requires test-takers to give an oral summary of the speaker's opinion whilst Task 2 and Task 4 ask test-takers to use the examples given to illustrate an academic concept presented by the speaker. In contrast, Task 3 is about providing solutions to a specific problem given [19,66]. The above task description is summarised in Table 1.

Table 1. Four Integrated Speaking Tasks.

T	Topic	LS	PT	TT
T1	Campus-life Situations	R-L-S	30 s	Oral summary
T2	Academic Lectures	R-L-S	30 s	Illustrating a concept with examples
T3	Campus-life Situations	L-S	20 s	Providing a solution to a problem
T4	Academic Lectures	L-S	20 s	Illustrating a concept with examples

Note. T = task, LS = language skills, R = reading, L = listening, S = speaking, PT = preparation time, s = seconds, TT = task type.

3.4. Data Collection

3.4.1. Survey Data

Data collection with the two aforementioned survey tools was conducted in multimedia laboratories, and the process took each participant approximately 30 min. The participants first performed the four speaking tasks on computers installed with the TPO software packages. Each time they finished a task, they answered the SCICASA on an on-line survey platform, Wenjuanxing (https://www.wjx.cn/index.aspx, accessed on 20 June 2018), through their mobile phones for their convenience, if they wished to use their phones. After they completed all the four test tasks, they responded to the self-rating scale. To counterbalance the carryover effect, a 20-min interval between tasks was offered, and risks from the order effect were minimised through a Latin square design [67].

3.4.2. Interview Data

We conducted the interviews with eight participants in standard Chinese (Mandarin Chinese) individually. We did not specify the report language, so interviewees could use whichever language (either English or standard Chinese) that they felt comfortable with [63]. At the beginning of the interviews, we presented a briefing to the interviewees on the research objectives and relevant ethical issues. Enough time between questions was given to the interviewees for recollecting past events, thereby increasing the validity of their responses. We used note-taking, audio-recording, and researcher diary-keeping to catch every detail of the interviewees' responses, as a strategy for methodological triangulation. Closing comments with gratitude were offered to the participants, and a research report to those who expressed interest was promised [63]. Each individual interview lasted approximately 30 min and was audio-recorded for later transcription. In the whole process of data collection, we appropriately addressed ethical issues after our study was approved by The University of Auckland Human Participants Ethics Committee (Reference Number 020972).

3.5. Data Analysis and Results

3.5.1. Learners' Perceptions of Task Difficulty across Tasks

Descriptive analysis revealed that the means of the participants' ratings on task difficulty in each of the four tasks were: Task 1 (M = 5.126; SD = 1.852), Task 2 (M = 5.937; SD = 1.700), Task 3 (M = 6.000; SD = 1.816), Task 4 (M = 5.926; SD = 2.012). With reference to the self-rating scale on which the number 9 indicates extremely difficult tasks, it is clear that these numerical values were overall far above 4.5, the median value for task difficulty. This suggested that the participants perceived the four integrated speaking tasks as very difficult [59,68].

To check if the participants' perceptions significantly varied across tasks, we ran one-way repeated ANOVA. Variances in the participants' perceptions of task difficulty across tasks were examined with reference to the *p*-value for the *F*-ratio ($p \leq 0.05$), and the η^2 (if η^2 is ≤ 0.01, it suggests a small effect size; a value ranging from 0.01 to 0.06 indicates a moderate effect size, and if η^2 is ≥ 0.14, it indicates a larger effect size) [59,68]. Based on this, the outcome of ANOVA showed that there was a significant variance in the participants' perceptions of task difficulty across tasks: (F (2.646, 1586.36) = 81.121, $p < 0.001$; η^2 = 0.119). Results of the two statistical tests were used to address RQ1.

3.5.2. Learners' Metacognitive Strategy Use across Tasks

Descriptive analysis was also used to investigate the participants' use of metacognitive strategies across tasks. This addressed RQ2. Results showed that the means of the four metacognitive strategies across the four tasks all fell in the range from 3 to 4. Among them, problem-solving (M = 3.70; SD = 0.68) was used the most by the participants, followed by planning (M = 3.51; SD = 0.62) and evaluating (M = 3.25; SD = 0.65), whereas monitoring (M = 3.21; SD = 0.64) was the least frequently used strategy [59,68]. With reference to the SCICASA, "3" stands for "often" and "4" represents "usually" on a 6-point Likert scale, and given the rather high language proficiency of the participants, this range value indicated that the participants were not active users of metacognitive strategies across the four tasks.

3.5.3. Effects of Learners' Perception of Task Difficulty on Metacognitive Strategy Use

In our study, as we defined the working mode of metacognitive strategies as working both independently, and interactively, we ran one-way repeated measures MANOVA to investigate the effects of learners' perceptions of task difficulty on their use of the four interactive metacognitive strategies. During MANOVA, we inspected the variances in the participants' use of the interactive metacognitive strategies across tasks with reference to the *p*-value for the *F*-ratio, and the η^2 as we did in ANOVA reported above. The output of the MANOVA showed a substantial effect of the participants' perceptions of task difficulty on their use of the four interactive metacognitive strategies: F (12, 1212) = 12, p = 0.007 (less than the threshold of 0.05), and η^2 = 0.022 [59,68].

To investigate the effects of the participants' perceptions of task difficulty on their use of the four metacognitive strategies, we conducted a series of separate ANOVAs. The outcome revealed that the participants' perceptions of task difficulty had considerable influence on their use of the problem-solving strategy: (F (3, 405) = 3.853, p = 0.010, η^2 = 0.022). By contrast, such an effect was not found in the other three individual metacognitive strategies: planning (F (3, 405) = 1. 205, p = 0.381, η^2 = 0.008); monitoring (F (3, 405) = 0.415, p = 0.743, η^2 = 0.003); evaluating (F (3, 405) = 0.730, p = 0.474, η^2 = 0.006) [59,67]. Through the two statistical procedures, we answered RQ3.

3.5.4. Interviewees' Reported Metacognitive Strategy Use in Response to Task Difficulty

The eight interviewees' responses to the semi-structured interview were subject to structure coding following content analysis with a deductive approach. This was for addressing RQ4. During the coding process, we strictly followed the coding scheme (see Appendix C), which we developed in accordance with the SCICASA. Data transcription and analysis for each interview followed the same guideline [69].

Coding results revealed that the interviewees did not use metacognitive strategies actively (see Appendices D–F). To be specific, the participants did not have a clear and particular goal as their use of planning. They had weak awareness of self-monitoring and evaluating. By contrast, the students resorted to problem-solving actively when they faced problems in performing the tasks. In addition, the participants' metacognitive strategy use was considerably influenced by their individual attributes such as motivation, anxiety, prior experiences, and knowledge, which were further influenced by their perceptions of task difficulty. Such relationships among metacognitive strategy use, individual attributes, and task difficulty suggested that the construct of the participants' individual attributes was a mediator between their metacognitive strategy use and their perceptions of task difficulty [68]. The mediator role of individual attributes is demonstrated in Figure 3.
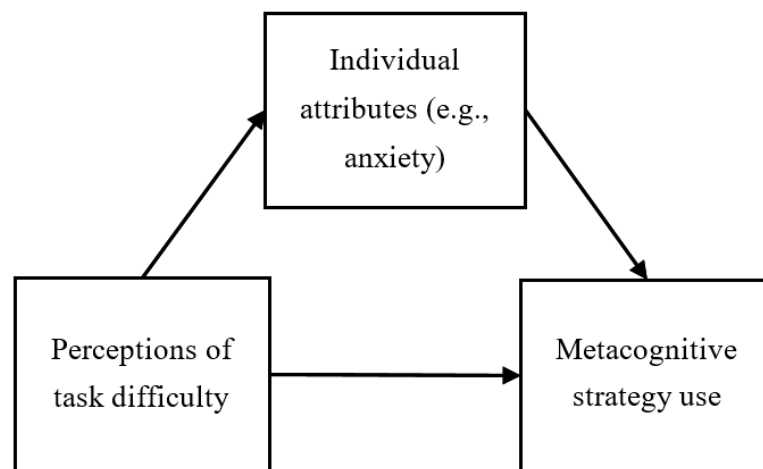

Figure 3. Individual attributes as a mediator.

In summary, the findings of our study show that the Chinese EFL participants perceived the four TOEFL iBT integrated speaking tasks as very difficult, and their perceptions varied significantly across the four tasks. It was also found that Chinese EFL participants did not use metacognitive strategies actively to tackle task difficulty in completing the four TOEFL iBT integrated speaking tasks. Among the four metacognitive strategies under investigation, they used the problem-solving strategy most frequently and monitoring least frequently. Finally, the Chinese EFL participants' use of interactive metacognitive strategies and the single problem-solving strategy was found to be substantially affected by their perceptions of task difficulty in completing the four TOEFL iBT integrated speaking tasks.

4. Discussions

4.1. Metacognitive Strategy Use and Metacognitive Instruction

As summarised above, the participants did not appear to be active users of metacognitive strategies in spite of their rather high English proficiency level. Such a finding is unexpected, because a positive correlation between language proficiency and strategy use has been evidenced in a large number of studies on L2 learning strategy use [70–73], especially in those that focus on university students [72], as was the case in this study. Therefore, we assumed that the participants in our study would have used metacognitive strategies actively.

The inconsistency between the present study and the previous literature may be due to the participants' lack of access to metacognitive instructions as they reported in the initial oral survey during the recruiting process. It is known that if individuals are able to use strategies, they need to learn them, just as they must expose themselves to language in order to acquire that language [71–73]. As the participants had not experienced any metacognitive instructions, it is possible that they were not competent and hence not active in metacognitive strategy use. In fact, in the actual Chinese educational context, EFL

teachers are not required to teach students how to systematically use different strategies, particularly metacognitive strategies. In other words, metacognitive instructions are not compulsory in Chinese EFL learning classrooms. As a result, teachers do not think these metacognitive strategies are important [73,74]. In such a context, it might be impossible for the participants to actively use metacognitive strategies when the repertoire of the strategies was not readily available to them, just as a Chinese idiom goes, "Qiǎo fù nán wéi wú mǐ zhī chuī", or, when translated into English, "it is impossible to make bricks without straw". This confirms what has been reported in the literature on the salience of metacognitive instruction in assisting learners to use metacognitive strategies, as discussed in Section 2.1.

On the other hand, the highest frequency of problem-solving reported by the participants may also reflect the influence of metacognitive instruction on EFL learners' metacognitive strategy use. Some scholars posit that the focus of teachers' instructions is on students' language proficiency, as is commonly observed in EFL classrooms worldwide [71,73,74]. Against this backdrop and as reported in the literature, problem-solving as a metacognitive strategy has the closest relationship with one's language proficiency among the four metacognitive strategies, and therefore, it has become one of the basic strategies that EFL teachers emphasise in the classroom [73,74]. As a result, the problem-solving strategy has become one of the compulsory strategies that EFL learners are required to master in their daily EFL learning activities. In fact, these teachers and students may not realise that such classroom instructions relate to metacognitive instruction [73,74]. This pedagogical practice inadvertently provided EFL learners with an opportunity to immerse themselves in a problem-solving strategy learning environment, which possibly explains why the participants used the strategy frequently in performing the four tasks.

4.2. Learners' Perceptions of Task Difficulty and Metacognitive Strategy Use

The participants' perceptions of task difficulty (very difficult) and their inactive use of metacognitive strategies are consistent with the findings reported in the literature about the effects of the former on the latter (see Section 2.2). With reference to Figure 1 that manifests the model on such effects, it is likely that when the participants perceived the four integrated speaking tasks as very difficult, their motivation and interest in engaging in these tasks were weakened. Consequently, their task motivation and emotional engagement in performing the tasks might have greatly decreased or even disappeared [37,38,47]. This may further negatively impact their cognitive engagement, and accordingly, their metacognitive strategy use might be negatively impacted, too. In light of this, it is understandable why the interviewees recalled that because of task difficulty, they experienced negative emotions such as low self-efficacy, low expectancy-value (lack of confidence in task completion), and negative emotional engagement (lack of motivation and interest; and little anxiety). Influenced by such emotions, as the interviewees reported, it was very hard for them to use metacognitive strategies actively and effectively to deal with the very difficult tasks [37,47].

Further, as Barkaoui et al. [17] and Oxford et al. [75] pointed out, if tasks are too difficult, task-takers tend to use whatever strategies that are available to tackle the tasks. In the end, their strategy use may not display variability in response to tasks with different degrees of difficulty. Their views lend some support to the result from our study that the participants' perceptions of task difficulty had no significant effects on their use of planning, monitoring, and evaluating strategies. However, the participants' use of the problem-solving strategy was found to be significantly affected by their perceptions of task difficulty. A possible explanation of this result is that problem-solving was reported as the most frequently used strategy across the four tasks. Since the participants' perceptions of task difficulty was found to demonstrate considerable variance across tasks (see Section 3.5.1), it is reasonable to speculate that the participants' use of this strategy experienced considerable variance in the four tasks with varying degrees of difficulty, which indicates the substantial effect of such task difficulty on the use of the problem-solving strategy.

4.3. Metacognitive Strategy Use in L2 Speaking

The highest frequency in the participants' use of the problem-solving strategy and the lowest frequency in their use of the monitoring strategy demonstrated the actual working mode of metacognitive strategies operating in the four integrated speaking tasks, which provides empirical evidence for Kormos' [14] model. As shown in Figure 2, because the participants may not have complete EFL knowledge, they were highly likely to use problem-solving in the whole L2 speech production to solve the problems caused by such incomplete knowledge [14]. This helps to explain the highest frequency of the strategy use, as reported by the participants.

In contrast with problem-solving, monitoring was reported as the least frequently used metacognitive strategy. Within Kormos' model, monitoring is one of the four fundamental stages in L2 speech production, and it engages in the whole procedure of L2 speaking in either a covert form or an overt form [15,16]. As the participants had no experiences in metacognitive instruction, it was possible that they had no awareness of their actual use of monitoring when the strategy was operating in their speaking process in a covert form. As a consequence, when the participants were recalling their metacognitive strategy use in tackling the speaking test tasks, they might not truly report their use of monitoring through the inventory.

In addition, Barkaoui et al. [17] have argued that compared with other language skills such as reading and writing, speaking has higher requirements of L2 speakers due to the immediate and online characteristics of speaking. In light of this, when performing the integrated speaking test tasks, the participants might have had to process a rather huge information load that involved not only speaking but also reading and listening. Challenged by such speaking tasks, the participants were unlikely to take useful notes, consciously apply learned or self-developed rules, or relate information to personal experiences, all related to the monitoring strategy under investigation (see Appendix C), especially when the participants had little knowledge of metacognitive strategies. The challenges might also further account for the highest frequency in the participants' use of the problem-solving strategy, as it is reasonable that more challenges commonly produce more problems that typically stimulate more frequent uses of the problem-solving strategy in L2 speaking [14–16]. Simply put, it is likely that it is the integrated speaking test tasks per se that elicited the lowest frequency of the participants' monitoring strategy use and the highest frequency in their employment of the problem-solving strategy.

5. Conclusions and Implications

Using a mixed-methods design, we investigated Chinese EFL learners' perceptions of task difficulty and their use of metacognitive strategies in completing four TOEFL iBT integrated speaking tasks. The findings showed substantial effects of EFL learners' perceptions of task difficulty on their use of metacognitive strategies interactively. Moreover, the findings also revealed that among the four metacognitive strategies, the participants' use of problem-solving was significantly affected by their perceptions of task difficulty. These findings are expected to provide pedagogical implications for metacognitive instruction in EFL speaking classrooms from the perspectives of learners and tasks. By the same token, the findings will offer validation implications for task development in L2 speaking assessment.

5.1. Implications for Metacognitive Instruction

From the perspective of learners' metacognitive strategy use, we propose a metacognitive instruction model for developing EFL speaking competence or skills, particularly for Chinese EFL teachers who might want to help their students in sustainable learning of EFL speaking.

As shown in Figure 4, the model encompasses problem-solving-centred strategy instruction, which emphasises teaching students how to use a single metacognitive strategy to start with, focusing on the problem-solving strategy, followed by a comprehensive

approach to metacognitive strategy instruction that aims at teaching students how to simultaneously use a cluster of four metacognitive strategies in the teaching plan, which can be termed "multiple-metacognitive-strategy instruction". In accordance with this model, teachers can take two steps in metacognitive instruction to develop students' EFL speaking competence or skills. The first step is problem-solving-centred (the arrow line highlighted in blue), single metacognitive strategy instruction, through which teachers teach their students one single metacognitive strategy at a time, and pay special attention to the use of the problem-solving strategy in designing a syllabus or a lesson plan and related classroom activities for teaching EFL speaking. By doing so, EFL teachers can impart to their students the knowledge of how to use a specific metacognitive strategy, in particular, the problem-solving strategy, to tackle a language learning task. As a result, the EFL students' awareness of adopting one specific metacognitive strategy, particularly the problem-solving strategy in task performance is likely to be raised. This type of metacognitive instruction is consistent with Oxford's [33] proposal which underscores EFL teachers' attention to their students' cognitive needs based on students' feedback on strategy use. Furthermore, Plonsky's [33–35] meta-analysis on strategy instruction also supports this type of metacognitive instruction, as his analysis reveals that when target metacognitive strategies are narrowed down in classroom instructions, students tend to learn these strategies in the most effective manner. What is more, as discussed earlier, problem-solving has always been emphasised in EFL classroom instructions, and against such a background, teaching this strategy to EFL learners is to teach metacognitive strategies according to their preference for strategy use. This will enhance learners' strategic competence as advocated by some scholars [76–79].
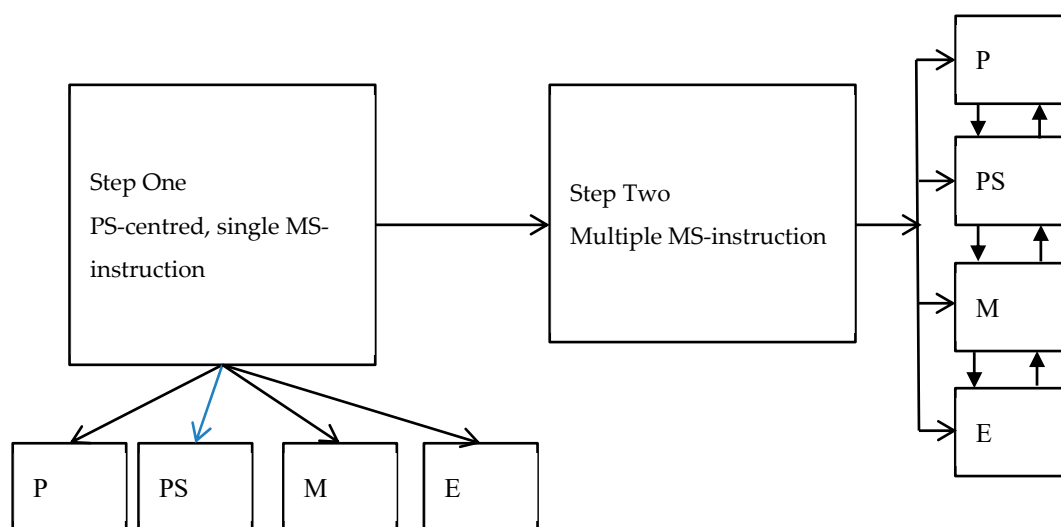

Figure 4. Metacognitive instruction model supported by this study. Note. MS = metacognitive strategy; P = planning; PS = problem-solving; M = monitoring; E = evaluating.

After their students master the four metacognitive strategies one by one, especially the problem-solving strategy, teachers can go further to take the second step: multiple-metacognitive-strategy instruction. This type of metacognitive instruction enables learners to learn and finally acquire the simultaneous use of the four metacognitive strategies to meet their actual needs in learning EFL speaking [33–35]. As evidenced by our study, in actual task performance, more often than not, learners had to use not only one single metacognitive strategy independently, but also, at the same time, multiple metacognitive strategies interactively. It is known that the latter form of metacognitive strategy use is more effective in helping learners to develop their metacognitive awareness, a critical metacognitive factor that contributes to learners' self-regulation [1,35]. This explains why multiple-metacognitive-strategy instruction is preferred by some scholars (e.g., [33–35]).

Despite this, it has been suggested that it is more learner-friendly and theoretically feasible for EFL teachers to apply multiple strategy instructions after their students have acquired each of the target metacognitive strategies as a foundation step [33]. Such a suggestion rationalises why we put problem-solving-centred instruction as the first step in our model.

From the perspective of tasks, we propose that in syllabus design for implementing metacognitive instruction in EFL speaking classes, EFL teachers need to consider selecting, grading, and sequencing speaking tasks in line with how their students evaluate these tasks, especially how they perceive the difficulty involved in them. By doing so, the teachers can make sure that the tasks that they prepare for their metacognitive instruction can be appropriate for their students in terms of language proficiency levels [9,10]. Otherwise, if the tasks are perceived by the students as too easy or too difficult, their motivation and engagement in performing these tasks may be weakened [37–39], and accordingly, they may not use metacognitive strategies as expected by their EFL teachers. In the end, the teachers' pedagogical purpose of metacognitive instruction may not be achieved due to the tasks per se. To gain the knowledge of how students perceive task difficulty, teachers can use their own teaching experience or use the self-rating scale, as endorsed by Ellis et al. [9] and Sasayama [49]. Figure 5 illustrates the possible procedure for ascertaining students' perceptions of task difficulty.
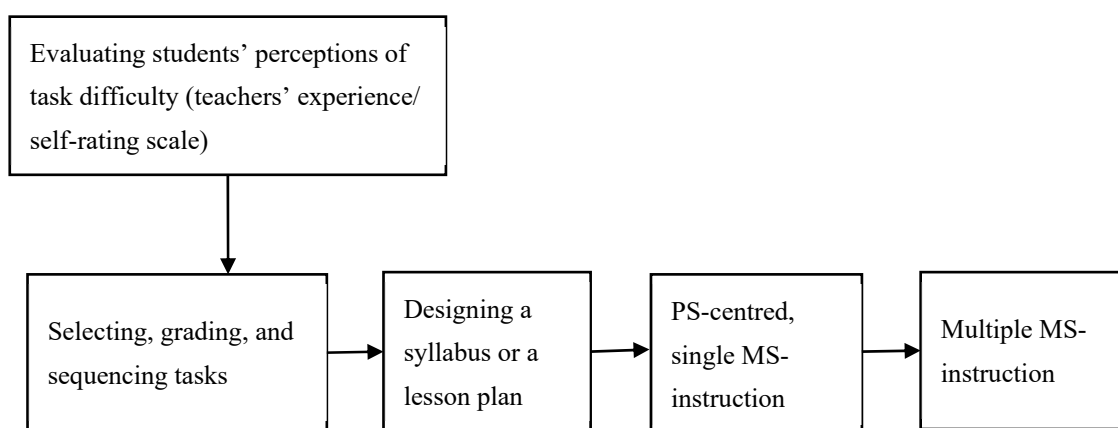

Figure 5. Metacognitive instruction procedure model for EFL speaking. Note. PS = problem-solving; MS = metacognitive strategy.

Moreover, our study may inspire EFL teachers to adopt a more holistic view in syllabus design and task development for teaching EFL speaking by integrating reading, listening, and speaking activities. As integrated speaking tasks can assist learners to develop familiarities with real-world language use tasks, and effectively activate their speaking ability [20], this holistic approach is expected to contribute positively to the success of EFL teachers' metacognitive instruction in EFL speaking classrooms.

Integrating the metacognitive instruction model into our proposal of syllabus design, we further propose another metacognitive instruction procedure model (see Figure 5) as a summary of the implications that our study can provide for metacognitive instruction in EFL speaking classrooms, particularly in the context of China.

5.2. Implications for L2 Speaking Assessment

In a similar vein, from the perspectives of test-takers and test tasks, the research findings of our study suggest that in L2 assessment, test developers should consider how test-takers perceive task difficulty. Such a consideration will ensure that the expected responses from the test-takers for examining their language ability can be prompted by the test tasks designed. If a test task is perceived by test-takers as too easy or too difficult, it may generate a test score that cannot truly reflect these test-takers' language ability in that test-takers may not use appropriate metacognitive strategies to tackle the test task [19,64].

In the end, the validity and reliability of the test, and accordingly its usefulness may be placed into question [23].

6. Limitations and Recommendations for Further Research

As convenience sampling was adopted in our study, the participants had similar backgrounds. Limitations caused by such sample homogeneity may restrict the generalisability of our research findings to other populations. This suggests that diverse sampling is preferred in future studies of relevance for better generalisability [80]. In addition, although individual attributes were found to work as a mediator between learners' perceptions of task difficulty and their use of metacognitive strategies, due to resource constraints, we did not carry out an in-depth investigation into this variable. Therefore, it is unknown what the individual attributes really were qualitatively, how they were impacted by task difficulty, and how they affected metacognitive strategy use quantitatively in integrated speaking assessment tasks. Given the salience of individual attributes in the research domains of metacognitive strategies [1–3,40], task research [9–11], and L2 assessment [23,81], further research into this construct is warranted.

Author Contributions: W.Z. conceived of the initial idea, fine-tuned by D.Z. and L.J.Z., W.Z. designed the study, collected and analyzed the data, drafted the manuscript. D.Z. and L.J.Z. revised and proofread the manuscript. L.J.Z. finalised and submitted it as the corresponding author. All authors have read and agreed to the published version of the manuscript.

Appendix A. The Strategic Competence Inventory for Computer-Assisted Speaking Assessment (English Version)

Part One Background Information

In this part please provide your information by ticking ($\sqrt{}$) in the box or write your responses in the space so we can better understand your answers.

1.　Code:
2.　Age:
3.　Gender: Male_□ Female_□ Gender diverse_□
4.　The years you have been learning English to present:
　　7~9 years_ □ 10~12 years_□ 13~15 years_ □ Others_____.
5.　English proficiency reflected by test
　　CET4_□ CET6_□_ BEC_□ IELTS_□ TOEFL_□

Part Two Task Difficulty Self-Rating Scale

Please give a rating on the task you just finished. The rating should range from 1 (the task is not difficult at all) to 9 (the task is extremely difficulty).

Tasks	This Task is Not Difficult at All	1	2	3	4	5	6	7	8	9	This Task is Extremely Difficult
Task 1											
Task 2											
Task 3											
Task 4											

Part Three The Strategic Competence Inventory for Computer-assisted Speaking Assessment

In this part, please read each of the following statements and indicate how you thought during the integrated speaking test by ticking (√) 0 (never), 1 (rarely), 2 (sometimes), 3 (often), 4 (usually), and 5 (always).

Your Thinking	0	1	2	3	4	5
1. I knew what the task questions required me to do.						
2. I was aware of the need to plan a course of action.						
3. I thought about what to do to complete the task well.						
4. I made sure I clarified the goals of the task.						
5. I understood the essential steps needed to complete the task.						
6. I organized the structure of what I was going to say before speaking.						
7. I guessed the meaning of the unknown words or expressions by using my knowledge (e.g., words in the context, knowledge of word information, knowledge of the topic.						
8. I used the context to guess the topic.						
9. I drew on my background knowledge to complete the task.						
10. I made up new words or guess if I didn't know the right ones to use.						
11. I used a word or phrase that means the same thing when I could not think of a word in English.						
12. I knew when I should complete a task more quickly.						
13. I knew when I should complete a task more carefully.						
14. I knew how much time had gone by.						
15. When I was speaking, I knew when I had spoken in a way that sounded like a native speaker.						
16. I related the incoming information to what I had known.						
17. When I was performing my task, I took notes on the important words and concepts.						
18. I knew what to do if my intended plan did not work efficiently during the task.						
19. I mentally give myself a grade after I finished my task.						
20. I checked whether I had accomplished my goal after completing my task.						
21. I checked the mistakes I had made in the task.						
22. I evaluated my performance satisfaction as I moved along the task.						
23. I evaluated whether my intended plans worked effectively.						

All participants who have completed the inventory, we invite you to participate in the interviews. If you want to join us for the interviews, please tick (√) the following statement to show your interest.

Yes, ____I want to participate in the interviews.

Appendix B. Semi-Structured Interview Guide (English Version)

1. Could you tell me what you did in the planning time for each task?
2. Could you tell me whether you had a particular goal before you started your task today?
3. Could you tell me whether you completely understood the reading and listening materials in the tests?
4. Could you tell me what you did to solve the problems in the tests if you had?

5. Could you tell me whether you had evaluated your performance in the test after the task was done?

Appendix C. Coding Scheme

Categories	Subcategories	Codes	Definitions
Metacognitive strategies	Planning	Setting goals	Identify the purpose of the task
		Directed attention	Decide in advance to focus on the given tasks and ignore distractions
		Activate background	Think about and use what you already know to help you do the task
		Prediction	Anticipate information to prepare and give direction for the task
		Organisational planning	Plan the task and content sequence
		Self-management	Arrange for conditions that help you learn
	Problem-solving	Inference	Make guesses based on previous knowledge
		Substitute	Use a synonym or descriptive phrase for unknown words
	Monitoring	Selective attention	Focus on key words, phrases, and ideas
		Deduction/induction	Consciously apply learned or self-developed rules
		Personalise/personal experience	Relate information to personal experiences
		Take notes	Write down important words and concepts
		Ask if it makes sense	Check understanding and production to keep track of progress and identify problems
		Self-talk	Talk to yourself to reduce anxiety
	Evaluating	Verify predictions and guesses	Check whether your predictions or guesses are correct
		Evaluating performance	Judge how well you do in the task
		Check goals	Decide whether goal was met

Appendix D. Coding Results

Subcategory	Interviewees	Metacognitive Strategies
Planning	All	Organisational planning (Outlining and organisation)
	In 4 and In 8	The activation of prior knowledge
	In 2 and In 3	Setting goals
	In 1 and In 3	Variation in setting goals response to task difficulty
	In 4, In 5, In 6, In 7, and In 8	No particular goals
	In 1 and In 3	Chance of oral practice
Problem-solving	In 6, In 7, In 1, In 2, and In 3	Inference
	In 8	The use of make-up
	In 8	The use of mother tongue
	In 8 and In 5	The use of gap fillers
	In 1 and In 7	Not knowing how to solve problems
	In 4	No problem-solving
Monitoring	In 1	Taking notes down in a messy manner
	In 6	Purposefully taking notes
	In 5	Making predictions and judgments in processing tasks
Evaluating	In 7 and In 8	Self-evaluation on task completion
	In 2 and In 6	Self-evaluation on understanding the tasks
	In 7	Self-evaluation on fluency of his speech
	In 1 and In 3	Self-evaluation on English learning
	In 5	Self-evaluation on his summary of the tasks
	All	Self-evaluation on personal feelings

Note. In = Interviewee.

Appendix E. New Themes with Examples

Categories	Subcategories	Themes	Examples
Metacognitive strategies	Planning	No particular goals Chance of oral practice Variance in setting goals based on task difficulty	I had no any goals, and I just thought about finishing the tasks carefully. Well, I just wanted to practice my oral English When the tasks became more difficult like Task 3, I felt hard to deal with them. And I had no goals at all. Actually, at the beginning, I had... But when the tasks were not what I expected, I felt stressful and lost all my goals in the end.
	Problem-solving	Not knowing how to solve problems No problems encountered	There was so much information, and if I couldn't understand the materials given, I had to ignore them and gave them up. I don't think I had problems. You (the researcher) has explained clearly the task requirements . . . I was very familiar with the test form and the tasks themselves

Appendix F. Unidentified Assumed Themes

Categories	Subcategories	Unidentified Codes
Metacognitive strategies	Planning	Directed attention Self-management
	Monitoring	Self-talk
	Evaluating	Verify predictions and guesses Check goals Evaluating performance

References

1. Zhang, D.L.; Zhang, L.J. Metacognition and self-regulated learning (SRL) in second/foreign language teaching. In *Second Handbook of English Language Teaching*; Gao, X., Ed.; Springer Nature International: Cham, Switzerland, 2019; pp. 883–897.
2. Zhang, L.J.; Zhang, D.L. Metacognition in TESOL: Theory and practice. In *The TESOL Encyclopedia of English Language Teaching, Vol. II: Approaches and Methods in English for Speakers of other Languages*; Liontas, J.I., Shehadeh, A., Eds.; Wiley-Blackwell: Hoboken, NJ, USA, 2018; pp. 682–792.
3. Seong, Y. Strategic competence and L2 speaking assessment. *Teach. Coll. Columbia Univ. Work. Pap. TESOL Appl. Linguist.* **2014**, *14*, 13–24.
4. Zhang, L.J. Constructivist pedagogy in strategic reading instruction: Exploring pathways to learner development in the English as a second language (ESL) classroom. *Instr. Sci.* **2008**, *36*, 89–116. [CrossRef]
5. Teng, L.S.; Zhang, L.J. Empowering learners in the second/foreign language classroom: Can self-regulated learning strategies-based writing instruction make a difference? *J. Second Lang. Writ.* **2020**, *48*, 1–12. [CrossRef]
6. Zhang, L.J.; Aryadoust, V.; Zhang, D.L. Taking stock of the effects of strategies-based instruction on writing in Chinese and English in Singapore primary schools. In *Quadrilingual Education in Singapore: Pedagogical Innovation in Language Education*; Silver, R.E., Bokhorst-Heng, W., Eds.; Springer Nature: New York, NY, USA, 2016; pp. 103–126.
7. Zhang, L.J. A dynamic metacognitive systems account of Chinese university students' knowledge about EFL reading. *TESOL Q.* **2010**, *44*, 320–353. [CrossRef]
8. Sato, M.; Loewen, S. Metacognitive instruction enhances the effectiveness of corrective feedback: Variable effects of feedback types and linguistic targets. *Lang. Learn.* **2018**, *68*, 507–545. [CrossRef]
9. Ellis, R.; Skehan, P.; Li, S.; Shintani, N.; Lambert, C. *Task-Based Language Teaching: Theory and Practice*; Cambridge University Press: Cambridge, UK, 2019.
10. Robinson, P.; Gilabert, R. Task-based learning: Cognitive underpinnings. In *The Concise Encyclopedia of Applied Linguistics*; Chapelle, C., Ed.; Blackwell: Oxford, UK, 2020; pp. 1046–1051.
11. Robinson, P.; Gilabert, R. Task complexity, the Cognition Hypothesis and second language learning and performance. *Int. Rev. Appl. Linguist. Lang. Teach.* **2007**, *45*, 161–176. [CrossRef]
12. Révész, A.; Michel, M.; Gilabert, R. Measuring cognitive task demands using dual-task methodology, subjective self-ratings, and expert judgments: A validation study. *Studies Second Lang. Acquis.* **2016**, *38*, 703–737. [CrossRef]

13. Kim, T.; Kim, J. Metacognitive Instruction Using Web 2.0 Technologies in an Adult ESL speaking course. *J. Res. Pract. Adult Lit. Second. Basic Educ.* **2017**, *6*, 50–59.

14. Kormos, J. Speech production and the Cognition Hypothesis. In *Second Language Task Complexity: Researching the Cognition Hypothesis of Language Learning and Performance*; Robinson, P., Ed.; John Benjamins: Amsterdam, The Netherlands, 2011.

15. Bygate, M. Teaching and testing speaking. In *The handbook of Language Teaching*; Long, M.H., Doughty, C.J., Eds.; Wiley: West Sussex, UK, 2011; pp. 412–440.

16. Luoma, S. *Assessing Speaking*; Cambridge University Press: Cambridge, UK, 2004.

17. Barkaoui, K.; Brooks, L.; Swain, M.; Lapkin, S. Test-takers' strategic behaviours in independent and integrated speaking tasks. *Appl. Linguist.* **2013**, *34*, 304–324. [CrossRef]

18. Cohen, A.D. *Strategies in Learning and Using a Second Language*; Routledge: London, UK, 2014.

19. Hung, S.A.; Plakans, L. Topical knowledge in L2 speaking assessment: Comparing independent and integrated speaking test tasks. *Lang. Test.* **2018**, *35*, 27–49. [CrossRef]

20. Frost, K.; Clothier, J.; Huisman, A.; Wigglesworth, G. Responding to a TOEFL iBT integrated speaking task: Mapping task demands and test takers' use of stimulus content. *Lang. Test.* **2020**, *37*, 133–155. [CrossRef]

21. Skehan, P. Foreign language aptitude, acquisitional sequences, and psycholinguistic processes. In *Cognitive Individual Differences in Second Language Processing and Acquisition*; Granena, G., Jackson, D., Yilmaz, Y., Eds.; Benjamins: Amsterdam, The Netherlands, 2016; pp. 7–40.

22. Skehan, P. *Second Language Task-Based Performance: Theory, Research, Assessment*; Routledge: New York, NY, USA, 2018.

23. Bachman, L.F.; Palmer, A.S. *Language Assessment in Practice: Developing Language Assessments and Justifying Their Use in the Real World*; Oxford University Press: Oxford, UK, 2010.

24. Chamot, A.U.; Harris, V. Language learning strategy instruction for critical cultural awareness. In *Learning strategy instruction in the Language Classroom: Issues and Implementation*; Chamot, A.U., Harris, V., Eds.; Multilingual Matters: Bristol, UK, 2019; pp. 123–139.

25. Jumaat, N.F.; Tasir, Z. A framework of metacognitive scaffolding in learning authoring system through Facebook. *J. Educ. Comput. Res.* **2016**, *54*, 619–659. [CrossRef]

26. Dabarera, C.; Renandya, W.A.; Zhang, L.J. The impact of metacognitive scaffolding and monitoring on reading comprehension. *System* **2014**, *42*, 462–473. [CrossRef]

27. An, Y.; Cao, L. Examining the effects of metacognitive scaffolding on students' design problem solving and metacognitive skills in an online environment. *J. Online Learn. Teach.* **2014**, *10*, 552–568.

28. Cohen, A.D. Moving from theory to practice: A close look at language learner strategies. In *Language Learning Strategies and Individual Learner Characteristics: Situating Strategy Use in Diverse Contexts*; Oxford, R.L., Amerstorfer, C.M., Eds.; Bloomsbury: London, UK, 2018; pp. 31–54.

29. Zhang, L.J.; Thomas, N.; Qin, T.L. Language learning strategy research in *System*: Looking back and looking forward. *System* **2019**, *84*, 87–93. [CrossRef]

30. Flavell, J.H. Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *Am. Psychol.* **1979**, *34*, 906–911. [CrossRef]

31. Flavell, J.H. Cognitive monitoring. In *Children's Oral Communication Skills*; Dickson, P., Ed.; Academic Press: New York, NY, USA, 1981; pp. 35–60.

32. Takeuchi, O. Language learning strategies: Insights from the past and directions for the future. In *Second Handbook of English Language Teaching*; Gao, X., Ed.; Springer Nature: Cham, Switzerland, 2020; pp. 684–699.

33. Oxford, R.L. *Teaching and Researching Language Learning Strategies*; Routledge: New York, NY, USA, 2017.

34. Plonsky, L. The effectiveness of second language strategy instruction: A meta-analysis. *Lang. Learn.* **2011**, *61*, 993–1038. [CrossRef]

35. Plosky, L. Language learning strategy instructing: Recent research and future directions. In *Learning Strategy Instruction in the Language Classroom: Issues and Implementation*; Chamot, A.U., Harris, V., Eds.; Multilingual Matters: Bristol, UK, 2019; pp. 3–21.

36. Goh, C.C. Reconceptualising second language oracy instruction: Metacognitive engagement and direct teaching in listening and speaking. *Asian J. Engl. Lang. Pedagog.* **2014**, *2*, 1–20.

37. Dörnyei, Z. From integrative motivation to directed motivational currents: The evolution of the understanding of L2 motivation over three decades. In *The Palgrave Handbook of Motivation for Language Learning*; Lamb, M., Csizér, K., Henry, A., Ryan, S., Eds.; Palgrave: Cham, Switzerland, 2019; pp. 39–69.

38. Cho, M. Task complexity and modality: Exploring learners' experience from the perspective of flow. *Mod. Lang. J.* **2018**, *102*, 162–180. [CrossRef]

39. Julkunen, K. Situation- and Task-Specific Motivation in Foreign Language Learning and Teaching. Ph.D. Thesis, University of Joensuu, Joensuu, Finland, 1989. Unpublished.

40. Rahimi, M.; Zhang, L.J. Writing task complexity, students' motivational beliefs, anxiety and their writing production in English as a second language. *Read. Writ. Interdiscip. J.* **2019**, *32*, 761–786. [CrossRef]

41. Kormos, J.; Wilby, J. Task motivation. In *The Palgrave Handbook of Motivation for Language Learning*; Lamb, M., Csizér, K., Henry, A., Ryan, S., Eds.; Palgrave: Cham, Switzerland, 2019; pp. 267–286.

42. Ajzen, I. Perceived behavioural control, self-efficacy, locus of control, and the theory of planned behaviour. *J. Appl. Soc. Psychol.* **2002**, *32*, 665–683. [CrossRef]

43. Horvath, M.; Herleman, H.; Lee McKie, R. Goal orientation, task difficulty, and task interest: A multilevel analysis. *Motiv Emot* **2006**, *30*, 169–176. [CrossRef]
44. Ma, J.H. Autonomy, Competence, and Relatedness in L2 Learners' Task Motivation: A Self-Determination Theory Perspective. Ph.D. Thesis, University of Hawaii, Hawaii, HI, USA, 2009. Unpublished.
45. Poupore, G. Measuring group work dynamics and its relation with L2 learners' task motivation and language production. *Lang. Teach. Res.* **2016**, *20*, 719–740. [CrossRef]
46. Lynch, R.; Hurley, A.; Cumiskey, O.; Nolan, B.; McGlynn, B. Exploring the relationship between homework task difficulty, student engagement and performance. *Ir. Educ. Stud.* **2019**, *38*, 89–103. [CrossRef]
47. Li, W.; Lee, A.; Solmon, M. The role of perceptions of task difficulty in relation to self-perceptions of ability, intrinsic value, attainment value, and performance. *Eur. Phys. Educ. Rev.* **2007**, *13*, 301–318. [CrossRef]
48. Robinson, P. Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Appl. Linguist.* **2001**, *22*, 27–57. [CrossRef]
49. Sasayama, S. Is a "complex" task really complex? Validating the assumption of cognitive task complexity. *Mod. Lang. J.* **2016**, *100*, 231–254. [CrossRef]
50. Yahya, M.Y. Improving Speaking Performance and L2 Motivation through Task-Based LANGUAGE Teaching on Malaysian Undergraduate Students. Ph.D. Thesis, The University of Reading, Reading, UK, 2019. Unpublished.
51. Skehan, P. *Researching Tasks: Performance, Assessment and Pedagogy*; Shanghai Foreign Language Education Press: Shanghai, China, 2011.
52. Wang, R.; Liu, Y. Literature review of syntactic priming experiment methods and bilingual speech production models. *Adv. Soc. Sci. Educ. Humanit. Res.* **2018**, *283*, 394–398.
53. Levelt, W.J. *Speaking: From Intention to Articulation*; MIT Press: Cambridge, MA, USA, 1989.
54. Creswell, J.W.; Creswell, J.D. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, 5th ed.; Sage: Los Angeles, CA, USA, 2018.
55. Zhang, L.M. *Metacognitive and Cognitive Strategy Use in Reading Comprehension: A Structural Equation Modelling Approach*; Springer Nature: Singapore, 2017.
56. National Education Examinations Authorities. How to Interpret Scores for College English Test. Beijing, China. 2020. Available online: http://cet.neea.edu.cn/html1/folder/19081/5124-1.htm (accessed on 29 May 2018).
57. Rahimi, M. Task Complexity, Affective Factors, and Pre-Task Planning: Effects on L2 Writing Production. Ph.D. Thesis, The University of Auckland, Auckland, New Zealand, 2016. Unpublished.
58. Kyle, K.; Crossley, S.A.; Mcnamara, D.S. Construct validity in TOEFL iBT speaking tasks: Insights from natural language processing. *Lang. Test.* **2016**, *33*, 319–340. [CrossRef]
59. Pallant, J. *SPSS Survival Manual: A Step by Step Guide to Data Analysis Using IBM SPSS*, 6th ed.; Allen & Unwin: Sydney, NSW, Australia, 2016.
60. Craig, K.; Hale, D.; Grainger, C.; Stewart, M.E. Evaluating metacognitive self-reports: Systematic reviews of the value of self-report in metacognitive research. *Metacognit. Learn.* **2020**, *15*, 155–213. [CrossRef]
61. Veenman, M.V.; van Cleef, D. Measuring metacognitive skills for mathematics: Students' self-reports versus on-line assessment methods. *Zdm Int. J. Math. Educ.* **2019**, *51*, 691–701. [CrossRef]
62. Zhang, W.W.; Zhang, L.J.; Wilson, A.J. Supporting learner success: Revisiting strategic competence through developing an inventory for computer-assisted speaking assessment. *Front. Psychol.* **2021**, *12*, 1811.
63. Dörnyei, Z.; Taguchi, T. *Questionnaires in Second Language Research: Construction, Administration, and Processing*, 2nd ed.; Routledge: New York, NY, USA, 2009.
64. Zhang, W.W. Strategic Competence, Task Complexity, and Learner Performance in Computer-delivered Integrated Speaking Test Tasks: A Study of English-as-a-Foreign-Language (EFL) Learners in China. Ph.D. Thesis, The University of Auckland, Auckland, New Zealand, 2021. Unpublished.
65. ETS (2021): TOEFL iBT®Practice Tests: TOEFL iBT®Practice Sets. Available online: https://www.ets.org/toefl/testtakers/ibt/prepare/tests/ (accessed on 1 May 2019).
66. Huang, H.D.; Hung, S.A. Comparing the effects of test anxiety on independent and integrated speaking test performance. *TESOL Q.* **2013**, *47*, 244–269. [CrossRef]
67. Corriero, E.F. Counterbalancing. In *The Sage Encyclopedia of Communication Research Methods*; Allen, M., Ed.; Sage: Los Angeles, CA, USA, 2017; pp. 278–281.
68. Frey, B.B. *The Sage Encyclopedia of Educational Research, Measurement, and Evaluation*; Sage: Los Angeles, CA, USA, 2018.
69. Richards, L. *Handling Qualitative Data: A Practical Guide*, 3rd ed.; Sage: Los Angeles, CA, USA, 2015.
70. Psaltou-Joycey, A.; Gavriilidou, Z. Language Learning Strategies in Greek Primary and Secondary School Learners. In *Language Learning Strategies and Individual Learner Characteristics: Situating Strategy Use in Diverse Contexts*; Oxford, R.L., Amerstorfer, C.M., Eds.; Bloomsbury: London, UK, 2018; pp. 248–269.
71. Griffiths, C. *The Strategy Factor in Successful Language Learning*; Multilingual Matters: Bristol, UK, 2013.
72. Fazilatfar, A.M. A study of reading strategies using task-based strategy assessment. *J. Engl. Lang. Teach. Learn.* **2010**, *53*, 20–44.
73. Zhou, L. On how to make a breakthrough in teaching Chinese high school EFL learners speaking skill in international schools. *Curric. Educ. Res.* **2020**, *8*, 112–113.

74. O'Malley, J.M.; Chamot, A.U. *Learning Strategies in Second Language Acquisition*; Cambridge University Press: Cambridge, UK, 1990.

75. Oxford, R.L.; Cho, Y.; Leung, S.; Hae-Jin, K. Effect of the presence and difficulty of task on strategy use: An exploratory study. *Int. Rev. Appl. Linguist. Lang. Teach.* **2004**, *42*, 1–47. [CrossRef]

76. Yi, J.I. Comparing Strategic Processes in the iBT Speaking Test and in the Academic Classroom. Ph.D. Thesis, The University of Leicester, Leicester, UK, 2012. Unpublished.

77. Zhang, D.L. Singaporean Secondary School Students' Metacognitive Knowledge about Oral Communication Skills Learning. Master's Thesis, National Institute of Education, Nanyang Technological University, Singapore, 2001. Unpublished.

78. Zhang, D.L. "Your English Is Too Cheem!": Singaporean Student Listening Difficulties and Tackling Strategies. *Asian Engl.* **2004**, *7*, 74–91.

79. Zhang, D.L.; Goh, C.C.M. Strategy knowledge and perceived strategy use: Singaporean students' awareness of listening and speaking strategies. *Lang. Aware.* **2006**, *15*, 199-119. [CrossRef]

80. Gurven, M. Broadening horizons: Sample diversity and socioecological theory are essential to the future of psychological science. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 11420–11427. [CrossRef]

81. O'Sullivan, B. (Ed.) *Language Testing: Theories and Practices*; Palgrave: Oxford, UK, 2011.