1. An investigation of the effectiveness and validity of planning time in Part 2 of the IELTS Speaking Test

Authors
Catherine Elder
The University of Melbourne, Australia

Gillian Wigglesworth
The University of Melbourne, Australia

This study addresses the question of whether the use of planning time for the IELTS Speaking Test assists in candidate performance.

ABSTRACT

This study investigates the relationship between three variables in the oral IELTS test – planning, proficiency and task – and was designed to enhance our understanding of how or whether these variables interact. The study questioned whether differences in performance resulted from one or two minutes of planning time. The study also aimed to identify the most effective strategies used by candidates in their planning.

Ninety candidates, in two groups – intermediate and advanced – each undertook three tasks with no planning time, one minute or two minutes' planning time. All tasks were rated by two raters, and the transcripts of the speech samples subjected to a discourse analysis.

Neither the analysis of the scores, nor the discourse analysis revealed any significant differences in performance according to the amount of planning time provided. While this suggests that planning time does not positively advantage candidates, we argue that one minute of pre-task planning should continue to be included on Task 2 of the IELTS test in the interests of fairness, and to enhance the face validity of the test. The report concludes with a discussion of possible reasons for the null findings and proposes avenues for further research.

AUTHOR BIODATA

GILLIAN WIGGLESWORTH

Gillian Wigglesworth is Associate Professor and Head of the School of Languages and Linguistics at The University of Melbourne. She has a wide range of research interests which broadly include both first and second language acquisition, language testing and assessment, and bilingualism. Gillian has several edited book publications, and numerous journal articles and book chapters which reflect her research interests.

CATHERINE ELDER

Catherine Elder is Associate Professor of Applied Linguistics and Director of the Language Testing Research Centre in the School of Languages and Linguistics at The University of Melbourne. Previously, and while undertaking this research study, she was with Monash University. Catherine is co-author of the *Dictionary of Language Testing* (CUP 1999) and co-editor of *Experimenting with Uncertainty: Essays in Honour of Alan Davies* (CUP 2001) and of the *Handbook of Applied Linguistics* (Blackwell 2004).

CONTENTS

1   BACKGROUND TO THE RESEARCH

The time variable is critical in information processing theories of speech production, and there is now a substantial body of Second Language Acquisition (SLA) research within this cognitive tradition investigating the effects of pre-task planning time on oral performance. This research has yielded fairly convincing evidence that opportunities for planning before a task impact both on the content of learners' speech and also on the quality of the language they produce. With regard to the latter, planning is seen as important because of the role it can play in helping learners access their L2 knowledge through controlled processing, promoting selective attention to form and monitoring (Skehan, 1988).

A review of the effects of planning time by Ellis (2005) shows that planning generally enhances the fluency and complexity of L2 learners' spoken performance (eg Foster, 1996; Foster and Skehan, 1996; Skehan and Foster, 1997; Wendel, 1997; Mehnert, 1998; Ortega, 1999; Yuan and Ellis, 2003). Results pertaining to accuracy are less consistent, but some studies (eg Ellis, 1997; Mehnert, 1998) show that planning also reduces the incidence of error in learner speech. This inconsistency has been attributed to a number of variables, including the characteristics of the tasks used to elicit learner speech and to the conditions under which these tasks are performed. Performance on structured tasks, for example, has been found to be more responsive to planning than is the case with unstructured tasks (Foster and Skehan, 1996; Skehan and Foster, 1997). The type of planning which learners engage in may also be important as Sangarun (2005) showed. Finally, the time allowed for planning needs also has an impact with some aspects of speech improving after only one minute of planning time and others requiring more sustained rehearsal. In Ortega's (1999) study, for example, fluency improvements were evident only after 10 minutes of pre-task planning.

One of the reasons for the intense interest in planning amongst SLA researchers is that it is believed to foster pushed output (Swain, 1993) and hence to aid acquisition, although firm evidence in support of this belief is yet to emerge. Whether or not this is the case, the different qualities of speech produced under planned and unplanned conditions provide insight into the psycholinguistic constraints on L2 production, and lend support to the distinction made by Ellis (2005) and others between implicit (automated) and explicit (analytic) knowledge. These constructs are regarded by many as central to psycholinguistic theories of second language production.

The justification for researching planning time in language testing contexts, such as the one investigated in this study, is somewhat different. Skehan (1998) invokes test validity, claiming that speaking tests need to sample language produced under planned and unplanned conditions if test scores are to be considered representative of a broad range of real world performances. Such a position begs the question of how much planning time will produce the desired variation in the quality of speech. Elder et al (2002) propose that tests like IELTS and TOEFL, which are used to predict language performance in academic settings, should include planning time for authenticity reasons, given that academic speech is more often than not planned prior to delivery. There are, however, obvious constraints on how closely test tasks can mirror the requirements of academia where students may spend several hours or days preparing for an academic presentation. In a testing context, the amount of planning time must be limited to what is practical given the resources available. It should also be acknowledged that the majority of speaking taking place in academic contexts is entirely spontaneous, so it seems logical to also include some tasks with no planning time. This however raises fairness issues – a further argument for allowing planning in testing contexts. In the highly stressful test situation, planning time may serve to reduce anxiety, a possible source of construct-irrelevant variance on a test. It may thereby give candidates opportunities to produce their best possible performance (see Swain's (1985) arguments about "biasing for best" in the test situation). However, what is not clear is either whether planning does reduce anxiety, or whether planning in fact makes a difference to test performance, as the SLA research would lead us to believe.

The few studies which have been conducted into the effects of planning in language testing contexts have produced less consistent results than is the case with classroom-based SLA research. The first study to be undertaken was that of Wigglesworth (1997), which explored the effects of planning on the oral proficiency component of the *access:* test (used to screen immigrants for entry to Australia) and found that pre-task planning increased the accuracy of certain grammatical features, such as verb tenses and articles, particularly amongst the higher proficiency candidates when performing cognitively demanding tasks. But while she found significant effects for planning at the discourse level, giving candidates pre-task planning time made no difference to their scores.

Two recent studies have also found that planning time can have a positive impact on performance. The first, by Tavakoli and Skehan (2005), which was conducted in what the authors claim to be a testing environment, found consistent benefits for planning on discourse measures of accuracy, complexity and "breakdown" fluency. The impact of planning on scores however, is not reported. Proficiency again interacted with planning time, as in the Wigglesworth study, but this time it was the less-proficient learners who gained the most (elementary planners in some cases outperformed the intermediate non-planners). Learners also found task performance easier under the planned condition. The second study by Xi (2005), which focused on a graph description task from the taped-mediated SPEAK (Speaking Proficiency English Assessment Kit), found that planning time had the effect of increasing holistic scores on some line graph tasks and also served to mitigate the effects of task familiarity on performance. Qualitative analyses revealed that candidates described more line segments and offered more complex information when planning was provided.

However, these findings are at odds with those of other test-based research, namely that of Wigglesworth (2000) and Iwashita et al (2001). In Wigglesworth's study, which focused only on test scores, planning was found to be counterproductive in the case of unstructured tasks and had little impact on learner performance on other task types. Iwashita et al (2001) found that planning before a monologic story-telling task had no impact on either the quality of test discourse or test scores, or indeed on candidates' perceptions of task difficulty. Elder and Iwashita (2005) offered a variety of tentative explanations for the discrepancy between the findings of classroom and language testing research, including the nature of the tasks themselves, of the instructions given to candidates and the opportunities for on-line planning during task performance (which they speculate may obscure the effects of pre-task preparation). They also suggest that the use of planning time by test-takers may be ineffective. Although some classroom-based research has investigated how learners use their planning time (Wendel, 1997; Ortega, 1999 & 2005; Sangarun, 2005), this issue is yet to be explored in a language testing context.

2   THE CURRENT STUDY

The current study was motivated by a desire to probe these issues further in the context of a face-to-face oral interview (the previous studies were conducted with tests requiring tape-based performances). Particular attention was paid to the design features of the study to avoid some weaknesses of previous research efforts in this area. As well as investigating the effect of different levels of planning time on learners at different levels of proficiency, we were interested in investigating the nature and effectiveness of test-taker planning processes and also in canvassing test-takers' perceptions of planning time (ie its adequacy and usefulness).

3   CONTEXT FOR THE RESEARCH

The study (funded from an IELTS Australia grant awarded in 2003) explored the effects of pre-task planning time on performance on Part 2 of the International English Language Testing System (IELTS) oral interview. The interview offers one minute's preparation time to all candidates and allows them to prepare notes which they can refer to during the actual interview. We will hereafter use Ellis's term "strategic planning" (2005: 3-5) to make it clear that we are talking about the preparation time given to candidates immediately before performing a test task rather than to pre-task rehearsal (Bygate and Samuda, 2005) in which the candidate actually practises the task prior to performing it.

The following questions were investigated in the study.

3.1   Research questions

1. Does the amount of strategic planning time provided make a difference to the scores awarded to candidates in Part 2 of the oral test?

2. Does the amount of strategic planning time make a difference to the quality of candidate discourse in Part 2 of the oral test?

3. How do candidates' perceive the usefulness and validity of strategic planning time?

4. How do candidates use their strategic planning time?

5. What are the most effective strategies for the use of strategic planning time?

3.2   Variables

Three variables were manipulated in the study design:

1. Proficiency level

2. Amount of planning time

3. Task.

3.2.1   Proficiency level

There were two groups of candidates at different levels of proficiency. Group A were intermediate level candidates as determined by previous scoring on IELTS (band 5.0-5.5) and/or institutional estimates derived from in-house measures used for placement purposes. Group B were advanced candidates (ie previous scores of 6.0 or more in the IELTS band or institutionally determined equivalent). Items from Nation's 3,000-5,000 level Academic Word List were also administered to candidates in each group to confirm the validity of these proficiency groupings. The vocabulary test was used as a surrogate for general language proficiency, which was the basis for the institutional groupings, to confirm that the candidates belonged to two distinct proficiency groupings.

3.2.2   Amount of planning time

The instructions for the IELTS Part 2 of the oral test indicate that candidates should be given "one to two minutes to prepare". Given previous research which has indicated that as little as one minute can affect performance on some discourse measures (see Mehnert, 1998; Wigglesworth, 1997), this study set out to investigate if there were any differences according to whether candidates are instructed to perform with a) no planning time, b) one minute or c) two minutes of planning time. In each case, 15 seconds was provided for the candidate to read the task.

3.2.3   Task

Three tasks were developed in line with the specifications for the Part 2 task. These were then sent to TESOL Cambridge for feedback. Modifications were made after suggestions by test developers to ensure that the tasks did indeed correspond very closely to what might be used in operational test conditions. (See Appendix 1 for the tasks and accompanying prompts to candidates). The design of the study was set up to control for variations in performance that might occur as a result of differences between the tasks, rather than as a result of the planning or proficiency variables (for details of the study design see under Interview Procedure below). This builds on previous research which has suggested that the impact of planning time on performance may be sensitive to relatively small differences in tasks (Foster, 1996; Foster and Skehan, 1996; Skehan and Foster, 1997; Mehnert, 1998; Ortega, 1999; Wigglesworth, 2001).

4   METHODOLOGY

4.1   Participants

Participants for the study were recruited from three different Australian tertiary institutions which offered English for Academic Purposes (EAP) and IELTS training. The candidates were given a small payment in compensation for time spent and also promised feedback on their performance against the various IELTS criteria (although it was explained that the resultant score was only roughly indicative of their IELTS level). The explanatory letter to participants is Appendix 2.

The participants were aged between 19 and 36 years of age, and came from a range of language backgrounds. Approximately 60% were Chinese speakers (Mandarin vs Cantonese not specified), and the remainder included Korean, Japanese, Thai, Arabic and Vietnamese speakers. Most participants had taken the IELTS test before and all were university bound, for either undergraduate or postgraduate study. All were intending to take the IELTS test in the near future. This study provided an important opportunity to practise an IELTS-like task and therefore motivation to participate was generally very high. There were 90 candidates in all, equally distributed across advanced and intermediate levels.

4.2   Study design

Each candidate did all three Part 2 task versions. In one task they were allowed no planning time; in another, one minute of planning time; and in the other, two minutes. Tasks, planning time and order were counterbalanced across candidates using a Latin Square design as indicated in Tables 1 and 2 below. There were 45 candidates in Group A (intermediate), and 45 candidates in Group B (advanced).

In each group, the candidates were divided into 3 subgroups (i, ii and iii), and within each subgroup, the candidates were divided into groups of five to avoid any practice effect. So, for example, in group Bi, all 15 candidates did Task 1 with no planning time, but five did this task first, five did it second and five did it third. Thus each student did each of the three tasks, and each student experienced one task with no planning time, one task with one minute of planning time and one task with two minutes of planning time.

Planning time	Group Ai	Group Aii	Group Aiii
0 minutes	Task 1	Task 2	Task 3
1 minute	Task 2	Task 3	Task 1
2 minutes	Task 3	Task 1	Task 2
1 minute	Task 2	Task 3	Task 1
2 minutes	Task 3	Task 1	Task 2
0 minutes	Task 1	Task 2	Task 3
2 minutes	Task 3	Task 1	Task 2
0 minutes	Task 1	Task 2	Task 3
1 minute	Task 2	Task 3	Task 1

Table 1: Research design (Group A: intermediate candidates)

Planning time	Group Bi	Group Bii	Group Biii
0 minutes	Task 1	Task 2	Task 3
1 minute	Task 2	Task 3	Task 1
2 minutes	Task 3	Task 1	Task 2
1 minute	Task 2	Task 3	Task 1
2 minutes	Task 3	Task 1	Task 2
0 minutes	Task 1	Task 2	Task 3
2 minutes	Task 3	Task 1	Task 2
0 minutes	Task 1	Task 2	Task 3
1 minute	Task 2	Task 3	Task 1

Table 2: Research design (Group B: advanced candidates)

4.3   Data collection procedures

4.3.1   Interviews

A total of eight trained and experienced IELTS interviewers were recruited for the study and were thoroughly briefed on the interview procedures. Candidates within each proficiency grouping (advanced and intermediate) were assigned randomly to interviewers who were issued with a bundle of pre-prepared student packs for their candidates. These packs contained the task prompts in the order in which they were to be administered, together with instructions for the candidates about the amount of planning time allowed (Appendix 3).

Apart from the differences in planning time, each task was administered under conditions which simulated as closely as possible the operational conditions of the IELTS interview. In the one and two minute planning conditions, candidates were given a sheet of paper and a pen to do their planning and were allowed to refer to their notes during task performance (as is normal during the IELTS interview). On completion of each task they were asked to hand the paper to the interviewer, who wrote the amount of planning time on the same sheet as well as any difficulties or notable features of candidate's behaviour that had been observed.

All interviews were tape-recorded so that any breach of the planning time instructions by either candidate or interviewer could be detected, and so that additional retrospective ratings of performance could be arranged.

Standard IELTS analytic criteria were used to rate each task performance separately as soon as the candidates completed each task (see Appendix 4 for the rating sheet). Ratings were assigned concurrently by the interviewer for feedback purposes but it was decided not to use these ratings for our research investigation given a) informal feedback indicating that interviewers found it difficult to rate one task at a time (under normal operational conditions ratings are completed once only after the interview is over) and b) our fear that ratings might be contaminated by interviewers' attitudes to planning time. (Interviewers have been found to compensate candidates for what they perceive as a difficult task or interlocutor and we believed the same might be true for task conditions perceived by raters to pose challenges to candidates.)

4.3.2   Post-interview questionnaires

On completion of the interview, all candidates filled out a questionnaire (see Appendix 6) which canvassed their perceptions of planning time. It asked about any prior strategy training the candidates had experienced (eg in IELTS preparation classes) and asked them to indicate which strategies they used during planning time. The planning strategies adopted for the questionnaire were based on those identified by Rutherford (2001) on the basis of feedback from a focus group of students very similar to the participants in the current study. Both micro level (language-related) and macro level (content-related) strategies were included. The questionnaire was administered on completion of the three-task sequence to avoid the risk of a learning effect (ie candidates using some of the strategies included on the questionnaire in subsequent task performance). Candidates then completed the vocabulary test (described above).

4.3.3   Focus groups

Candidates' perceptions regarding the difficulty/fairness of the task under the three different conditions and the utility of planning time were further probed during two focus group interviews each involving 8–10 participants from the larger study who volunteered to stay on for a further hour after the questionnaire and vocabulary test. The questions which guided the focus groups are given in Appendix 5. These focus group discussions were recorded on tape. For the purposes of this study, focus group interviews were preferred over individual interviews for two main reasons. Firstly, for entirely practical reasons, the focus group meant that the candidates were not required to wait for a long period of time. Secondly, focus groups allow for a dynamic interaction between the members of the group (Greenbaum, 1998; Bryman, 2001), which was considered to be productive in terms of drawing out candidates' views, particularly given that they were second language learners. We acknowledge, however, that the views expressed by focus group participants are not necessarily representative of the broader sample.

4.4   Data compilation and analysis

4.4.1   Transcription and digitisation of tapes

All 90 tapes were transcribed so that transcripts could be analysed and coded (see further details below). The tapes were then sent to a laboratory for digitization and a CD-Rom created of all 90 performances. Instructions from the interviewer and silences for planning were removed from the CD so that raters would be unaware of the conditions under which each task was performed.

4.4.2   Post-performance ratings

Two trained IELTS raters were recruited to rate all three tasks on each of the 90 tapes using the IELTS analytic criteria. They were instructed to take a break at least once an hour to avoid fatigue. The ratings of both assessors were then entered into a database. Inter-rater reliability was calculated (using the Spearman's correlation coefficient). The data were first analysed using the Facets rating scale model (Linacre 1990) with rater, task and proficiency and planning time entered as separate facets in the file. Univariate F tests (using SPSS) were then calculated with task and planning time entered as independent variables and average (of the two raters) scores on each of the analytic rating criteria as the outcome measures. Due to the Latin Square design, whereby candidates were randomly assigned to different planning conditions within each proficiency grouping rather than across groupings, these analyses were conducted separately for high and low proficiency candidates.

4.4.3   Discourse analysis

A subset of speech samples was selected for further analysis of the discourse. Two candidates from each of the nine cells in tables 1 and 2 were randomly selected. Thus 18 advanced and 18 intermediate candidates' speech samples were selected. Transcribed speech samples for each candidate were coded for the following categories.

1. Fluency
    ▪ fluent versus disfluent speech
    ▪ filled and unfilled pauses
    ▪ self repairs

2. Accuracy: global measures in terms of:
    ▪ error-free AS-units
    ▪ error-free clauses

3. Complexity
    ▪ proportion of dependent clauses per AS-unit
    ▪ percentage of subordinate clauses to AS-units

Fluency features were coded on the WAV files using the EMU Speech Data Base System and the R Statistical package for extracting the statistics. The EMU System offers a more accurate means of measuring fluency than does the traditional approach based solely on written transcripts. It allows data to be coded in real time on a variety of different levels chosen by the investigator, and the R package allows these to be read once the features have been labelled. In other words, stretches of fluent speech were marked at beginning and end, as were filled and unfilled pauses and self-repairs. Although much more detailed labeling is available (eg syllables can be marked and thus counted) this process was very time-consuming. It was decided not to do this in the first instance, and only to focus on a more detailed analysis in the event of significant differences between groups on the broader categories.

For the measures of accuracy and complexity, the transcripts were coded into AS-units (Foster, Tonkyn and Wigglesworth, 2000) and clauses. Following Foster et al (2000), an AS-unit was defined as an utterance consisting of an independent clause together with any subordinate clause associated with it. An independent clause was defined minimally as a clause which included a finite verb, while a

subordinate clause was defined as a clause consisting of a finite or non-finite verbal element with at least one other clausal element such as a subject, object, complement or adverbial (pp 365). Subordinate clauses were divided into two types which we labeled subordinate (when introduced by a subordinating discourse marker, eg because, before, after) and dependent, consisting of non-finite and other non-independent clauses.

Twelve speech samples were coded by the two chief investigators and reliability checks were then conducted. Areas of discrepancy were discussed and modifications were made to the coding system where necessary. The remaining speech samples were coded by a single researcher only.

4.4.4   Questionnaire responses

Questionnaire responses were entered into a database and descriptive statistics (frequencies and percentages) were calculated for the various items. T-tests were used to compare the mean number of strategies used under the one and two minute planning conditions. The relative frequency of micro- and macro- planning strategies at each proficiency level was also calculated and these frequencies were compared using the Chi Square statistic. Correlations were also computed to determine whether there was a significant relationship between number of micro- and macro-planning strategies used and test scores. The questionnaires also yielded qualitative data about candidates' attitudes to planning time. These comments were thematically coded and summarised with reference to findings from the focus group interviews (see below).

4.4.5   Focus group responses

Focus group interviews were replayed and coded for keywords based on themes emerging from the data. These themes were exemplified with verbatim quotes where appropriate.

5   RESULTS

The results of the vocabulary test, which we used as a surrogate for proficiency, confirmed that the intermediate and advanced students came from different groups, with the intermediate students averaging 46.15 (standard deviation 13.21) and the advanced students averaging 56.50 (sd 9.43). This difference was significant (t= 4.243, df 87, p<.0001). An inter-rater reliability check on the two trained IELTS raters was calculated on each of the rating categories and yielded coefficients ranging from .51 (for intelligibility) to .73 (for accuracy). However, it should be pointed out that while candidates – the majority of whom were from mainland China – were at different levels of proficiency, their speaking skills were not highly variable. This is likely to be as a result of their lack of exposure to spoken English in their previous instructional contexts. In future studies it might be useful to pre-test students' oral proficiency, rather than their general proficiency, as a means of forming the different groupings.

5.1   Research question 1

Does the amount of strategic planning time provided make a difference to the scores awarded to candidates in Part 2 of the oral test?

The first research question addressed the issue of whether strategic planning time made a difference to the scores awarded to the candidates. Mean IELTS scores and standard deviations for advanced and intermediate groups are presented in Table 3 below. The univariate analysis revealed no significant effects for either task or planning time at either level of proficiency on the global ratings, a null finding that was confirmed in the facets analysis.

Planning time	none		1 minute		2 minutes	
	mean	SD	mean	SD	mean	SD
Intermediate	23.6	2.2	23.6	2.2	23.8	2.1
Advanced	23.9	2.2	24	1.9	24	2

Table 3: Total IELTS score (N=90)

Similarly, descriptive statistics presented in Tables 4 and 5 below show only minimal mean differences according to planning time on each component of the analytic rating scale at each proficiency level. The univariate F test again confirmed that there were no significant effects for either task or planning time.

Planning time	none		1 minute		2 minutes	
	mean	SD	mean	SD	mean	SD
Fluency	5.8	0.9	5.8	0.9	5.8	0.8
Lexis	6.0	0.7	5.9	0.7	6.0	0.7
Grammar	5.8	0.6	5.8	0.7	5.8	0.7
Pronunciation	6.0	0.3	6.1	0.3	6.1	0.3

Table 4: Analytic measures for intermediate candidates (N=45)

Planning time	none		1 minute		2 minutes	
	mean	SD	mean	SD	mean	SD
Fluency	6.1	0.7	6.1	0.7	6.0	0.7
Lexis	6.0	0.7	6.1	0.5	6.1	0.6
Grammar	5.9	0.7	5.9	0.5	5.9	0.6
Pronunciation	5.9	0.6	6.0	0.5	6.0	0.5

Table 5: Analytic measures for advanced candidates (N=45)

5.2   Research question 2

Does the amount of strategic planning time make a difference to the quality of candidate discourse in Part 2 of the oral test?

The discourse analytic measures were used to determine whether planning time made a difference to the quality of the discourse in these tasks. As indicated above, the discourse of a subset of candidates was assessed on measures of fluency, accuracy and complexity. The fluency measures identified the percentage of fluent versus non fluent speech, filled and unfilled pauses, and duration of reformulations, repetitions and false starts (self repairs). The results for the intermediate candidates are given in Table 6, and those for the advanced candidates in Table 7. The univariate analyses yielded no significant differences for either task or planning time across any of these measures.

Planning time	none		1 minute		2 minutes	
	mean	SD	mean	SD	mean	SD
% fluent vs non fluent speech	65.20	10.87	66.83	9.96	65.85	10.73
Unfilled pauses	25.96	10.86	27.59	12.80	27.76	14.89
Filled pauses	9.58	4.79	8.07	5.91	8.47	4.19
Reformulations (duration seconds)	2.74	1.94	2.49	1.65	3.48	2.61
Repetitions (duration secs)	3.64	2.70	3.78	1.60	3.64	2.74
False starts (duration secs)	3.34	2.14	3.55	1.71	3.09	1.59

Table 6: Fluency measures for intermediate candidates (N=18)

Planning time	none		1 minute		2 minutes	
	mean	SD	mean	SD	mean	SD
% fluent vs non fluent speech	69.49	9.41	68.83	9.26	70.13	7.53
Unfilled pauses	22.49	9.05	21.62	8.58	21.69	8.61
Filled pauses	6.79	3.92	7.17	2.58	7.72	3.88
Reformulations (duration seconds)	2.04	1.61	2.09	1.83	2.86	2.37
Repetitions (duration secs)	3.58	5.42	2.80	2.30	3.47	3.18
False starts (duration secs)	2.63	2.34	2.51	1.81	3.54	3.33

Table 7: Fluency measures for advanced candidates (N=18)

There were two measures of complexity – proportion of dependent clauses per AS-unit, and percentage of subordinate clauses per AS-unit. Once again, as shown in Tables 8 and 9, the mean scores for each planning condition were fairly close, although there does appear to be an increase in the number of subordinate clauses per AS-unit in the one minute planning condition for both intermediate and advanced candidates. However, this difference was not large enough to reach statistical significance.

Planning time	none		1 minute		2 minutes	
	mean	SD	mean	SD	mean	SD
Dependent clauses / AS-unit	1.4	0.4	1.5	0.5	1.5	0.5
Subordinate clauses / AS-unit	16.5	10.2	26.9	21.3	21.8	15.1

Table 8: Complexity measures for intermediate candidates (N=18)

Planning time	none		1 minute		2 minutes	
	mean	SD	mean	SD	mean	SD
Dependent clauses / AS-unit	1.7	0.2	1.7	0.2	1.7	0.4
Subordinate clauses / AS-unit	21.9	12.8	27.1	14.5	20.4	20.6

Table 9: Complexity measures for advanced candidates (N=18)

The global measures for accuracy (error free AS-units and error free clauses) are presented in Tables 10 and 11. Statistical analyses again indicated that there were no significant differences according to either task or the amount of planning time provided.

Planning time	none		1 minute		2 minutes	
	mean	SD	mean	SD	mean	SD
% error free AS-units	26.5	21.7	27.3	23.1	24.8	22.7
% error free clauses	40.4	9.9	40.1	21.4	39.1	12

Table 10: Accuracy measures for intermediate candidates (N=18)

Planning time	none		1 minute		2 minutes	
	mean	SD	mean	SD	mean	SD
% error free AS-units	26.1	16.8	26.5	16.8	30	16.7
% error free Clauses	39.1	16.6	42	18.7	40.3	15.8

Table 11: Accuracy measures for advanced candidates (N=18)

To summarise, there were no significant differences in any of the score measures, or in the discourse measures between groups depending upon whether they had had access to one or two minutes' planning time, or whether they had had no planning time.

The implications of these results for continuing to include planning time in Part 2 of the IELTS test are discussed further below.

5.3   Research question 3

How do candidates' perceive the usefulness and validity of strategic planning time?

Candidates were asked in the questionnaire whether they felt that planning time helped them, to which 89% responded positively. This was reiterated in the focus group interviews where most of the students said that they found it easier when planning time was available "*Planning time is important ... you can organise your idea and prepare what you want to say*". One candidate stated that planning time is useful not only for organising ideas but also for providing time in which to calm nerves in the stressful testing situation.

The comment section of the questionnaire provided some interesting insights. The candidates were asked to comment on three aspects of their performance in the questionnaire: a) why planning time had not helped them b) which task they thought they had performed best on and why and c) which task they had performed worst on and why. Very generally, the candidate responses can be broken down as follows.

Planning time was used to:	Number	% of candidates
Organise	21	23.59
Improve ideas/think about topic	18	20.22
Improve speaking	16	17.98
Structure	5	5.61
Nervousness	3	3.37
Other	16	17.98
Negative	10	11.24
Total	**89**	**100%**

Table 12: Use of planning time by candidates

Negative responses indicating that planning time was not useful or even counterproductive were few in number, although one candidate at the focus group interview suggested that having to prepare in front of the interviewer made him more anxious than when he spoke without any planning. Typical responses from the major categories are given below.

Organise

- planning time helps organise ideas (candidate 11)
- planning lead to organise my ideas (cand 23)
- helped me know what I have to say and what is first, second… (cand 34)
- I can decide on ideas and organise them (cand 40)
- had time to organise topic and write down my idea (cand 49)
- it can help to organise my ideas (cand 54)
- I can prepare and organise my ideas to explain better (cand 63)
- I can organise my thinking and ideas before speaking (cand 64)
- helped me organise my idea (cand 85)

Improve ideas/think about topic

- more time allows you to better use ideas (cand 13)
- makes me brainstorm (cand 15)
- think about the topic step-by-step (cand 33)
- can prepare and think more to say about the topic (cand 44)
- helps me think about the content of the topic (cand 45)
- I spent time thinking about how to extend my topic (cand 53)
- thought about more things to talk about (cand 68)
- I can describe more about the topic (cand 80)

Improve speaking

- can make speaking more clearly (cand 6)
- because I can speak well planning the tasks (cand 27)
- improve my speaking in English (cand 43)
- I know what I am going to talk, making me more fluent (cand 55)
- successful speech – smooth (cand 60)
- I didn't think about the topic, but it helped to speak calmly (cand 70)
- helps speak clearly (cand 79)
- I wrote the points and then I was able to speak clearly (cand 82)

Structure

- organised sentences better (cand 3)
- can think about how to make sentences correctly then word form (cand 65)
- tried to write down words relating to my topic (cand 66)
- better arrangement, grammar structure and fewer awkward sentences (cand 73)

Perception of 'worst' task

Interestingly, when asked to identify which of the three tasks they did *worst* on, many commented that they did worst on a particular task because they did not have time to prepare their response properly. (Task 1 was describing a subject they had studied; Task 2 was describing a book or movie; and Task 3 was describing an important event in their lives.)

- The last task. I wasn't able to take notes, so I had to think immediately (cand 10, task 1, no planning)
- Subject. I had little time to prepare (cand 13, task 1, no planning)
- The first one. The time was only enough to remember my event (cand 20, task 2, no planning)
- The second one. No enough time even to read the topic (cand 28, task 2, no planning)
- The first one due to no enough time (cand 32, task 3, no planning)
- Event due to no time to plan (cand 33, task 3, no planning)
- The last task due to not enough time to get ready (cand 36, task 3, no planning)
- The last one. No time to prepare (cand 39, task 3, no planning)
- Task 1. I had no time to organise or think (cand 49, task 1, no planning)
- Subject. I had no time to think about my ideas (cand 54, task 1, no planning)
- Subject. I had no time to think about the topic (cand 60, task 1, no planning)
- The first one due to no enough time (cand 65, task 2, no planning)

- The last one. Without preparation, I kept repeating the same information (cand 82, task 3, no planning)
- Event. I had neither time nor ideas (cand 85, task 3, no planning)

As can be seen from the examples above, this was often when the candidates had no planning time available, but this was not always the case, as the following examples show.

- Task 2: time was short and the topic was hard for me (cand 4, task 2, one minute)
- Task 3. I didn't have time to think (cand 5, task 3 (event), two minutes)
- The last one. Not enough time (cand 26, task 3, one minute)
- Task 1. I had no time and didn't know what to say (cand 78, task 1, one minute)
- Task 1. Not enough time (cand 84, task 1, one minute)

Topic was another important factor which impacted on the activity. As can be seen from the responses below, the task they found most difficult was often where they found the topic difficult, and the presence or absence of planning time was unlikely to make much difference.

- 2nd. In the middle of that task, I couldn't talk about anything (cand 5, task 2, one minute)
- Event. I don't have information about it (cand 7, task 3, two minutes)
- Task 2, subject. I had no idea (cand 14, task 2, one minute)
- The third one. I seldom watch movies (cand 15, task 2, one minute)
- The first one. I couldn't think of anything (cand 16, task 2, no planning)
- Book/movie. I had no idea about it (cand 21, task 2, no planning)
- Subject. I've never thought about this (cand 22, task 1, two minutes)
- Book. I was confused (cand 35, task 2, two minutes)
- Subject. I have no idea about it, even when I use my own language (cand 4, task 1, one minute)
- Subject. I've never thought about this task (cand 43, task 1, one minute)
- Subject. I have no idea to describe a subject (cand 44, task 1, one minute)
- First one. I have never thought of it before (cand 47, task 1, no planning)
- Book. I had no idea about the book (cand 51, task 2, one minute)
- The last one. I didn't know about the topic very well (59, task 2, one minute)
- Subject. I've never done this task before (cand 71, task 1, two minutes)
- The third one. I had nothing to say (cand 75, task 3, one minute)
- Movie/book. It was hard to describe a book, especially some Chinese book (cand 79, task 2, two minutes)
- Event. This topic is too big (cand 80, task 3, no planning)

5.3.1 Topic as a factor

Topic was also identified as a salient factor in responses to the question about which task they felt they had performed best on, with 56 of the candidates (62.9%) mentioning this, compared to 21 (23.5%) who identified the presence of planning time as the major determinant of their performance. However, 19 of the responses in the latter category indicated that it was the two minutes of planning time which they perceived to have made the difference and of these, five mentioned both planning time and topic as contributing. Some typical responses are below.

- Maybe the movie because I am interested in it (cand 3)
- Task 1. The topic was easier than others (cand 6)
- Event. I have a lot of events in my life, that I can explain very well (cand 11)
- Book. I just read the book recently, so I can remember (cand 24)

- Task 2. My memorable event in my life that I never forget (cand 41)
- Last one. It was a part of my life (cand 47)
- Movie. There were many ideas in my mind (cand 64)
- Event. It was the most important event in my life (cand 72)
- Task 1. I got many points to talk about (cand 81)
- Task 3. I was familiar with the topic (cand 83)
- Movie. I'm interested in it (cand 86)

5.3.2 Planning time as a factor

- Second one. Enough time to think (cand 8, task 3, two minutes)
- Event. I had more time to prepare (cand 13, task 3, two minutes)
- Task 3 about subject. I could use two mins to plan (cand 16)
- The first one. I had time to prepare (cand 28, task 1, two minutes)
- Task 3. I had more time to prepare (cand 34, task 3, two minutes)
- Event. I had more time to think about my ideas and how to say them (cand 54, task 3, two minutes)
- Event. I had more time to prepare.
- With two mins, you have enough time to think of it (cand 59)
- The last one. I had time to prepare it (cand 65, task 1, two minutes)

5.3.3 Planning and topic as a factor

- 3: I had time to organise my ideas and the topic was familiar with me (cand 4)
- Task 3. I had enough time to think and the test name "talking about movies" was interesting (cand 14)
- Task 3. I had more time to organise and the topic was easier for me (cand 49)
- Subject. Enough time to prepare and familiar subject (cand 30)

It should be pointed out that in the questionnaire almost 50% of the candidates claimed to be familiar with the tasks. Almost 15% of the candidates reported they had previously practised the tasks in their responses to a subsequent question "Which task do you think you performed best on? Why?"

- Subject. I am familiar with it (cand 21)
- Movie. I have done this topic before. I'm familiar with it (cand 31)
- I practised it before (cand 62).
- Subject. I prepared this topic before and am familiar with the vocabulary (cand 80)

As discussed below, this may be a factor which contributes to the null findings presented in this study.

Overall, from the candidates' responses above, it appears that although planning time does not seem to affect scores, or engender differences in the discourse measures investigated above, the majority of the candidates clearly found it useful, and identified difficulties when it was not present. Nevertheless, the topic of the task emerged as the most important factor in how candidates perceive themselves performing on these tasks.

5.4   Research question 4

How do candidates use their planning time?

The analysis of the strategy questionnaires revealed that the candidates used a variety of strategies when they had planning time available. The most common strategies used are given in Table 13; the six most popular strategies are shaded.

Strategy	one minute planning		two minutes planning	
	Number	%	Number	%
I tried to decide what topic I would talk about	72	80.9	68	76.4
I thought about the content and ideas needed for the task	58	65.2	57	64.0
I read the task card again	57	64.0	53	59.5
I thought about how to organise my ideas	53	59.5	61	68.5
I wrote down vocabulary on paper	42	47.2	51	57.3
I wrote down useful sentences or phrases on paper	40	44.9	44	49.4
I thought about grammar (eg verb forms) in my head	32	35.9	37	41.6
I made notes about grammar on paper	11	12.4	17	19.1
I practised useful sentences or phrases in my head	33	37.1	42	47.2
I made a list of vocabulary in my head	26	29.2	31	34.8
I made a list of useful organising and/or linking language in my head	32	35.9	43	48.3
I wrote down useful organising and/or linking language on paper	22	24.7	35	39.3
I practised the task in my head	27	30.3	35	39.3
I practised pronunciation in my head	15	16.8	21	23.6
I wrote down ideas in my first language and then translated them	10	11.2	13	14.6
I thought about nothing	12	13.5	9	10.1

Table 13: Use of strategies by candidates with one and two minutes planning time

While the pattern of strategy use is similar for both the one and two minute planning condition, there was a significant difference between the number of strategies candidates reported using when more planning time was available (t=2.575, df=88, p=0.012).

5.5   Research question 5

What are the most effective strategies for the use of strategic planning time?

Given the results of the previous analyses, it was anticipated that there would be no significant correlations between the number of planning strategies and either the global or analytic scores given by the raters for planning condition. This proved to be the case.

A further analysis was undertaken which involved identifying the strategies as either macro-strategies (those concerned with topic, content and organisation), and micro-strategies (those concerned with language level issues such as grammar, structure, vocabulary, etc). The last strategy ('I thought about nothing'), which attracted very few responses, was omitted. (See Figure 1.)

Macro-strategies	Micro-strategies
I read the task card again	I thought about grammar (eg verb forms) in my head
I practiced the task in my head	I made notes about grammar on paper
I practiced pronunciation in my head	I practiced useful sentences or phrases in my head
I tried to decide what topic I would talk about	I wrote down useful sentences or phrases on paper
I thought about how to organise my ideas	I made a list of vocabulary in my head
I thought about the content and ideas needed for the task	I wrote down vocabulary on paper
I wrote down ideas in my first language and then translated them	

Figure 1: Macro and micro strategies

Table 14 summarises the strategy use by proficiency level and amount of planning time provided. While it appears that macro-strategies were used more frequently than micro-strategies under the one minute planning condition and that the reverse was true when two minutes of planning was allowed, a Chi Square analysis revealed no significant differences across any of the groupings.

		Macro-strategies used	Micro-strategies used
Intermediate	1 minute	149	138
	2 minutes	139	166
Advanced	1 minute	128	115
	2 minutes	148	155

Table 14: Use of micro and macro strategies by group

Finally, the results of a t-test analysis indicated that there was no significant difference in the mean level of performance between micro and macro planners (ie candidates who reported using more language related strategies and those who reported focusing more on content and organisation).

6   DISCUSSION AND CONCLUSION

The null findings in this study mirror those of Iwashita et al (2001) and Wigglesworth (2000). As noted in our earlier literature review, test-based research has produced scant evidence of benefits for strategic planning time on the quality of the subsequent speaking performance. In this study, the lack of any effect for planning time was consistent across all measures used, including the different categories of the IELTS rating scale and the various discourse dimensions. While there was some trend towards greater discourse complexity (as measured by the ratio of subordinate clauses to AS-units) under the one minute planning condition for both intermediate and advanced level candidates, this finding did not prove to be statistically significant. It therefore seems reasonable to conclude that planning time has limited utility for Part 2 of the IELTS oral test, which uses very similar tasks.

Does this mean that the one minute of planning time currently available to prepare performance on Part 2 of the IELTS oral is superfluous? We think not. Candidates' expressed preference for planning time is worth taking notice of, if only for face validity reasons. Providing opportunities for planning may engender greater confidence in the IELTS Speaking Test on the part of candidates and, accordingly, greater acceptance of the scores obtained. However, while candidates' questionnaire and interview responses suggest that removing the currently offered one minute of planning time from IELTS task 2 is likely to be unwelcome, there is surely no point in extending the amount of planning time provided, since the longer (two minute) planning condition yielded no additional benefit on any performance measure. Even for complexity, the marginal gains observed under the one minute condition disappeared completely when two minutes of planning were provided.

As far as strategies are concerned, the results of this study (and indeed from most other studies of planning in a test situation) suggest that, while candidates appreciate being given planning time before speaking, they make poor use of it. There was no evidence that either the number of strategies or the particular type of strategy (macro or micro) used by learners made a significant difference to performance. Interviewer feedback after administering the test indicated that many learners appeared lost during the planning period, or were too anxious to make use of what they had prepared. This is supported by comments made by one of the focus group interviewees, who reported that the presence of the interviewer distracted him from his planning efforts. Another commented that she was unable to read the notes she had made.

Another possibility (also reflected in comments from focus group candidates) is that the benefits of planning are constrained by memory, and that improvements in the fluency, accuracy or complexity of the discourse cannot be sustained beyond the first few utterances of candidate speech. It seems likely that raters are also constrained by memory and that it is the final impression which informs their judgement. This would explain the lack of any impact for planning on scores and on the discourse measures which are averaged across the whole stretch of performance.

It is also possible that in an unpressured monologic performance such as this one, candidates are able to monitor their speech as they go, and that this produces benefits even in the zero planning condition (see Yuan and Ellis 2005). The effects of strategic planning may therefore be discernible only under highly-pressured performance conditions where on-line planning is not possible. Further investigation of this may be warranted using the approach adopted by Yuan and Ellis (2005) in which on-line planning is sharply differentiated from pre-task planning and no planning by introducing a time limit for both the pre-task and no planning conditions, but providing unlimited time for the on-line planning condition.

Alternatively, it may be that there is a mismatch between the focus of candidate planning and what is valued by the IELTS rater and captured by our discourse measures. The strategies which candidates reported using most frequently in both the one and two minute planning conditions were those directed to planning the message content, whereas the main focus of the IELTS analytic rating scale categories is on form, or, to be more precise, candidates' accuracy, fluency, pronunciation and the lexical resources they deploy. It might therefore be instructive to devise some means of measuring the propositional complexity of the discourse, to see if planning makes a difference to this dimension of performance (although it is debatable whether propositional complexity is of interest in a language testing context).

It might also be useful to examine in more detail those individuals who benefit from planning to determine what planning strategies these candidates engage in. However, to do so, we would need to devise a more fine-grained taxonomy of strategy use (see Ortega, 2005) and to gather rich think-aloud data (of the kind elicited by Sangarun, 2005). Such a study would be of interest to those involved in teaching test preparation courses and could form the basis for further research on the role of strategy training in boosting performance.

As pointed out above, many of the candidates reported having practised these or similar tasks before. It may be that planning is to no avail when candidates are already familiar with the task, particularly simple ones (like those used in this study) which require a description or commentary on past experience. Indeed it may be that on a high-stakes test such as IELTS, some candidates have prepared so well that much of what we are really measuring on this test is pre-rehearsed rather than spontaneous unplanned discourse (although this study provides no direct evidence of such a phenomenon, which should be the subject of further research). On the other hand, we saw comments from a number of test-takers indicating they were unprepared for the topics and in these instances, as was suggested earlier, planning time may do little to improve their performance.

The current study adds to the weight of evidence suggesting that planning time is not conducive to producing better performance in a testing environment. However, Xi's (2005) recent findings in relation to the graph task on the SPEAK exam nevertheless give some grounds for believing that planning time may interact with task type. Before definitive conclusions are drawn, further research needs to be conducted using more complex and cognitively demanding tasks. In this respect, integrated tasks in which candidates may be required to integrate specific features of aural and written input in their oral response, may mean that planning is more beneficial than in other types of task. This would mitigate again, for example, the situation found in this study where some candidates find the *topic* difficult and this overrode the availability or not of planning time. In integrated tasks, where familiarity (or not) with the task is likely to be less of an issue since input material is given, planning would certainly be warranted, not only for reasons of fairness, but also on authenticity grounds.

In summary, the findings of this study offer positive support for the inclusion of a small amount of planning time on oral proficiency tests. However, the null findings on all measures of both rater evaluations and of the discourse suggest that the rationale for this relates more to fairness and face validity, than to the ability of candidates to improve their performance as a result of planning time. As already noted, it is clear that further research into the effects of planning time in testing contexts is warranted if we are to fully understand the impact that the provision of planning time may have in oral proficiency tests, and the ways in which it may impact on the test construct.

REFERENCES

Bryman, A, 2001, *Social Research Methods,* Oxford University Press, Oxford

Bygate, M and Samuda, V, 2005, 'Integrative planning through the use of task-repetition' in *Planning and task performance in a second language,* ed R Ellis, John Benjamins, Amsterdam and Philadelphia

Crookes, G, 1989, 'Planning and interlanguage variation', *Studies in Second Language Acquisition,* vol 11, pp 183-199

Elder, C and Iwashita, N, 2005, 'Planning for test performance: What difference does it make?' in *Planning and task performance in a second language,* ed R Ellis, John Benjamins, Amsterdam and Philadelphia

Elder, C, Iwashita, N and McNamara, T, 2002, 'Estimating the difficulty of oral proficiency tasks: what does the test-taker have to offer?', *Language Testing,* vol 19, no 4, pp 347-368

Ellis, R, 1987, 'Interlanguage variability in narrative discourse: style shifting in the use of the past tense', *Studies in Second Language Acquisition*, vol 9, no 1, pp 1-20

Ellis, R, ed, 2005, *Planning and task performance in a second language,* John Benjamins, Amsterdam and Philadelphia

Ellis, R and Yuan, F, 2005, 'The effects of careful within-task planning on oral and written task performance' in *Planning and task performance in a second language,* ed R Ellis, John Benjamins, Amsterdam and Philadelphia

Foster, P, 1996, 'Doing the task better: how planning time influences students' performance' in *Challenge and change in language teaching,* eds J Willis and D Willis, Heinemann, London

Foster, P and Skehan, P, 1996, 'The influence of planning and task-type on second language performance', *Studies in Second Language Acquisition,* vol 18, pp 299-323

Foster, P, Tonkyn and Wigglesworth, G, 2001, 'Measuring spoken language: a unit for all reasons', *Applied Linguistics,* vol 21, no 3, pp 354-375

Greenbaum, TL, 1998, *The handbook for focus group research*, Thousand Oaks, Sage, California

Iwashita, N, McNamara, T and Elder, C, 2001, 'Can we predict task difficulty in an oral proficiency test? Exploring the potential of an information processing approach to task design', *Language Learning* vol 21, no 3, pp 401-436

Linacre, M, 1990, *FACETS,* computer program for many faceted Rasch measurement, Mesa Press, Chicago

Mehnert, U, 1998, 'The effects of different lengths of time for planning on second language performance', *Studies in Second Language Acquisition,* vol 20, no 1, pp 83-108

Ortega, L, 1999, 'Planning and focus on form in L2 oral performance', *Studies in Second Language Acquisition,* vol 21, pp 109-148

Ortega, L, 2005, 'What do learners plan? Learner-drive attention to form during pre-task planning' in *Planning and task performance in a second language,* ed R Ellis, John Benjamins, Amsterdam and Philadelphia

Rutherford, K, 'An investigation into the effects of planning on oral production in a second language', unpublished masters dissertation, University of Auckland, New Zealand

Sangarun, J, 2005, 'The effects of focusing on meaning and form in strategic planning' in *Planning and task performance in a second language,* ed R Ellis, John Benjamins, Amsterdam and Philadelphia

Skehan, P, 1996, 'A framework for the implementation of task-based instruction', *Applied Linguistics*, vol 17, pp 38-62

Skehan, P, 1998, *A cognitive approach to language learning,* Oxford University Press, Oxford

Skehan, P and Foster, P, 1997, 'Task type and task processing conditions as influences on foreign language performance', *Language Teaching Research,* vol 13, pp 185-211

Skehan, P and Foster, P, 1999, 'The influence of task structure and processing conditions on narrative retellings', *Language Learning*, vol 49, no 1, pp 93-120

Swain, M, 1985, 'Large-scale communicative language testing: A case study' in *Testing*, Pergamon Press, Oxford, pp. 35-46

Swain, M, 1993, 'The output hypothesis: Just speaking and writing aren't enough', *The Canadian Modern Language Review,* vol 50, pp 158-164

Tavarkoli, P and Skehan, P, 2005, 'Strategic planning, task structure and performance testing' in *Planning and task performance in a second language,* ed R Ellis, John Benjamins, Amsterdam and Philadelphia

Wendel, JN, 1997, *'*Planning and second language narrative production', unpublished doctoral dissertation, Temple University, Japan

Wigglesworth, G, 1997, 'An investigation of planning time and proficiency level on oral test discourse', *Language Testing*, vol 14, no 1, pp 101-122

Wigglesworth, G, 2000, 'Issues in the development of oral tasks for competency-based assessments of second language performance' in *Studies in immigrant English language assessment,* Vol 1, Research series 11, ed G Brindley, National Centre for English Language Teaching and Research Macquarie University, Sydney, pp 81-124

Wigglesworth, G, 2001, 'Influences on performance in task-based oral assessments' in *Task based learning,* eds M Bygate, P Skehan and M Swain, Addison Wesley Longman, London, pp 186-209

Yuan, F and Ellis, R, 2003, 'The effects of pre-task planning and on-line planning on fluency, complexity and accuracy in L2 monologic oral production', *Applied Linguistics,* vol 24, no 1, pp 1-27

Xi, X, 2005, 'Do visual chunks and planning impact performance on the graph description task in the SPEAK exam?', *Language Testing,* vol 22, no 4, pp 463-508

APPENDIX 1: TASK PROMPTS PROVIDED FOR CANDIDATES

TASK 1   SUBJECT

Describe a subject you have studied which has had a great influence on your life:

You should say

> ❖   what the subject was
>
> ❖   where you learned the subject
>
> ❖   who your teacher was

and explain how it has influenced your life.

TASK 2   BOOK OR MOVIE

Talk about a book or a movie that you found interesting.

You should say:

> ❖   what the book or movie was about
>
> ❖   who the main characters were
>
> ❖   what you liked and/or disliked about it

and explain why you found the book or movie interesting.

TASK 3   EVENT

Describe an event in your life (eg holiday or childhood experience) which made a great impression on you.

You should say:

> ❖   what the event was
>
> ❖   where and when it took place
>
> ❖   who you were with

and explain why it made a great impression on you.

APPENDIX 2: TASK ADMINISTRATION INSTRUCTIONS FOR INTERVIEWER

When there is NO PLANNING TIME you should say the following:

Now, I'm going to give you a topic and I'd like you to talk about it for one to two minutes. I'd like you to start talking straight away. Do you understand?

Here's your topic [hand over the relevant task card and give students 15 seconds to read the card]

I'd like you to talk about X (mention the topic of the task)

All right? Remember you have one to two minutes for this so don't worry if I stop you. I'll tell you when the time is up. Can you start speaking now please?

When there is ONE MINUTE OF PLANNING TIME you should say the following:

Now, I'm going to give you a topic and I'd like you to talk about it for one to two minutes. Before you talk, you'll have one minute to think about what you are going to say. You can make some notes if you wish. Do you understand?

Here's some paper and a pen for making notes *[hand over spare paper and a pencil]* and here's your topic *[hand over the relevant task card]*

I'd like you to talk about X *(mention the topic of the task)*

Allow up to a minute for preparation, but the candidate can start earlier if he/she wants.

When the time is up or the student signals readiness to begin you should say:

All right? Remember you have one to two minutes for this, so don't worry if I stop you. I'll tell you when the time is up. Can you start speaking now please?

When there is TWO MINUTES OF PLANNING TIME you should say the following:

Now, I'm going to give you a topic and I'd like you to talk about it for one to two minutes. Before you talk, you'll have two minutes to think about what you are going to say. You can make some notes if you wish. Do you understand?

Here's some paper and a pen for making notes*[hand over spare paper and a pencil]and* here's your topic *[hand over the relevant task]*

I'd like you to talk about X *(mention the topic of the task)*

Allow up to two minutes for preparation, but the candidate can start earlier if he/she wants. When the time is up or the student signals readiness to begin you should say:

All right? Remember you have one to two minutes for this so don't worry if I stop you. I'll tell you when the time is up. Can you start speaking now please?

When the student has finished the task you should retrieve the notes he has made and attach your own notes (if relevant) to them. Say:

Thank you very much.

APPENDIX 3: MARKING SHEET

Student's number_____

Interviewer name_____

Please give 4 ratings for each task, using the normal IELTS criteria, namely:

FC =   **Fluency and coherence**
LR =   **Lexical resources**
GRA =   **Grammatical range and accuracy**
P =   **Pronunciation**

	FC	**LR**	**GRA**	**P**
Task 1	_____	_____	_____	_____
	_____	_____	_____	_____
Task 2	_____	_____	_____	_____
	_____	_____	_____	_____
Task 3	_____	_____	_____	_____
	_____	_____	_____	_____

Tasks are to be rated one at a time in order of performance.

APPENDIX 4: FOCUS GROUP INTERVIEW QUESTIONS

1. Did you the think the tasks used for this study were a good measure of your ability to use language in university settings? (Give reasons for your answer)

2. Did you find planning time made the tasks easier? If no, please explain why. If yes indicate how you see the benefits of planning time (ie how did it help you?)

3. Which planning activities were most helpful in performing the task?

4. Do you think you used the planning time as well as you could have? Say why/why not.

5. If you took notes during the planning session did you use these when performing the task? If yes, did having the notes in front of you help you?

6. Have you ever been given instruction/training on how to use pre task planning time? If yes, how useful was it? If no, do you think it would help to have this kind of training?

APPENDIX 5: STUDENT QUESTIONNAIRE

A Task Feedback

1. Have you practised any of the three tasks you have just done before? (Tick yes or no)

Talking about a SUBJECT	Yes ☐	No ☐
Talking about a BOOK/MOVIE	Yes ☐	No ☐
Talking about an EVENT	Yes ☐	No ☐

2. Have any of your teachers taught you how to plan before speaking?   Yes ☐   No ☐

3. For two of the three tasks you have just performed some planning time was given.

Indicate (by ticking all the relevant boxes) which of the following things you did during your planning time before you started speaking.

	With 1 minute	With 2 minutes
TASK NAME	_____	_____
I read the task card again	☐	☐
I thought about grammar (eg verb forms) in my head	☐	☐
I made notes about grammar on paper	☐	☐
I practised useful sentences or phrases in my head	☐	☐
I wrote down useful sentences or phrases on paper	☐	☐
I made a list of vocabulary in my head	☐	☐
I wrote down vocabulary on paper	☐	☐
I made a list of useful organising and/or linking language in my head	☐	☐
I wrote down useful organising and/or linking language on paper	☐	☐
I practised the task in my head	☐	☐
I practised pronunciation in my head	☐	☐
I tried to decide what topic I would talk about	☐	☐
I thought about how to organise my ideas	☐	☐
I thought about the content and ideas needed for the task	☐	☐
I wrote down ideas in my first language and then translated them	☐	☐
I thought about nothing	☐	☐

I did other things (please tell us what you did) _____

Do you think the planning helped you?   Yes ☐   No ☐

Explain why/why not _____

Which task do you think your performed best on? Why? _____

Which of the three tasks do you think you performed worst on? Why?_____