John Benjamins Publishing Company



This is a contribution from TASK 1:2
© 2021. John Benjamins Publishing Company

This electronic file may not be altered in any way. The author(s) of this article is/are permitted to use this PDF file to generate printed copies to be used by way of offprints, for their personal use only.

Permission is granted by the publishers to post this file on a closed server which is accessible only to members (students and faculty) of the author's/s' institute. It is not permitted to post this PDF on the internet, or to share it on sites such as Mendeley, ResearchGate, Academia.edu.

Please see our rights policy on *https://benjamins.com/content/customers/rights*
For any other use of this material prior written permission should be obtained from the publishers or through the Copyright Clearance Center (for USA: *www.copyright.com*).

Please contact *rights@benjamins.nl* or consult our website: *www.benjamins.com*

Impact of task preparedness on L2 oral performance and strategy use

The case of Chinese EFL learners

Qianqian Gu
KU Leuven, Belgium

The present study sets out to explore the effects of pre-task planning and unpressured on-line planning on L2 learners' oral performance and their choices of planning strategies in a dialogic task condition. Forty-eight intermediate Chinese EFL learners were invited to perform the task and were then assigned to four groups, each with a different planning condition. Complexity, accuracy, and fluency of their oral production were measured. Results indicated that in the dialogic task condition, unpressured on-line planning raised syntactic complexity. Strikingly, pre-task planning did not improve L2 performance in all dimensions. Additionally, a trade-off effect was found between complexity and accuracy. Retrospective interviews were conducted to explore strategies employed by the participants and their perceptions of task preparedness. Results showed that the participants preferred to use metacognitive strategies and social/affective strategies in the dialogic task. Both advantages and limitations were identified by the participants regarding different planning conditions.

Keywords: pre-task planning, on-line planning, CAF measures, planning strategy

Introduction

Over the last few decades, researchers in the field of task-based language teaching (TBLT) have paid growing attention to the effects of task planning on L2 performance. As stated in Ellis (2018), investigations in task preparedness are of great importance to language teachers since these studies provide empirical evidence that could inform them how their tasks should be implemented. Apart from the pedagogic implications, for researchers in the field of TBLT, exploring the effects of task preparedness helps validate models of language processing and speech production (Ellis, 2005).

https://doi.org/10.1075/task.20020.gu
TASK 1:2 (2021), pp. 196–226. ISSN 2666-1748 | E-ISSN 2666-1756 © John Benjamins Publishing Company

Studies on task preparedness mainly investigate the impact of pre-task planning (or strategic planning) and on-line planning, with a focus on three aspects of L2 production identified in Skehan (1996), namely complexity, accuracy, and fluency (CAF). Positive effects of task planning on L2 performance have been found but there are important issues that need to be addressed. One major issue is that most planning research employed monologic tasks but interactive tasks are the norm in task-based language teaching classrooms (Ellis, 2018). Several studies have found that whether a task is monologic or interactive will significantly influence learners' performance. Michel, Kuiken, and Vedder (2007) found that when the task was dialogic, participants tended to produce more accurate but less complex speech. They also spoke more fluently in the dialogic task, which might be a result of turn-taking behaviors. Tavakoli (2016) compared various fluency measures in monologic and dialogic conditions. It was reported that the dialogic task performance was more fluent in all aspects of fluency except for the number of clause-internal and clause-external pauses. Regarding the different impacts of monologic and dialogic tasks on L2 performance, it is necessary to investigate the effects of pre-task planning and on-line planning in dialogic task conditions.

Another issue is that only a few studies have employed a qualitative approach to investigate planning behaviors. Ortega (2005) was the first to explore the planning process from the language learners' perspective. As emphasized by Ortega, it is essential to investigate how learners perceive and utilize pre-task planning since this will further researchers' understanding of why pre-task planning helps some but not all learners improve their performance. Another advantage of investigating the planning process qualitatively is that it provides researchers with more direct evidence on which aspects of performance are prioritized by the learners. The CAF measures can reflect how learners allocate their attention but it relies heavily on the researcher's interpretation (Ellis et al., 2019). Finally, the qualitative approach is important in pedagogical terms as it will inform language teachers about which planning strategies should be taught and which should be avoided (Pang & Skehan, 2014). Nevertheless, it is striking to see that after more than a decade since Ortega's investigation, the number of studies that follow the qualitative approach remains extremely limited. In view of this, the current study will carry out retrospective interviews to qualitatively explore the strategies used by learners in different planning conditions when they engage in an interactive task.

Planning and task performance

Ellis (2005, 2018) distinguished two major options of task preparedness: pre-task planning and on-line planning. The former is often operationalized by giving participants sufficient time to prepare for the task prior to task performance, while

© 2021. John Benjamins Publishing Company
All rights reserved

the latter is operationalized by providing participants with sufficient time to perform the task.

For pre-task planning, it has been investigated with learner factors such as proficiency level. Kawauchi (2005) examined the effects of three types of pre-task planning on low-intermediate, high-intermediate, and advanced Japanese EFL learners. It was found that pre-task planning benefited the high-intermediate learners regarding fluency and complexity, whereas the low-intermediate group produced significantly more accurate speech after planning. Surprisingly, advanced EFL learners did not benefit much from pre-task planning, which might be due to the ceiling effect mentioned in the study.

Though most studies were carried out in a more relaxed setting such as classrooms, some studies investigated pre-task planning in a testing setting (Ellis, 2009). Elder and Iwashita (2005) created a testing environment that was similar to the Test of Spoken English (TSE). In the planning condition, participants were given 75 seconds to read the instructions and 3 minutes to prepare for the task. The study, however, failed to find any statistically significant differences between planners and non-planners. Several explanations were provided such as inefficient instructions, learners' lack of knowledge on how to utilize the planning opportunity, and insufficient planning time. Tavakoli and Skehan (2005) also explored the effects of planning in a testing setting. Surprisingly, the results were entirely different from Elder and Iwashita (2005). In their study, the guided planning group produced fewer end-of-clause pauses, more subordinate clauses, and more accurate speech. In other words, guided planning in this study benefited complexity, accuracy, and fluency of L2 performance. Ellis (2005) explained that this was because the participants in Tavakoli and Skehan (2005) were not guided to treat the tasks as a test.

The investigation of guided and unguided pre-task planning is another research area that provides pedagogic implications. Sangarun (2005) aimed at finding out whether learners' attention could be directed to focus on specific aspects of performance when they engage in pre-task planning. Three guided planning conditions were designed: focus on meaning, focus on form and focus on meaning/form. Detailed instructions were provided and participants were trained to plan aloud. Results indicated that the participants' attention could be directed as the meaning-focused group engaged in extensive meaning planning and the form-focused group planned more on lexical forms. An important finding was that the meaning/form-focus group achieved a balance between meaning planning and form planning, which offered more insights for language teachers regarding instructional design.

Few studies have been carried out to explore the effects of on-line planning on L2 performance. Two studies should be mentioned here as they operationalized

© 2021. John Benjamins Publishing Company
All rights reserved

on-line planning differently. Yuan and Ellis (2003) were the first to investigate on-line planning in TBLT. They operationalized on-line planning by giving the participants unlimited performing time. Results showed that on-line planning raised syntactic complexity and accuracy. Skehan and Foster (2005) operationalized on-line planning by giving the participants extra information mid-task so that they had to replan on-line. To ensure that on-line planning was elicited, repair measures such as reformulation and false-starts were added. Results indicated that the provision of extra information did not influence the participants' performance.

It should be mentioned that the differences in choosing specific CAF measures may have contributed to the mixed results obtained in planning studies (Ellis, 2018). Michel (2017) suggests that complexity can be measured in terms of lexical complexity and grammatical complexity. Almost all studies on task planning measured grammatical complexity but the integration of lexis is a more recent development (Skehan, 2009). Fluency measures are more complicated in planning research. Some studies only measured the speed of L2 performance (e.g., Ortega, 1999; Yuan & Ellis, 2003), while other studies added measures of pauses and repairs (e.g., Tavokoli & Skehan, 2005; Skehan & Foster, 2005). Measuring accuracy, on the other hand, tends to be more consistent as most studies calculated errors globally (Ellis, 2009).

Another key issue of performance in planning research is that it is rare to find all areas of performance being raised (Ellis, 2009). A possible explanation is the trade-off effect. Skehan (1998) proposed the Trade-off Hypothesis, which argues that L2 learners' attentional resources are limited and that they cannot attend to all aspects of performance simultaneously. Typically, a trade-off effect may occur between complexity and accuracy. If learners prioritize complexity as they try to produce more complex linguistic forms, this may go at the expense of accuracy, and vice versa. It is also possible that when learners try to produce more complex or accurate speech, they will slow down their speech and consequently influence their fluency. However, this does not mean that learners cannot overcome their limited attentional resources (Skehan, 2009). For instance, the provision of pre-task planning offers learners more time to attend to the Conceptualizer stage in Levelt's (1989) Speech Production Model, which is beneficial for complexity and fluency (Skehan, 2009). The operationalization of on-line planning in Yuan and Ellis (2003), on the other hand, provides learners with more time to attend to the Formulator stage, which may lead to the improvement in syntactic complexity and accuracy. Following the literature, the current study will combine pre-task planning and unpressured on-line planning as a full-planning condition to find out whether all dimensions can be raised in the interactive task.

© 2021. John Benjamins Publishing Company
All rights reserved

Planning and strategy use

The study conducted by Ortega (2005) is remarkable given that it was the first study to qualitatively investigate pre-task planning strategies employed by L2 learners. An important contribution of the study is the coding scheme developed based on O'Malley & Chamot (1990) and Oxford (1990). Three categories of planning strategies were identified, namely metacognitive strategies, cognitive strategies, and social/affective strategies. Metacognitive strategies refer to strategies that are higher order (Brown et al. 1983), such as organizing the content and monitoring the speech (O'Malley & Chamot, 1990). For L2 learners, it is essential to apply metacognitive strategies as these strategies help allocate their attention when the incoming information is overwhelming (Oxford, 1990). Cognitive strategies in language learning are important as they can be used to directly deal with specific language tasks. Social/affective strategies such as lowering anxiety and cooperating with peers are useful given that communication is essentially a social activity (O'Malley & Chamot, 1990). Ortega reported that, in general, the participants viewed pre-task planning positively. They tended to use the time to write notes for later retrieval, rehearse their speech, and organize the storyline. Interestingly, some participants did not believe that pre-task planning was necessary, which might be that the narrative tasks did not require too much planning to perform well.

Building on Ortega's study, Pang and Skehan (2014) carried out another qualitative study that tapped into planning strategies. An advantage of this study is that the participants' performance was mapped with the planning strategies. For example, strategies related to structure building were found to favor accuracy and complexity. Another interesting finding is that the study reported several strategies that learners should not use, such as *use accurate words* and *plan how to express better*. This suggests that when learners engage in pre-task planning, they need to be realistic about their ability and not to plan ambitiously. Building on these studies, the current research will qualitatively investigate the strategies employed by the participants as well as their opinions on task preparedness.

Research questions

To explore the impact of pre-task planning and unpressured on-line planning on L2 oral performance in dialogic task conditions, the following research questions were formulated:

1. What are the effects of pre-task planning and on-line planning on the complexity, accuracy, and fluency of L2 oral production in the dialogic task?

© 2021. John Benjamins Publishing Company
All rights reserved

2. Do any trade-off effects occur in pre-task planning and unpressured on-line planning conditions?

Another three questions were added regarding the types of strategies adopted by the L2 learners and their perceptions of task preparedness:

3. What are the strategies used by L2 learners when they have time for pre-task planning?
4. What are the strategies used by L2 learners when they have time for on-line planning?
5. What are the L2 learners' perceptions of the two types of planning?

Methodology

Participants

The present study invited 48 Chinese EFL learners who were university students to perform the task. Since the researcher was living in Belgium at the time of the study, twenty-eight of the participants who were also living in Belgium were invited. These participants completed the experiment and the retrospective interviews face-to-face with the researcher. Due to the Covid-19 pandemic, the other 20 participants stayed in China and performed the task and the retrospective interviews online via the software Tencent Meeting. The participants' speaking proficiency was equal to the B2 level based on the Common European Framework of Reference for Languages (CEFR) (Council of Europe, 2001). More specifically, all of them have obtained 5.5–6.5 in the speaking section of the IELTS exam or 19–21 in the speaking section of the TOEFL exam. Forty-four of the participants were female, while four of them were male. Their age ranged from 20–24 years old. Most of the participants had learned English as a foreign language in China for 10–13 years and none of them had been to English-speaking countries for education purposes. In addition, forty-four of the participants were Humanities students, while the other four participants studied Statistics and Engineering. All of the participants received a financial reward for participating in the experiment.

Design of the study

This study adopted a between-subjects design with four planning conditions. Participants were paired with a partner and were asked to perform a dialogic task. Immediately after the task performance, each participant was invited to participate in an interview on their planning strategies as well as their perceptions of

© 2021. John Benjamins Publishing Company
All rights reserved

task preparedness. Different sets of questions were designed for each planning condition, including follow-up questions (Appendix B). All the interviews were conducted in the participants' L1 to ensure that the students were able to clearly describe their strategies and perceptions. Those interviews were then transcribed and translated by the researcher. A coding scheme on planning strategies was used to categorize and order the participants' answers.

Coding scheme

The coding scheme for planning strategies was originally developed in Ortega (2005), which was based on O'Malley & Chamot (1990) and Oxford (1990). Three categories were distinguished, namely metacognitive strategies, cognitive strategies, and social/affective strategies. Table 1 presents the coding scheme with examples, which were labeled with codes indicating planning conditions. For instance, *PP-2-1* refers to pre-task planning group, pair 2, speaker 1.

Table 1. The coding scheme for planning strategies

Type of strategies		Example
1. Metacognitive strategies		
1.1	Organizational planning	*I was trying to read and digest information about the candidates, like their educational background, expected salaries. And I tried to find which one was better.* (PP-2-1)
1.2	Problem identification	*I knew I was going to make mistakes on third-person pronouns, so I sort of wrote it down on my note.* (PP-1-2)
1.3	Production monitoring	(Interviewer: Were you paying attention to what you said?) *Sure, sure. I was told to correct grammatical mistakes since primary school. I think I was very aware when I made mistakes on third-person pronouns.* (NP-1-1)
1.4	Style monitoring	*I didn't want to sound like I was just chatting with friends, you know, with all these simple words.* (OP-1-2)
1.5	Prediction*	*I kind of pre-assumed that my partner would make a different choice.* (PP-3-1)
2. Cognitive strategies		
2.1	Writing for retrieval	*I listed some similarities and differences among those candidates.* (PP-3-2)
2.2	Rehearsing	*I think I tried to list their advantages in my mind and then say it out.* (PP-6-2)
2.3	Paraphrasing*	(Interviewer: Are you trying to say budget?) *Yeah, yeah. I forgot how to say it so I said our expected salary.* (FP-1-1)
2.4	Translating	*I wrote down the note in Chinese and then translate them when I was talking. I think it's faster, um, to take note.* (PP-2-2)

© 2021. John Benjamins Publishing Company
All rights reserved

Table 1. *(continued)*

Type of strategies		Example
3. Social/affective strategies		
3.1	Empathizing with the listener	*I think it's more important to make it easier for my partner to understand. So, I tried to use simple words.* (NP-2-1)
3.2	Avoiding embarrassment*	*I was thinking about whether I should ask more questions or talk more because I was afraid that it would be very embarrassing if none of us was talking.* (FP-6-1)
3.3	Lowering anxiety	*I was really nervous cuz I didn't have time to prepare at all. So, I knew it was sort of messy, but I had to stay calm.* (NP-4-1)

* strategies emerged in the current study

The task

To elicit interactions, a decision-making task was developed by the researcher. The participants were asked to act as human resource managers in an international high school and were given basic information (in English) about four candidates regarding their educational background, work experiences, academic awards, and expected salaries (Appendix A). They were asked to discuss the qualities of the four candidates and to reach a decision on the most appropriate candidate for a teaching job.

We opted for a decision-making task for several reasons. First, this type of task is commonly used in interactive language classrooms. Second, it is likely to elicit interaction between the participants since they have to express their opinions, exchange reasons, and negotiate with each other. Finally, decision-making tasks provide reasonable challenges for learners' attentional and other cognitive resources as they have to learn and comprehend new information, find reasons to defend themselves, and search for proper linguistic resources (Foster & Skehan, 1996).

The procedure

Participants were first paired with their partners and were then randomly assigned to four groups, each with one planning condition. It should be mentioned that the participants were allowed to choose the partner that they were familiar with so that they would be more willing to interact with each other. Instructions of the task were given in the participants' L1 to ensure that all the participants understood the task clearly. They were allowed to ask questions regarding the task and the planning conditions before performing the task. After the

© 2021. John Benjamins Publishing Company
All rights reserved

participants confirmed that they fully understood the task, they were asked to perform it and were audio-recorded. All of the recordings were transcribed and coded by the researcher.

The first group was the non-planning group (NP). Participants in this group were asked to go through the task material for thirty seconds. They were told that they had to discuss and select the best candidate after the thirty seconds and were asked to immediately start performing the task. Participants in this group had a maximum of three minutes to complete the task. This group was considered as the baseline when comparing the performance with the other groups.

The second group was the pre-task planning group (PP). Participants in the second group were given three minutes to prepare for the task individually. This time limit was determined based on a pilot study that consisted of six participants who were also first-year postgraduate students at the same university. Participants were allowed to make some notes but were not allowed to look at their notes while performing the task to ensure pressured on-line planning. This group also had three minutes to complete the task.

The third group was the unpressured on-line planning group (OP). Like the first group, participants in this group were told to go through the material for thirty seconds. After the thirty seconds, they were told to choose the best candidate and were required to immediately start performing the task. This group was given unlimited time for their performance. On average, each pair of participants used approximately 6 minutes to complete the task.

The fourth group was the full-planning group (FP). Participants in this group were told to read the task material for thirty seconds. They were then given three minutes for pre-task planning and unlimited time to complete the task. On average, each pair of participants in the fourth group used around 5 minutes to complete the task.

Table 2 summarizes the planning conditions.

Table 2. Planning conditions

	Pre-task planning	Unpressured on-line planning
NP, $n = 12$	–	–
PP, $n = 12$	+	–
OP, $n = 12$	–	+
FP, $n = 12$	+	+

NP = non-planning; PP = pre-task planning; OP = on-line planning; FP = full-planning

© 2021. John Benjamins Publishing Company
All rights reserved

Measures

Complexity

Syntactic complexity and lexical complexity were measured in the present study. Two measures of syntactic complexity were employed: the average number of words per AS-unit and the average number of subordinate clauses per AS-unit. The two measures are expected to reflect two dimensions of grammatical complexity emphasized in Michel (2017), namely length and interdependence. To ensure the reliability of AS-unit segmentation, a second researcher who is familiar with AS-units was invited. 25% of the data were coded by the second researcher and the interrater reliability was 98.5%. Discrepancies were resolved by discussions on the principles of AS-units. Minimal utterances were excluded before performing the two measures since Foster et al. (2000) suggested that utterances such as *yeah, yes, right*, etc. might distort interactional data. Lexical complexity was measured by the index of Guiraud (Guiraud, 1959), which is defined as $G = types/\sqrt{tokens}$. This measure is more text length-independent compared to the traditional raw type-token ratio (Vermeer, 2000).

Accuracy

Accuracy was calculated based on the percentage of error-free AS-units. All syntactic and lexical errors were found and the total number of error-free units in each text was calculated and divided by the total number of AS-units. Though previous studies used specific measures to calculate accuracy (e.g., Ellis, 1987; Crooks, 1989), it was found that specific measures such as measuring verb tense or the use of articles were not sensitive in finding significant differences among different planning conditions. Therefore, this study decided to apply the global measure rather than specific ones. Another reason for opting for this general measure was that the task in the present paper was not designed to elicit the use of any specific linguistic forms.

Fluency

Two aspects of fluency were measured in this study: speech rate and repair fluency. Speech rate was calculated by dividing the total number of syllables in unpruned speech (i.e., including self-repetitions, self-repairs, replacement) by the total speaking time (Yuan & Ellis, 2003). Repair fluency was measured by the average number of filled pauses (e.g., *um, er*) per minute (Tavakoli, 2016).

© 2021. John Benjamins Publishing Company
All rights reserved

Results

One-way ANOVA tests were conducted, followed by post-hoc pairwise *t*-tests when statistically significant differences were found. Eta-squared (η^2) was used to measure effect sizes of the one-way ANOVA tests. Bonferroni correction was used to adjust the *p*-value obtained from the pairwise *t*-tests. Before performing the one-way ANOVA tests, Shapiro-Wilk tests were employed to check the normality of the data, and Levene's tests were performed to check the homogeneity of variance. Results of the Shapiro-Wilk tests showed that all variables were normally distributed ($p > 0.05$) and the Levene's tests confirmed that the variances were equal ($p > 0.05$). Table 3 shows the descriptive statistics of the means and standard deviations of the chosen CAF measures.

Table 3. Descriptive statistics of means (standard deviations) of all measures

	Planning condition			
Measures	**NP**	**PP**	**OP**	**FP**
Complexity				
words/AS	6.60 (0.49)	7.34 (0.75)	8.88 (1.65)	8.57 (0.68)
subord.clauses/AS	0.16 (0.06)	0.27 (0.10)	0.34 (0.10)	0.33 (0.05)
Guiraud's Index	6.96 (0.26)	6.92 (0.70)	7.63 (0.76)	7.29 (0.70)
Accuracy				
%error-free AS	77.17 (0.12)	73.67 (0.12)	59.50 (0.13)	64.00 (0.05)
Fluency				
syllables/min	112.28 (29.19)	117.22 (33.79)	104.40 (30.63)	125.40 (15.48)
filled pauses/min	7.93 (3.52)	6.43 (1.58)	6.23 (3.07)	5.56 (2.19)

subord.clauses = subordinate clauses; % = percentage; min = minute

Complexity

For grammatical complexity, one-way ANOVA tests showed statistically significant differences regarding the average number of words per AS-unit ($F(3, 44) = 6.82$, $p < 0.01$, $\eta^2 = 0.51$) and the average number of subordinate clauses per AS-unit ($F(3, 44) = 6.51$, $p < 0.01$, $\eta^2 = 0.49$). Post-hoc pairwise *t*-tests showed that the unpressured on-line planning group and the full-planning group produced significantly longer AS-units than the non-planning group, with $p < 0.01$ and $p < 0.05$ respectively. The two groups also produced significantly more subordinate clauses than the non-planning group ($p < 0.01$). No statistically significant differences were found regarding Guiraud's Index ($F(3, 44) = 1.62$, $p = 0.22$, $\eta^2 = 0.20$).

© 2021. John Benjamins Publishing Company
All rights reserved

Accuracy

For the percentage of error-free AS-units, a statistically significant difference was determined by the one-way ANOVA test ($F(3,44) = 3.30$, $p < 0.05$, $\eta^2 = 0.33$). Post-hoc pairwise *t*-test showed that the unpressured on-line planning group produced significantly less accurate AS-units than the non-planning group ($p < 0.05$) and the pre-task planning group ($p < 0.05$).

Fluency

No statistically significant differences were found between the four groups in terms of speech rate ($F(3,44) = 0.59$, $p = 0.63$, $\eta^2 = 0.08$) and filled pauses ($F(3,44) = 0.82$, $p = 0.50$, $\eta^2 = 0.11$). This suggested that task planning operationalized in the present study did not raise fluency though the non-planning group produced more filled pauses on average.

Trade-off effect

Pearson's correlation was used to explore whether any trade-off effects occurred across the four groups. Table 4 gives the complete matrix of Pearson's correlation coefficient.

Table 4. Pearson correlation coefficient matrix of all measures ($N = 48$)

		1	2	3	4	5	6
1.	words/AS	–					
2.	subordination ratio	0.85***	–				
3.	Guiraud's Index	0.24	0.18	–			
4.	%error-free AS	−0.64***	−0.66***	0.01	–		
5.	syllables/min	0.06	0.17	0.29	0.04	–	
6.	pauses/min	0.11	−0.14	−0.28	−0.10	−0.22	–

*** $p < 0.001$

Results showed that there were no significant correlations between accuracy and fluency, nor between complexity and fluency. For accuracy and complexity, a negative correlation was found between the percentage of error-free AS-units and the average number of words per AS-unit with a large effect size ($r = -0.64$, $p < 0.001$). It was also found that the percentage of error-free AS-units was negatively correlated with the subordination ratio with a large effect size ($r = -0.66$, $p < 0.001$).

© 2021. John Benjamins Publishing Company
All rights reserved

To further explore whether any trade-off effects occurred within each group, several Pearson's correlation analyses were conducted. Results showed that, in the unpressured on-line planning group, the average number of words per AS-unit was negatively correlated with the percentage of error-free AS-units and the effect size was large ($r = -0.81$, $p = 0.05$). Additionally, in this group, a negative correlation was found between the subordination ratio and the percentage of error-free AS-units with a large effect size ($r = -0.80$, $p = 0.05$). For other groups, no statistically significant correlations were found.

The retrospective interviews

Table 5 presents the descriptive statistics of the strategies adopted by the participants. For all groups, the percentage of participants who used each type of strategy was calculated with the actual number in the brackets. The total number of strategies used in each group and the total number of strategies used in each category were provided. It should be mentioned that participants in the non-planning group and the unpressured on-line planning group had no time for *organizational planning, problem identification, prediction, writing for retrieval*, and *rehearsing*.

Table 5. Descriptive statistics on planning strategies

Strategy types	NP $n = 12$	PP $n = 12$	OP $n = 12$	FP $n = 12$	Total
1. Metacognitive strategies					
1.1 Organizational planning	not available	100% (12)	not available	100% (12)	24
1.2 Problem identification	not available	8.3% (1)	not available	8.3% (1)	2
1.3 Production monitoring	42% (5)	58% (7)	83% (10)	58% (7)	29
1.4 Style monitoring	0	8.3% (1)	16.7% (2)	16.7% (2)	5
1.5 Prediction*	not available	8.3% (1)	not available	8.3% (1)	2
Total	**5**	**22**	**12**	**23**	**62**
2. Cognitive strategies					
2.1 Writing for retrieval	not available	50% (6)	not available	33.3% (4)	10
2.2 Rehearsing	not available	58% (7)	not available	8.3% (1)	8
2.3 Paraphrasing*	8.3% (1)	25% (3)	16.7% (2)	25% (3)	9
2.4 Translating	16.7% (2)	8.3% (1)	0	33.3% (4)	7
Total	**3**	**17**	**2**	**12**	**34**

© 2021. John Benjamins Publishing Company
All rights reserved

Table 5. *(continued)*

Strategy types	NP $n=12$	PP $n=12$	OP $n=12$	FP $n=12$	Total
3. Social / affective strategies					
3.1 Empathizing with the listener	42% (5)	50% (6)	66.7% (8)	50% (6)	25
3.2 Avoiding embarrassment*	8.3% (1)	8.3% (1)	8.3% (1)	33.3% (4)	7
3.3 Lowering anxiety	75% (9)	16.7% (2)	8.3% (1)	8.3% (1)	13
Total	**15**	**9**	**10**	**11**	**45**

Similarities and differences between the pre-task planning group and the full-planning group could be found regarding pre-task planning strategies. The pattern of metacognitive strategies was quite similar, with every participant reporting that they used part of the time for *organizational planning*. By contrast, participants in the pre-task planning group reported that they relied on *writing for retrieval* and *rehearsing*, with 50% and 58% respectively, but only 33.3% of the participants in the full-planning group believed that writing down notes was important and most of them did not rehearse while preparing for the task.

For metacognitive strategies, *production monitoring* was popular among all groups, ranging from 42% (NP) to 83% (OP). For cognitive strategies, participants from all groups reported that they tried to use *paraphrasing* when they had difficulty with certain lexical choices, though the percentage remained fairly low.

For social/affective strategies, it seemed important for the participants to empathize with their partners, which ranged from 42% (NP) to 66.7% (OP). An interesting phenomenon could be observed when participants tried to lower their anxiety. For the non-planning group, 75% of the participants reported that they felt extremely nervous but the percentage was much lower among other groups, with only 1 participant reporting being anxious in the unpressured on-line planning group and the full-planning group.

To compare the use of different types of strategies in the four planning conditions, the total number of the employed metacognitive strategies, cognitive strategies, and social/affective strategies were divided by the total number of participants in each group to establish the average number of adopted strategies. The results are presented in Table 6.

As can be inferred from the table, in total, the participants relied more on metacognitive strategies than cognitive strategies: participants used 1.29 types of metacognitive strategies on average and 0.71 types of cognitive strategies. An interesting pattern was found regarding the use of social/affective strategies. In the present study, participants used more social/affective strategies (Mean = 0.94)

© 2021. John Benjamins Publishing Company
All rights reserved

Table 6. Average number of strategies employed in each group

	Metacognitive strategies	Cognitive strategies	Social/affective strategies
All groups (*N*=48)			
Total	**62**	**34**	**45**
Mean	1.29	0.71	0.94
Non-planning group (*N*=12)			
Total	**5**	**3**	**15**
Mean	0.42	0.25	1.25
Pre-task planning group (*N*=12)			
Total	**22**	**17**	**9**
Mean	1.83	1.42	0.75
Unpressured on-line planning group (*N*=12)			
Total	**12**	**2**	**10**
Mean	1.00	0.17	0.83
Full-planning group (*N*=12)			
Total	**23**	**12**	**11**
Mean	1.92	1.00	0.92

than cognitive strategies (Mean = 0.71). Within each group, participants without planning opportunities did not rely on metacognitive and cognitive strategies, with only 0.42 types and 0.25 types used respectively. This group, however, used more social/affective strategies than the other groups (Mean = 1.25). For the pre-task planning group, participants relied equally on metacognitive (Mean = 1.83) and cognitive (Mean = 1.42) strategies. By contrast, participants from the unpressured on-line planning group employed metacognitive strategies (Mean = 1.00) much more frequently than cognitive strategies (Mean = 0.17). Similarly, the number of metacognitive strategies (Mean = 1.92) employed in the full-planning group almost doubled compared to cognitive strategies (Mean = 1.00). Participants from the three groups also frequently used social/affective strategies, with the average number ranging from 0.75 types to 0.92 types.

© 2021. John Benjamins Publishing Company
All rights reserved

Discussion

Effects of task planning on L2 performance (RQ 1)

The first research question concerns the effects of different planning conditions on L2 performance in a dialogic task condition. The results are somewhat surprising when comparing the performance of the non-planning group and the pre-task planning group: no significant differences were found between the two groups in any of the three dimensions. This differs widely from previous studies on pre-task planning (e.g., Wendel, 1997; Ortega, 1999; Yuan & Ellis, 2003; Sangarun, 2005; Wang, 2014, etc.). Skehan (2009) argued that pre-task planning helps complexify syntactic structure since it allows sufficient time for learners to attend to the Conceptualizer stage. It also benefits lexical complexity and fluency as learners will have the time to recall existing lexical resources and syntactic rules prior to task performance and thus prime lexical and syntactic retrieval at the Formulator stage (Skehan, 2009). However, these effects of pre-task planning were not found in this research. One possible explanation is the influence of interaction. In monologic tasks, the availability of pre-task planning is important because what is prepared can be used for the actual performance. In other words, monologic task performance is more predictable, making it easier to prepare. The dialogic task in this study, however, caused more unforeseen situations that required the participants to change what was originally planned. This was reported in one of the interviews:

> *Sometimes my partner said something that I wanted to say. So, I had to think about something new at the moment. It's really improvised, I think.* (PP-3-2)

As pointed out by the participant, it is hard to predict and anticipate the contributions, strategies and problems of a partner. This means that all the advantages brought by pre-task planning in monologic tasks may have been diluted in interactive tasks because students sometimes will revisit the Conceptualizer and Formulator stages, leading to a larger proportion of improvisation, which resembles the non-planning condition.

For the unpressured on-line planning group, participants produced syntactically more complex structures than the non-planning group. This is consistent with previous studies (e.g., Yuan & Ellis, 2003; Ahmadian, 2012; Wang, 2014). To produce complex speech, several conditions need to be met. First, learners have to be willing to take risks because it is possible to make mistakes when they try to stretch and reconstruct their interlanguage (Foster & Skehan, 1996). Participants in this group reported that they did not feel anxious at all because they knew that

© 2021. John Benjamins Publishing Company
All rights reserved

they would have plenty of time producing their language, which made them more willing to speak more:

> *You told us that we had unlimited time. So, I knew I could take my time and say whatever I want to say. Even if I forget something, I could always go back and check, and say it later. So, I was pretty relaxed. […] Also, I knew this was not a language test because normally in a language test the time was limited, but I had plenty of time just now.* (OP-2-2)

Second, learners need sufficient time to search for abstract syntactic rules that are needed to structure a more complex speech. Skehan (1998) argued that the abstract language patterns are only available in the rule-based system, which requires more cognitive resources. The unlimited on-line planning time might have eased the pressure on learners' attentional resources and made it easier for them to mobilize their rule-based system. It is surprising that, although they produced more syntactically complex speech, lexical complexity did not increase in this group. A similar result was found in Yuan and Ellis (2003). They argued that this could be a reflection of a trade-off effect between grammatical complexity and lexical complexity. In other words, they believed that it is possible that when learners try to search for more complex grammatical forms, they do not have extra attentional resources to find advanced lexical resources. This negative correlation between grammatical complexity and lexical complexity, however, was not found in the current study. Another reason may be that the participants performed on the same topic in the current research. Studies on lexical complexity (e.g., Vermeer, 2000; Yu, 2010) suggested that lexical diversity and richness are strongly associated with the assigned topics. Since the participants all performed the same task, it is possible that their choice of words will not differ too much.

Interestingly, no statistically significant differences were found regarding speech rate. Previous studies have shown a detrimental effect of unpressured online planning on fluency (e.g., Yuan & Ellis, 2003; Ahmadian, 2012) but it was not found in the present study. This result can be explained by the presence of the interlocutor. Michel (2011) showed that having an interlocutor can strongly influence the participants' performance. Gilabert et al. (2011) also reported that learners tended to adjust their performance for their interlocutor in dialogic tasks. Regarding speech rate, the influence of the interlocutor is twofold. First, the existence of the interlocutor may have pushed participants to speak faster. Tavakoli (2016) found that the participants were more willing to speak faster when they took their interlocutors into consideration. It was found in this study that when the participants spoke too slow, they felt worried about being embarrassed in front of their partner or wasting their partner's time. In fact, it can be observed in the participants' performance that some of them immediately apologized after

© 2021. John Benjamins Publishing Company
All rights reserved

they paused or repeated themselves more than once. By contrast, learners can speak slowly without worrying about their interlocutors in monologic tasks. Second, the interlocutor may provide help when the speaker pauses (Michel, 2011). In the current study, the participants tried to immediately help their partner by continuing their speech or providing vocabulary when they paused. On the contrary, this advantage is not available in monologic tasks. The other fluency measure, filled pauses, indicated that all the participants actively engaged in on-line planning regardless of planning conditions. This is pointed out in Tavakoli and Foster (2008) and Michel (2011) that speakers in dialogic task conditions will take the time to plan on-line when it is their partner's turn. Tavakoli (2016) suggested that filled pauses can be another indicator of considering the need of the interlocutor since the speakers may be trying to convey their speech in a more precise way, being it form-precise or meaning-precise. It seems that the influences of interaction on fluency were not affected by planning conditions in this research.

For accuracy, the unpressured on-line planning group produced significantly less accurate speech than the non-planning group and the pre-task planning group, which differs from previous studies (e.g., Yuan & Ellis, 2003; Ahmadian, 2012). Ahmadian (2012) explained that the unlimited performing time generated sufficient self-monitoring behaviors that were necessary for producing accurate L2 speech. In the dialogic task condition, however, the self-monitoring process may be constrained by the quick turn-taking behaviors. That is to say, due to the presence of the interlocutor, learners have to speak faster and may not have enough time to engage in self-monitoring. A more plausible explanation of the findings regarding accuracy is provided in the next section in which a trade-off effect is discussed.

Unexpectedly, results obtained from the full-planning group were similar to the unpressured on-line planning group. Participants produced significantly more complex language than the non-planning group. They also made more mistakes (Mean = 64.00) than the non-planning group (Mean = 77.17) though the difference was not statistically significant. The above-mentioned limited function of pre-task planning in dialogic tasks can explain the findings. In other words, participants in this group seem to have only benefited from unpressured on-line planning.

Trade-off effect occurred in the current study (RQ 2)

A trade-off effect was found between syntactic complexity and global accuracy as determined by Pearson's correlation coefficient. This means that the participants who made fewer mistakes also produced less complex speech, which is the case of the non-planning group. On average, participants in this group produced

© 2021. John Benjamins Publishing Company
All rights reserved

the most accurate but relatively less complex speech, although the Pearson analysis did not find negative correlations between complexity and accuracy in this group. Skehan (1998) argued that accuracy reflects the control of already-known linguistic knowledge, while complexity concerns restructuring that knowledge. It is possible that when the participants tried to create more complex language, their limited attentional resources did not allow them to monitor the accuracy of their language simultaneously. This limited attentional capacity (Skehan, 1998, 2009, 2018) is supported by the retrospective interviews. Most participants from the non-planning group stated that they had experienced anxiety because they did not have time to plan their ideas and language at the same time. Consequently, they found a conflict between organizing their answer, which requires metacognitive attention and searching for linguistic items to express that answer, which requires cognitive resources. They somehow found a way to solve this conflict, namely to use the simplest structure:

> *I was very nervous and I think my partner was as well. […] I mean I didn't have enough time to digest everything on the paper. So, I had to make a choice and organize my reasons while I was talking. I had never done this before cuz normally we would have time to read materials. So, this was quite difficult for me. So, I think I should just use simple words.* (NP-2-1)

This suggests that when no planning time was available, the participants could not attend to all dimensions of performance simultaneously and had to choose to prioritize meaning over form for communicative purposes.

In contrast to the non-planning group, participants from the unpressured on-line planning group produced more complex language but this went at the expense of accuracy. The negative correlations between the grammatical complexity measures and the accuracy measure in this group further supported this. However, this trade-off effect is not in line with previous studies such as Yuan and Ellis (2003) and Ahmadian (2012) as they reported a trade-off effect between accuracy and fluency but not complexity and accuracy. This difference can be explained by the fact that the previous studies employed monologic tasks, which did not provide the above-mentioned advantages of having an interlocutor. This indicates that the trade-off effect between accuracy and fluency may have been eliminated in the dialogic task. Another interesting indication is that in dialogic conditions, simply giving learners unlimited time to perform the tasks may not be sufficient to achieve the balance between complexity and accuracy.

To explain the trade-off effect between complexity and accuracy in the unpressured on-line planning group, we now turn to the retrospective interviews. It was found that when the participants had sufficient time to complete the task, they chose to prioritize the content of their speech, such as how to convince their

© 2021. John Benjamins Publishing Company
All rights reserved

partners and how to explain themselves better. This, again, indicates that in dialogic tasks, learners tend to focus on meaning not form. Another reason is their attitude towards making mistakes in real-time communication. Many participants in the group reported that they did not mind making mistakes during the performance, especially when the mistake would not influence the smoothness of that conversation.

Strategies employed during pre-task planning (RQ 3)

The third research question sets out to find the patterns of strategies employed during pre-task planning. This concerns the pre-task planning group and the full-planning group. For the pre-task planning group, the number of employed metacognitive and cognitive strategies is almost equal. This is in line with Ortega (2005) in which a balance between the two categories was achieved. The full-planning group, though employed the highest number of metacognitive strategies, relied much less on cognitive strategies. This is in contrast with findings from Ortega (2005) and the series of studies conducted in O'Malley and Chamot (1990), which reported that the learners adopted more cognitive strategies than metacognitive strategies.

Interesting patterns can be found when comparing specific strategies used by the two groups. For metacognitive strategies during pre-task planning, all participants reported using *organizational planning*. In Ortega's study, the most frequently used strategy was *summarizing* but none of the participants used this strategy in the current study. O'Malley and Chamot (1990) pointed out that the choice of strategies is strongly affected by the types of tasks. The interactive task in this study is highly unstructured, especially when taking into account the influence of the interlocutor. This makes it almost impossible to write a summary prior to the task performance. On the other hand, *organizational planning* is useful because the dialogic task did not provide a clear organization of content. Another metacognitive strategy that should be mentioned is the emerged strategy *prediction*. Two of the participants reported that they would predict the answer from their partners during pre-task planning and would prepare their answers based on the prediction. It is clear that this strategy does not always produce satisfactory effects. If the prediction is proven correct, students' preparation could be useful to some extent, but if the prediction is wrong, students may feel anxious because they have to reconsider their answers while they are talking.

Regarding cognitive strategies, the pre-task planning group frequently used *writing for retrieval* and *rehearsing*. This is interesting because if we turn back to the performance produced by this group, we can find that the effects of rehearsal and retrieval on interactive task performance are limited. In addition, Skehan

© 2021. John Benjamins Publishing Company
All rights reserved

(2018) suggested that pre-task planning only provides what he called "shallow rehearsal", which is constrained by retention issues. As a result, it is difficult for the learners to transfer what was originally planned into the actual performance.

The full-planning group, by contrast, only reported limited use of *writing for retrieval* and *rehearsing*. It seems that sufficient performing time affected the participants' choice of planning strategies. Oxford (1990) argued that among cognitive strategies, the most important ones are *practicing* (e.g., *repeating, rehearsing*) and *retrieving*. However, when giving learners unlimited time to perform the task, they may believe that being fully prepared is no longer necessary and that they can search for necessary words and syntactic rules during the performance.

Strategies employed during on-line planning (RQ 4)

Yuan and Ellis (2003) and Ellis (2018) argued that on-line planning cannot be excluded from task performance and can only be pressured or unpressured. In this view, on-line planning strategies concern all groups in this study. For metacognitive strategies used during task performance, the most popular one was *production monitoring*. This is in line with the findings reported in Ortega's study. A noticeable pattern is that the unpressured on-line planning group frequently used *production monitoring*. This result is striking because this group produced the least accurate performance. One possible explanation is that instead of monitoring the accuracy of their language, they paid more attention to monitoring the content clearness of their speech as one of the participants reported that she thought more about finding better expressions to precisely convey her ideas. Interestingly, Pang and Skehan (2014) coded *plan how to express better* as an independent strategy and reported that this strategy showed a negative relationship with accuracy and complexity, which suggested that their participants planned too ambitiously. Similarly, in this study, when the participants were given unlimited time to search for more complex expressions, they seemed to be less realistic about their proficiency level and planned too much.

Fewer cognitive strategies were used during on-line planning. In each group, no more than three participants reported that they used *paraphrasing*. Oxford (1990, p. 50) coded *paraphrasing* as *adjusting or approximating the message*. She suggested that this strategy is often used to compensate learners' lack of lexical knowledge. It is possible that the vocabulary required in the dialogic task was not challenging and thus the participants did not have to paraphrase. For *translating*, only seven participants in total reported that they had translated their speech from Chinese to English. This can be explained by the participants' attitude towards translation: they believed that using translation is a sign of incompetence in English.

© 2021. John Benjamins Publishing Company
All rights reserved

Social/affective strategies were more frequently used than cognitive strategies during the performance. This is not in line with Ortega's study but can be explained by the different nature of monologic tasks and dialogic tasks. Though participants in Ortega's study were asked to perform the task in front of a listener, there were no interactions between them, which was different from having an actual interlocutor. Furthermore, the chosen dialogic task required the participants to negotiate with each other and make decisions together, meaning that they had to cooperate with each other to achieve better performance. As a result, more than half of the participants chose to empathize with their listeners. Oxford (1990) stated that empathy is the key to communication because it helps the speakers understand their interlocutor's perspective. Griffiths (2003), a more recent study on language learning strategies, reported that *strategies relating to interaction with others* were emphasized by language learners and therefore worth more attention from language teachers. These findings suggest that it is important to promote the use of social/affective strategies to achieve successful communications. In this respect, the use of dialogic tasks should be encouraged in language classrooms.

Having an actual interlocutor has led to the use of *avoidance of embarrassment*. As some participants reported, they were worried about experiencing silence while talking to their partners. They would therefore try to speak more in order to not make each other feel awkward. Finally, it seems that giving students planning time, whether pre-task planning or on-line planning, helps ease their anxiety while performing the task: nine of the participants from the first group stated that they felt really nervous and insecure when they did not have enough time to prepare and perform the task, but the number was much lower in the other groups. O'Malley and Chamot (1990) and Oxford (1990) both emphasized that learning to reduce anxiety in a task is important because when there is considerable anxiety, it may stop learners from learning the language. This holds true for specific language skills such as speaking. A recent study conducted by Syafryadin (2020) found that controlling emotional state by reducing anxiety was helpful in motivating the participants to continue learning speaking skills. In this regard, it is necessary to allow some planning time for the learners, otherwise, it will impose too much anxiety and consequently discourage them to perform better.

Learners' perceptions of task preparedness (RQ 5)

The last question concerns the participants' perceptions of pre-task planning and unpressured on-line planning. Most of the participants demonstrated a positive attitude towards pre-task planning. Two major advantages were reported by the participants. First, pre-task planning enabled them to better digest new informa-

© 2021. John Benjamins Publishing Company
All rights reserved

tion and helped them find related expressions in advance to facilitate communication:

> *I think my answer was a bit messy. If I had more time to prepare, I think it could help me understand the information better so that my partner could understand me better.* (PP-4-1)

Second, pre-task planning gave them the chance to take notes. Some participants believed that note-taking helped them remember the new information and unfamiliar words better. These two advantages seem to associate with task material familiarity and task type familiarity proposed in Skehan (1996). Bui (2014) further divided familiarity into three categories: topic familiarity, schematic familiarity, and task familiarity. It is possible that pre-task planning familiarized the participants with the topic and type of the interactive task as they tried to digest and memorize the new information contained in the task and task instructions. Also, it provided time for the participants to retrieve previous linguistic and content knowledge stored in their long-term memory, which is associated with schematic familiarity.

Nevertheless, there are participants who believed that taking notes made their performance less natural:

> *I think I relied too much on the note. I tried very hard to recall my notes but it made the conversation more like a speech. Just not that natural.* (PP-3-2)

Similarly, some participants in Ortega's study reported that "extemporaneous performance feels good" and they felt that their speech was false after being fully prepared. This suggests that individual differences may exist in understanding successful communication. Some learners believe that it is necessary to be fully prepared in order to better convey their ideas and achieve smooth communication even if the preparation will make their speech less natural. But other learners value naturalness and believe that taking notes will lead to over-preparation.

For on-line planning, participants' perceptions were quite mixed. Some participants believed that unpressured on-line planning was beneficial:

> *For me, it's better to have unlimited performing time. Because normally, I would speak less fluently at the beginning. But I could speak more frequently and logically as time went by.* (NP-1-1)

It seems that a mental warm-up effect could be boosted by unpressured on-line planning, which is not found in previous studies. According to Brown (2006), mental warm-up in language learning refers to the activation of learners' background knowledge that could be utilized to understand new information. This is

© 2021. John Benjamins Publishing Company
All rights reserved

similar to the activation of schematic knowledge during pre-task planning. The difference lies in the fact that pre-task planning does not provide the opportunity to articulate the speech. But for participants with the unpressured on-line planning opportunity, it is possible that they activated previous linguistic knowledge including information related to articulation in the first few minutes of their performance so that they gradually became more fluent. This warm-up effect also prepared them in emotional terms: some participants reported that they gradually felt relaxed as they spoke more.

Another advantage brought by unpressured online planning is the opportunity to refine the output. Some participants stated that because of the unlimited performing time, they could logically organize their answers while they were talking. And others said that due to the unlimited time, they had the opportunity to come up with more pieces of information, which made their answers more complete.

By contrast, some participants reported that the unlimited time made them less focused:

> *I think if I speak for, like, 10 minutes, I would definitely become tired. I mean, I would not be interested in performing the task anymore. I think it would be more efficient if there was a time limit. I could just give my choice and the reasons immediately instead of casually chatting with my partner.* (FP-2-1)

Another drawback of unpressured on-line planning is that it makes the task less challenging. Some participants, unlike others, preferred challenges while performing the task. This can be seen as a reflection of individual learning styles. Many studies have found that personality type can strongly affect the choice of learning strategies (e.g., Oxford, 1989, 2003; Chen, 2009). Those participants who embrace challenges believed that the unlimited time made the task less demanding. As a result, they lost the motivation to perform it. This is in line with Dörnyei's (2019) theory on task motivation, which argues that it is necessary to provide an "optimal level" of challenge in order to boost language learning motivation. Furthermore, the disfavor of unlimited performing time can be explained by the educational context in China. Since the participants in this study are all Chinese EFL learners, they are more used to pressured performance because English tests in China usually provide relatively short completion time. These participants reported that they were afraid of losing the ability to cope with the pressure if they always speak English without a time limit.

Finally, one benefit that is shared by both types of planning is the effect of lowering anxiety. Pre-task planning made the participants less nervous since it gave them the time to organize the content and prepare unfamiliar words in advance. Surprisingly, Ortega (2005) found that having time to prepare in advance

© 2021. John Benjamins Publishing Company
All rights reserved

increased the pressure as some of the participants believed that more time to prepare means more time to be worried. For unpressured on-line planning, participants reported that they felt more at ease because they did not have to rush to a conclusion and could modify their answer until they felt confident to do so. This is particularly true for dialogic tasks given that the learners need time to react to their interlocutors' responses and have to adjust their answers accordingly.

Conclusion

Very few studies have been carried out to explore the effects of task preparedness in dialogic task conditions (Ellis, 2009, 2018). This study makes a contribution in this respect. The overall results suggest that the effects of pre-task planning are limited in dialogic task conditions. This is because the interactions make the performance less predictable and thus difficult to prepare in advance. For unpressured on-line planning, it benefits syntactic complexity. A trade-off effect is found between complexity and accuracy, which may be a result of the limited attentional resources proposed in the Trade-off Hypothesis (Skehan, 1998, 2009). Interestingly, the influence of interactions seems to eliminate the trade-off effect between accuracy and fluency that has been reported in previous studies (e.g., Yuan & Ellis, 2003; Ahmadian, 2012).

Following Ortega (2005), this study has qualitatively investigated the strategies employed by L2 learners during pre-task planning and on-line planning. Overall, metacognitive strategies are the most popular for all planning conditions. For pre-task planning, learners prefer to utilize the opportunity for organizing ideas. For on-line planning, learners use the opportunity to monitor their speech but more attention is paid to meaning not form. This suggests that in dialogic tasks, learners tend to prioritize meaning for the sake of successful communications. Cognitive strategies are less popular among L2 learners, especially during on-line planning. For pre-task planning, learners prefer to retrieve their linguistic knowledge and rehearse. But these strategies are not effective in this study because (1) what is prepared cannot always be used in dialogic tasks; (2) what can be used cannot always be transferred to the actual performance due to retention issues (Skehan, 2018). Social/affective strategies, on the other hand, are favored in dialogic task conditions since learners tend to take their interlocuter's needs into consideration.

For L2 learners' perceptions of task preparedness, it is found that most of the learners value pre-task planning given its possible effects on helping learners familiarize themselves with the topic and type of the task and active schematic knowledge stored in their long-term memory. However, individual learning styles

© 2021. John Benjamins Publishing Company
All rights reserved

may influence these perceptions as some L2 learners value naturalness over preparedness. Turning to unpressured on-line planning, some learners believe that it helps refine their performance. But others hold that the unlimited time will lower the efficiency of communication and is likely to eliminate a sense of challenge. Both planning conditions are perceived positively regarding their effects on reducing anxiety.

Limitations and future directions

This study is exploratory in nature given the lack of investigations on dialogic tasks in planning research. There are limitations regarding the experiment and the interviews. For the experiment, the number of participants involved in this study is limited. If more participants are involved, it will be more warranted to make generalizations. Second, more measures could have been employed regarding fluency, such as mean length of run. Finally, it will make the results more rigorous if inter-rater reliability is calculated regarding the CAF measures, especially for accuracy. This is due to that raters do not always reach agreements on the accurate use of the target language (Michel, 2017).

With regard to the retrospective interview, the coding scheme could be further developed. For example, the other categories that are mentioned in Oxford (1990) can be included. The coding scheme will be more sophisticated if categories that specifically target at planning are developed. Currently, the categories are derived from language learning strategies but not planning strategies.

For future research, there are many possibilities. For example, to date, most studies on task planning only focus on its impact on task performance. It will be insightful to add studies exploring the effect of task planning on L2 acquisition. Also, the difference between guided and unguided planning in dialogic task conditions can be investigated to see if guided planning will lead to improvements in accuracy. Furthermore, studies on task preparedness typically use the CAF measures, but other measures can be utilized to investigate L2 performance such as pragmatic and discourse features (Ellis, 2018). Finally, it will be interesting to employ different types of tasks (e.g., narrative tasks, decision-making tasks, information-gap tasks, etc.) in a study to explore whether task planning has any effects on L2 performance based on the task types.

© 2021. John Benjamins Publishing Company
All rights reserved

References

Ahmadian, M. J. (2012). The effects of guided careful online planning on complexity, accuracy and fluency in intermediate EFL learners' oral production: The case of English articles. *Language Teaching Research*, *16*(1), 129–149. https://doi.org/10.1177/1362168811425433

Brown, A. L., Bransford, J. D., Ferrara, R. A., & Campione, J. C. (1983). Learning, remembering, and understanding. In J. H. Flavell & M. Markman. (Eds.). *Carmichael's manual of child psychology* (pp. 77–166). John Wiley & Sons.

Brown, S. (2006). *Teaching listening*. Cambridge University Press.

Bui, G. (2014). Task readiness: Theoretical framework and empirical evidence from topic familiarity, strategic planning, and proficiency levels. In P. Skehan. (Ed.), *Processing perspectives on task performance* (pp. 63–94). John Benjamins.

Chen, M. L. (2009). Influence of grade level on perceptual learning style preferences and language learning strategies of Taiwanese English as a foreign language learners. *Learning and Individual Differences*, *19*(2), 304–308. https://doi.org/10.1016/j.lindif.2009.02.004

Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge University Press.

Crookes, G. (1989). Planning and interlanguage variation. *Studies in Second Language Acquisition*, *11*, 367–383. https://doi.org/10.1017/S0272263100008391

Dörnyei, Z. (2019). Task motivation: What makes an L2 task engaging? In Z. Wen & M. J. Ahmadian. (Eds.). *Researching L2 task performance and pedagogy: In honor of Peter Skehan*. (pp. 53–66). John Benjamins. https://doi.org/10.1075/tblt.13.04dor

Elder, C. A., & Iwashita, N. (2005). Planning for test performance: Does it make a difference? In R. Ellis. (Ed.), *Planning and task performance in a second language* (pp. 219–238). John Benjamins. https://doi.org/10.1075/lllt.11.14eld

Ellis, R. (1987). Interlanguage variability in narrative discourse: Style shifting in the use of the past tense. *Studies in Second Language Acquisition*, *9*, 12–20. https://doi.org/10.1017/S0272263100006483

Ellis, R. (2005). Planning and task-based performance: Theory and research. In R. Ellis. (Ed.) *Planning and task performance in a second language* (pp. 3–34). Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.11.03ell

Ellis, R. (2009). The differential effects of three types of task planning on the fluency, complexity and accuracy in L2 oral production. *Applied Linguistics*, *30*, 474–509. https://doi.org/10.1093/applin/amp042

Ellis, R. (2018). *Reflections on task-based language teaching*. Multilingual Matters.

Ellis, R., Skehan, P., Li, S., Shintani, N., & Lambert, C. (2019). *Task-based language teaching: Theory and practice*. Cambridge University Press. https://doi.org/10.1017/9781108643689

Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, *18*, 299–323. https://doi.org/10.1017/S0272263100015047

Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics*, *23*(3), 354–375. https://doi.org/10.1093/applin/21.3.354

Gilabert, R., Barón, J., & Levkina, M. (2011). Manipulating task complexity across task types and modes. In P. Robinson. (Ed.), *Second language task complexity: Researching the Cognition Hypothesis of language learning and performance* (pp. 105–138). John Benjamins. https://doi.org/10.1075/tblt.2.10ch5

© 2021. John Benjamins Publishing Company
All rights reserved

Griffiths, C. (2003). Language learning strategy use and proficiency: The relationship between patterns of reported language learning strategy (LLS) use by speakers of other language (SOL) and proficiency with implications for the teaching/learning situation (Unpublished doctoral dissertation) University of Auckland.

Guiraud, P. (1959). *Problèmes et méthodes de la statistique linguistique*. Reidel.

Kawauchi, C. (2005). The effects of strategic planning on the oral narratives of learners with low and high intermediate L2 proficiency. In R. Ellis. (Ed.), *Planning and task performance in a second language* (pp. 143–164). John Benjamins. https://doi.org/10.1075/lllt.11.09kaw

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. The MIT Press.

Michel, M. (2011). Effects of task complexity and interaction in L2 performance. In P. Robinson. (Ed.), *Second language task complexity: Researching the Cognition Hypothesis of language learning and performance* (pp. 141–174). John Benjamins. https://doi.org/10.1075/tblt.2.12ch6

Michel, M. (2017). Complexity, Accuracy and Fluency (CAF). In S. Loewen & M. Sato. (Eds.), *The Routledge handbook of instructed second language acquisition* (pp. 50–68). Routledge. https://doi.org/10.4324/9781315676968-4

Michel, M., Kuiken, F., & Vedder, I. (2007). The influence of complexity in monologic versus dialogic tasks in Dutch L2. *IRAL – International Review of Applied Linguistics in Language Teaching*, *45*(3), 241–259. https://doi.org/10.1515/iral.2007.011

O'Malley, J. M. & Chamot, A. U. (1990). *Learning strategies in second language acquisition*. Cambridge University Press. https://doi.org/10.1017/CBO9781139524490

Ortega, L. (1999). Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition*, *21*(1), 109–148. https://doi.org/10.1017/S0272263199001047

Ortega, L. (2005). What do learners plan? Learner-driven attention to form during pre- task planning. In R. Ellis. (Ed.). *Planning and task performance in a second language* (pp. 77–109). Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.11.07ort

Oxford, R. (1989). Variables affecting choice of language learning strategies by university students. *The Modern Language Journal*, *73*(3), 291–300. https://doi.org/10.1111/j.1540-4781.1989.tb06367.x

Oxford, R. (1990). *Language learning strategies: What every teacher should know*. Newbury House.

Oxford, R. (2003). Language learning styles and strategies: Concepts and relationships. *International Review of Applied Linguistics in Language Teaching*, *41*(4), 271–278. https://doi.org/10.1515/iral.2003.012

Pang, F., & Skehan, P. (2014). Self-reported planning behavior and second language performance in narrative retelling. In P. Skehan. (Ed.), *Processing perspectives on task performance* (pp. 95–127). John Benjamins. https://doi.org/10.1075/tblt.5.04pan

Sangarun, J. (2005). The effects of focusing on meaning and form in strategic planning. In R. Ellis. (Ed.). *Planning and task performance in a second language* (pp. 111–141). John Benjamins. https://doi.org/10.1075/lllt.11.08san

Skehan, P. (1996). Second-language acquisition research and task-based instruction in J. Willis & D. Willis. (Eds.), *Challenge and change in language teaching* (pp. 17–30). Heinemann.

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

Skehan, P. (2009). Models of speaking and the assessment of second language proficiency. In A. Benati. (Ed.), *Issues in second language proficiency* (pp. 202–215). Continuum.

© 2021. John Benjamins Publishing Company
All rights reserved

Skehan, P. (2018). *Second Language Task-Based Performance: Theory, Research, Assessment*. Routledge. https://doi.org/10.4324/9781315629766

Skehan, P., & Foster, P. (2005). Strategic and on-line planning: The influence of surprise information and task time on second language performance. In R. Ellis. (Ed.). *Planning and task performance in a second language* (pp. 193–216). John Benjamins. https://doi.org/10.1075/lllt.11.12ske

Syafryadin, S. (2020). Students' strategies in learning speaking: Experience of two Indonesian schools. *Journal for Language and Foreign Language Learning*, *9*(1), 34–47.

Tavakoli, P. (2016). Fluency in monologic and dialogic task performance: Challenges in defining and measuring L2 fluency. *IRAL – International Review of Applied Linguistics in Language Teaching*, *54*(2): 133–150. https://doi.org/10.1515/iral-2016-9994

Tavakoli, P. & Foster, P. (2008). Task design and second language performance: The effect of narrative type on learner output. *Language Learning*, *58*(2), 439–473. https://doi.org/10.1111/j.1467-9922.2008.00446.x

Tavakoli, P., & Skehan, P. (2005). Strategic planning, task structure, and performance testing. In R. Ellis. (Ed.). *Planning and task performance in a second language* (239–273). John Benjamins. https://doi.org/10.1075/lllt.11.15tav

Vermeer, A. (2000). Coming to grips with lexical richness in spontaneous speech data. *Language Testing*, *17*(1), 65–83. https://doi.org/10.1177/026553220001700103

Wang, Z. (2014). On-line time pressure manipulations: L2 speaking performance under five types of planning and repetition conditions. In P. Skehan. (Ed.), *Processing perspectives on task performance* (pp. 27–62). John Benjamins. https://doi.org/10.1075/tblt.5.02wan

Wendel, J. (1997). Planning and second language narrative production (Unpublished doctoral dissertation). Temple University, Japan.

Yu, G. (2010). Lexical diversity in writing and speaking task performance. *Applied Linguistics*, *31*(2), 236–259. https://doi.org/10.1093/applin/amp024

Yuan, F., & Ellis, R. (2003). The effects of pre-task and on-line planning on fluency, complexity and accuracy in L2 monologic oral production. *Applied Linguistics*, *24*(1), 1–27. https://doi.org/10.1093/applin/24.1.1

Appendix

In this task, you are a human resource manager in an international high school. Now you are asked to hire a teacher for an English course at your school. The basic information of four candidates is listed as follows. You need to discuss with your partner and choose the best candidate for the job.

Candidate No.2

Name:	Li Lei
Educational Background:	M.A. in Education (Language), *The University of Oxford*, GPA: 3.5/4, 2016–2017
	B.A. in Education, *Beijing Normal University*, GPA: 3.9/4, 2012–2016
Work Experience:	2016-present, Teaching Assistant, *The University of Oxford*

© 2021. John Benjamins Publishing Company
All rights reserved

Candidate No.3

Name:	Lili Chou
Educational Background:	Ph.D. in Applied English Linguistics, *University of Cambridge*, 2017–2020
	M.A. in Applied English Linguistics, *University of Cambridge*, GPA: 3.5/4, 2016–2017
	B.A. in English, *Zhejiang University*, GPA: 3.9/4, 2012–2016
Work Experience:	2016-present, Teaching Assistant, *University of Cambridge*
	2015.06–2015.09, Teaching Assistant, *Zhejiang University*

Candidate No.4

Name:	Xiaoming Wong
Educational Background:	Ph.D. in Linguistics, *Shanghai International Studies University*, 2016–2020
	M.A. in Linguistics, *Shanghai International Studies University*, GPA: 3.7/4, 2013–2016
	B.A. in English, *Shanghai International Studies University*, GPA: 3.9/4, 2009–2013
Work Experience:	2018-present, Translator, *Shanghai International Business Group*
	2015.06–2015.09, TOEFL teacher, *New Oriental*
Academic Award:	University Translation Competition, 2nd place
Expected Salary:	￥15,000–￥20000

Appendix B. Questions for retrospective interviews

Warm-up questions:

1. What did you think of the task?
2. What do you think of your performance?

Questions for Pre-task Planning:
(for Group 2 & 4)

1. How did you spend your time preparing the task?
2. What was your focus when you were preparing the task? (for Group 1 & 2)
3. How do you feel about the time limit for completing the task?
4. Do you think you would perform differently if there had been more time allowed?

Questions for On-line Planning:
(for Group 3 &4)

1. What were you thinking while you were discussing (the options) with your partner?
2. Did you spend time analyzing and revising your expressions while discussing (ideas) with your partner? (for Group 1 & 3)

© 2021. John Benjamins Publishing Company
All rights reserved

3. How do you feel about the limited time for preparing the task?
4. Do you think you would perform differently if there had been more time allowed?

Address for correspondence

Qianqian Gu
KU Leuven
01.42, Blijde-Inkomststraat 21
3000, Leuven
Belgium
qianqian.gu@kuleuven.be

Biographical notes

Gu Qianqian (Shane) is currently a Ph.D. student in the research group Language and Education, KU Leuven. Her doctoral research project is related to the effects of on-line planning on second language performance and development, funded by the Chinese Scholarship Council. She received her first M.A. in Applied Linguistics from the Chinese University of Hong Kong and her second M.A. in Advanced Studies in Linguistics from KU Leuven. Her research interests include task-based language teaching, task planning behavior, and second language acquisition.

Publication history

Date received: 25 September 2020
Date accepted: 19 August 2021

© 2021. John Benjamins Publishing Company
All rights reserved