Examining the Effect of Manipulating Task Complexity Along the [±Planning Time] Dimension and the [±Here and Now] Dimension on Learner Perceptions of Task Difficulty

Tomohito Ishikawa and Maiko Katherine Nakano

Second language (L2) tasks have recently attracted considerable attention in L2 research and pedagogy (e.g., Cadierno & Robinson, 2009; Ellis, 2008; Ishikawa, 2005, 2006, 2007, 2008a, 2008b, to appear; Gilabert, 2007a, 2007b; Robinson, 2001a, 2001b, 2005, 2007a, 2007b, 2009; Robinson & Gilabert, 2007; Skehan, 1998, 2003). Historically, the advent of the use of L2 tasks as a unit of planning for language classroom activities was primed by several ideas influenced by Communicative Language Teaching (Bygate, Skehan, & Swain, 2001). Tasks have been recognized as a solution for different levels of problems that various task-users encounter, such as language teachers, learners, and testers (Bygate et al., 2001).

The 1980s fostered three major approaches to task-based syllabus design (Long & Crookes, 1992). These task-based syllabi include the procedural syllabus, the process syllabus, and task-based language teaching (TBLT) (see e.g., Long, 1985; Long & Crookes, 1992, 1993; Long & Robinson, 1998). In the 1990s, Long and Crookes (1992) emphasized several strengths of TBLT
over the other approaches with regard to their compatibility with SLA and classroom research findings, and notably their proposed methodological principle of “focus on form” (e.g., Long & Robinson, 1998). After advocating the use of tasks as a potentially effective means for language pedagogy, Long and Crookes (1992) voiced the concern of TBLT with grading and sequencing criteria for pedagogic tasks. They state that it “remains one of the oldest unsolved problems in language teaching of all kinds” (Long & Crookes, 1992, p. 46).

Closely related to Long and Crookes’ (1992) concern is the question: What effects does differential task complexity have on the L2 learner’s production; namely, linguistic complexity, fluency, and accuracy? Asking this question and the pursuit of its answers are crucial because pedagogic tasks first need to be categorized (Robinson, 2007a) in the light of empirical findings so that L2 development may be facilitated (Robinson, 2005). In order to examine task factors empirically, a TBLT framework is provided by Robinson.

Robinson (2007a, 2009), largely in agreement with Long and Crookes’ approach to TBLT, points out the necessity of categorizing and sequencing pedagogic tasks from an information processing perspective in a way that is congruent with behavioral descriptions of the target tasks identified by needs analysis (e.g., one-way vs. two way tasks). He further emphasizes the importance of respecting individual differences and considering the role they would play as they interact with task demands and conversational conditions, producing differential effects on L2 learning, retention, and performance (Robinson, 2007a). In order to accommodate these broad issues, Robinson (2007a) proposes his framework for TBLT, where three major task components and their respective roles are identified.
Robinson's triadic framework for TBLT

The present study adopted Robinson's triadic framework (Table 1) because it is one of the most comprehensive and systematic TBLT frameworks (also see Skehan, 1998, for another framework). The triadic componential framework specifies three broad categories of task characteristics (task complexity, condition, and difficulty), which can be used in designing pedagogic tasks that increasingly approximate the demands of target tasks identified via a needs analysis.

First, in Robinson's framework, task complexity refers to the information processing demands of the pedagogic tasks on memory, attention, and reasoning. This category of task dimensions is characterized as cognitive in nature, and each proposed dimension represents inherent and relatively fixed task demands of pedagogic tasks (e.g., ±here-and-now], ±intentional reasoning], see below for further description of variables of task complexity). These proposed dimensions are relatively stable and manipulable by teachers and syllabus designers prior to task performance. Therefore, they can contribute to prospective decisions in sequencing pedagogic tasks in the classroom. What is of crucial importance is that, according to Robinson, task complexity is to be the sole criterion for prospective task sequencing and it constitutes the heart of task-based syllabus design. No other task categories play a crucial role in considering prospective task sequencing decisions.

The second category of task condition refers to the interactive factors of pedagogic tasks, which consists of several descriptive, behavioral properties. These interactional, descriptive task characteristics include participation variables and participant variables. Some of the participation variables concern the distribution as well as the flow of information in performing pedagogic tasks (one-way vs. two-way) and the relative degree of freedom in the number of solutions (open vs. closed). Other participation variables concern the
number of participants, the degree of need for meaning negotiation and contributions. Participant variables on the other hand concern the participants’ attributes relevant to task performance. Included are attributes of participants such as proficiency level, gender. Robinson (2007a) states that the task condition is largely constrained by the target task that is identified by needs analysis and is to be approximated by a series of pedagogic tasks. Whereas task condition is a crucial factor if the series of pedagogic tasks are to bear a close similarity to the real-world target task, it does not play a central role when it comes to sequencing decisions.

Finally, according to Robinson (2007a), the task category of task difficulty refers to learner factors which can influence the perceptions of the difficulty of tasks. Task difficulty arises as a result of an interaction between strengths and weaknesses in such factors and the intrinsic cognitive complexity of tasks. Some variables within this category are subject to temporal change (e.g., motivation) in contrast to relatively stable ability variables (e.g., working memory capacity). Concerning prospective task sequencing decisions, the learner’s perceptions are impossible to assess prior to task implementation, and it is for these reasons that prospective sequencing decisions that are based on task difficulty are impossible to make. However, they are very important indicators for online decisions in the language classroom.

Moreover, interactions between the three categories may also take place. For instance, differential task complexity demands may induce differing levels of perceptions in the task difficulty on the part of the learner. Mediating the relationship between task complexity and task difficulty perceptions are the learner’s affective and ability variables. Affective variables of task difficulty such as higher motivation or anxiety may temporarily expand or reduce the learner’s available ability resources, making them perceive tasks to be easier or more difficult. Likewise, participation roles and participant factors
Table 1  Pedagogic L2 task classification- Categories, Criteria, Analytic procedures, and Characteristics (based on Robinson, 2007a)

<table>
<thead>
<tr>
<th>Task complexity</th>
<th>Task Condition</th>
<th>Task difficulty</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Cognitive factors)</td>
<td>(Interactive factors)</td>
<td>(Learner factors)</td>
</tr>
<tr>
<td>(Classification criteria: cognitive demands)</td>
<td>(Classification criteria: interactional demands)</td>
<td>(Classification criteria: ability requirements)</td>
</tr>
<tr>
<td>(Classification procedure: information-theoretic analyses)</td>
<td>(Classification procedure: behavior descriptive analyses)</td>
<td>(Classification procedure: ability assessment analyses)</td>
</tr>
</tbody>
</table>

Sub categories:

a) resource-directing variables
making cognitive/conceptual demands
+/- here and now
+/- few elements
+/- spatial reasoning
+/- causal reasoning
+/- intentional reasoning
+/- perspective-taking

b) resource-dispersing variables
making performative/procedural demands
+/- planning time
+/- prior knowledge
+/- single task
+/- task structure
+/- few steps
+/- independency of steps

Sub categories:

a) participation variables
making interactional demands
+/- open solution
+/- one way flow
+/- convergent solution
+/- few participants
+/- few contributions needed
+/- negotiation not needed

b) participant variables
making interactant demands
+/- same proficiency
+/- same gender
+/- familiar
+/- shared content knowledge
+/- equal status and role
+/- shared cultural knowledge

Sub categories:

a) ability variables
and task relevant resource differentials
h/l working memory
h/l reasoning
h/l task-switching
h/l aptitude
h/l field independence
h/l mind-reading

b) affective variables
and task relevant state-trait differentials
h/l openness
h/l control of emotion
h/l task motivation
l/h processing anxiety
h/l willingness to communicate
h/l self-efficacy
that belong to the category of task condition also affect task difficulty perceptions.

Returning to the category of task complexity, the degree of task complexity is, in Robinson's framework, defined according to what he calls "resource-directing" and "resource-dispersing" dimensions (see Table 1). The resource-directing dimensions of task complexity have the essential characteristic of directing the learner's cognitive resources to aspects of the target language code in response to task demands (Robinson, 2001a, 2001b, 2003, 2005a, 2007a). For example, when the task goal is to identify certain objects among a set of similar distracters (i.e., [+few elements]) (e.g., Brown & Yule, 1983), the learner's attention may be directed toward a set of linguistic options that can serve to distinguish the intended objects from the competing ones (e.g., attributive adjectives, relative clauses or locative expressions, depending on the history of discourse and the nature of the situational context). Alternatively, when the immediate speech situation is not shared by the participants (i.e., [- here and now], see Robinson, 1995; Robinson & Cadierno, & Shirai, 2009), spatio-temporally displaced referents (e.g., agents, events, and objects) need to be established, which requires the use of more complex linguistic coding systems such as the morphological markers than those required in the here-and-now speech situation. Further, tasks that impose IR demands on the speaker require the learner to verbalize the mental states of others in order to provide psychological accounts of their behavior. Consequently, greater production of the complex language structures (e.g., causative expressions such as "because" or mental verbs such as "think"), which may well entail syntactic subordination and complementation, is predicted to take place.

The resource-dispersing dimensions, on the other hand, have their characteristic of affecting the amount of cognitive resources available to the learner
EXAMINING THE EFFECT OF MANIPULATING TASK COMPLEXITY

during L2 task performance without necessarily affecting the attention allocation policy. For instance, giving planning time ([+ planning time]) or performing a single task ([− dual task]) makes cognitive resources available to the learner. However, this does not necessarily direct those cognitive resources to specific aspects of the target language code. Thus, increasing task complexity along the resource-dispersing dimensions places heavy demands on mental access to cognitive resources.

Research questions and hypotheses

The purpose of the present study is to investigate the effect of increasing task complexity along the [±planning time] and [±here and now] dimensions on the learner's perceptions of task difficulty. The research question of the present study is What effect do increasing task complexity along the [±planning time] and [±here and now] dimensions have on the learner perceptions of task difficulty?

The Cognition Hypothesis claims that more cognitively complex tasks will be perceived to be more difficult than less complex counterparts by all learners (e.g., Robinson, 2007b). For instance, it is expected that the task with the task design features of [−planning time] would be perceived more difficult by the learner than the pedagogic tasks with the task design feature of [+planning time]. This is also applicable to the cases of [±here and now] conditions. Thus, the task with the task design features of [−here and now] would be perceived more difficult by the learner than the pedagogic tasks with the task design feature of [+here and now].

Following Robinson (2001b), it is hypothesized that the task conditions associated with the task design feature of [−here and now] would be perceived for instance more difficult than those with [+here and now]. Similarly, the task conditions associated with the task design feature of [−planning time]
would be perceived for instance more difficult than those with [+planning time]. Similarly to Robinson (2001b) we hypothesized that these relations would be applicable to Difficulty and Frustration; the reverse relation would be applicable to the task difficulty item of Confidence (e.g., the learner would be least confident on the most complex task). We did not pose any directional hypotheses regarding Interest and Motivation. Consequently, the following 10 hypotheses were formulated:

H 1: The learner’s perceived task difficulty of Difficulty will be higher on the tasks associated with the task design feature of [(here and now] than those with the task design feature of [+here and now].

H 2: The learner’s perceived task difficulty of Frustration will be higher on the tasks associated with the task design feature of [—here and now] than those with the task design feature of [+here and now].

H 3: The learner’s perceived task difficulty of Confidence will be lower on the tasks associated with the task design feature of [—here and now] than those with the task design feature of [+here and now].

H 4: The learner’s perceived task difficulty of Interest will be as high on the tasks associated with the task design feature of [—here and now] as those with the task design feature of [+here and now].

H 5: The learner’s perceived task difficulty of Motivation will be as high on the tasks associated with the task design feature of [—here and now] as those with the task design feature of [+here and now].

H 6: The learner’s perceived task difficulty of Difficulty will be higher on the tasks associated with the task design feature of [—planning time] than those with the task design feature of [+planning time].

H 7: The learner’s perceived task difficulty of Frustration will be higher on the tasks associated with the task design feature of [—planning time]
than those with the task design feature of [+planning time].

H 8: The learner's perceived task difficulty of Confidence will be lower on the tasks associated with the task design feature of [-planning time] than those with the task design feature of [+planning time].

H 9: The learner's perceived task difficulty of Interest will be as high on the tasks associated with the task design feature of [-planning time] as those with the task design feature of [+planning time].

H 10: The learner's perceived task difficulty of Motivation will be as high on the tasks associated with the task design feature of [-planning time] as those with the task design feature of [+planning time].

Method

Participants
The participants for the present study were third-year high school students learning English as an L2, whose L1 was Japanese (N = 72; female = 41, male = 31). The participants' proficiency levels ranged between low to high intermediate and their ages ranged from 17 to 18 years old, with an average age of 17.6 years old. They participated in the experiment voluntarily and were paid 500 yen (circa 5 dollars).

Materials
Two picture cartoons were used in the present study. Picture A is in Appendix A and Picture B is in Appendix B.

Experimental design
Table 2 shows the experimental design of the present study. The factor of [±planning time] was a between-subject variable whereas the factor of [±here
Table 2  *Experimental design of the present study and the number of participants in each experimental condition*

<table>
<thead>
<tr>
<th>Planning condition</th>
<th>Picture order</th>
<th>[-planning time]</th>
<th>[+planning time]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task sequencing</td>
<td>A-B</td>
<td>B-A</td>
<td>A-B</td>
</tr>
<tr>
<td>[+here and now]-[-here and now]</td>
<td>9 9</td>
<td>9 9</td>
<td></td>
</tr>
<tr>
<td>[-here and now]-[+here and now]</td>
<td>9 9</td>
<td>9 9</td>
<td></td>
</tr>
</tbody>
</table>

and now] was a within-subject variable. In the case of the within-subject variable of [+here and now], the effect of picture sequencing was counterbalanced. As Table 2 shows, there are eight cells and each cell had 9 learners, amounting to 72 learners.

**Procedures**

In conducting the current study, the participants were informed that the speaking task would not be considered as a part of their school grades. They were also informed that the speech samples would be used for research purposes. The research was conducted on individual basis.

Before the experiment began, the researcher told the participant to perform the two tasks (i.e., cartoons A & B). In the [+here and now] condition, the learner was allowed to view a strip cartoon when they spoke and required to use the present tense during speaking whereas in the [(here and now] condition, the picture cartoon was not available to the learner when they were speaking, when the learner was required to use the past tense. In the [-planning time] condition, the learner was required to begin speaking one minute after the picture cartoon was passed on to them whereas in the [+planning time], the learner had 10 minutes of planning time. During the planning ses-
sion, the learner was allowed to take notes (words/phrases) but the notes were removed before speaking (see also Crookes, 1989). In all conditions, the learner was not allowed to name the characters in the strip cartoon (see S. Ishikawa (1995) for a similar treatment to elicit articles).

Task difficulty questionnaire

The task-difficulty questionnaire was adopted from Robinson (2001b), who included items assessing overall difficulty, frustration, confidence, interest, and motivation. The questionnaire was used to check whether the task complexity conceptualized by the task designer is matched by perceptions of difficulty on the part of the L2 learner.

i) Difficulty (I thought this task was easy/ thought this task was hard)
ii) Frustration (I felt relaxed doing this task/I felt frustrated doing this task)
iii) Confidence (I did not do well on this task well—I did well on this task)
iv) Interest (This task was not interesting/This task was interesting)
v) Motivation (I don’t want to do more tasks like this/I want to do more tasks like this)

The learner was asked to respond to each item along a seven-point Likert scale. The questionnaire items were written in the learners’ L1. In coding the data, a score of 0 was assigned, for example, to the semantically most negative learner response to each questionnaire item (easy, no confidence, no motivation etc.) and a score of 8 to the semantically most positive learner response (most difficult, most frustrated, most interesting etc.) to each questionnaire item.
Results

It was found that the distributions of the task difficulty scores were not distributed normally. Therefore, non-parametric statistical analyses were employed to test the hypotheses.

Table 3  Rank-order means of the [-here and now] task conditions and the [+here and now] task conditions

<table>
<thead>
<tr>
<th>Task difficulty measure</th>
<th>Task complexity</th>
<th>Mean rank</th>
<th>Mean rank</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[+here and now]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difficulty</td>
<td></td>
<td>1.43</td>
<td>1.57</td>
</tr>
<tr>
<td>Frustration</td>
<td></td>
<td>1.56</td>
<td>1.44</td>
</tr>
<tr>
<td>Confidence</td>
<td></td>
<td>1.53</td>
<td>1.47</td>
</tr>
<tr>
<td>Interest</td>
<td></td>
<td>1.47</td>
<td>1.53</td>
</tr>
<tr>
<td>Motivation</td>
<td></td>
<td>1.57</td>
<td>1.43</td>
</tr>
</tbody>
</table>

Figure 1. Rank-Order Means of Difficulty, Frustration, Confidence, Interest, and Motivation as a Function of the Task Design Feature of [+here and now]
**Descriptive statistics**

Tables 3 and 4 show the mean ranks of the [-here and now] task conditions and the [+here and now] task conditions and the mean ranks of the [-planning time] task conditions and the [+planning time] task conditions respectively. Figures 1 and 2 represent these results visually.

**Table 4**  *Rank-order means of the [-planning time] task conditions and the [+planning time] task conditions*

<table>
<thead>
<tr>
<th>Task difficulty measure</th>
<th>Task complexity</th>
<th>Mean rank</th>
<th>Mean rank</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[+here and now]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difficulty</td>
<td>73.86</td>
<td>71.14</td>
<td></td>
</tr>
<tr>
<td>Frustration</td>
<td>76.02</td>
<td>68.98</td>
<td></td>
</tr>
<tr>
<td>Confidence</td>
<td>70.40</td>
<td>74.60</td>
<td></td>
</tr>
<tr>
<td>Interest</td>
<td>68.68</td>
<td>76.32</td>
<td></td>
</tr>
<tr>
<td>Motivation</td>
<td>64.27</td>
<td>80.73</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[-here and now]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 2.** Rank-Order Means of Difficulty, Frustration, Confidence, Interest, and Motivation as a Function of the Task Design Feature of [+planning time]
Hypotheses testing

Friedman’s analyses of variance (ANOVAs) were used in order to test Hypotheses 1-5. To test Hypotheses 6-10, the Kruskal-Wallis tests were applied to the dataset. The alpha was set at the probability level of $p = .005$ for Hypotheses 4, 5, 9, and 10 as a result of Bonferroni correction. As for Hypotheses 1, 2, 3, 6, 7, and 8, which were directional or one tailed, the alpha was set at the probability level of $p = .01$.

The results of Friedman’s ANOVAs are summarized in Table 4. As Table 4 shows, none of Friedman’s ANOVA results reached the statistical significance level. This means that Hypotheses 1-3 were disconfirmed whereas Hypotheses 4 and 5 were confirmed.

Similar patterns of results were obtained in the case of the results of the Kruskal-Wallis test results. The results are summarized in Table 5. As Table 5 shows, none of the results were significant although the result of Motivation approximated the significance level of $p = .005$. Thus, Hypotheses 6-8 were disconfirmed whereas Hypotheses 9 and 10 were confirmed.

Discussion

The present study investigated the effect of increasing task complexity along the [±planning time] and [±here and now] dimensions on the learner’s

<table>
<thead>
<tr>
<th>Measure</th>
<th>df</th>
<th>$^2$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Difficulty</td>
<td>1</td>
<td>2.17</td>
<td>.14</td>
</tr>
<tr>
<td>Frustration</td>
<td>1</td>
<td>1.60</td>
<td>.21</td>
</tr>
<tr>
<td>Confidence</td>
<td>1</td>
<td>0.64</td>
<td>.42</td>
</tr>
<tr>
<td>Interest</td>
<td>1</td>
<td>0.71</td>
<td>.40</td>
</tr>
<tr>
<td>Motivation</td>
<td>1</td>
<td>2.78</td>
<td>.10</td>
</tr>
</tbody>
</table>

Table 5  Summary Table of Friedman’s ANOVA Results
perceptions of task difficulty. In order to answer the research question of What effect do increasing task complexity along the [±planning time] and [±here and now] dimensions have on learner perceptions of task difficulty?, 10 hypotheses were formulated and tested empirically. The results indicated that Hypotheses 1, 2, 3, 6, 7, and 8 were disconfirmed whereas Hypotheses 4, 5, 9, and 10 were confirmed.

The reason Hypotheses 1, 2, 3, 6, 7, and 8 were disconfirmed seem to be related to one of the interaction characteristics of the task; namely monologic aspects of the pedagogical tasks. Garrod and Pickering (2004) pointed out that monologic tasks impose greater processing demands than interactive tasks in certain respects. Most people have much less experience (i.e., practice opportunities) in monologue than dialogue (e.g., Garrod & Pickering, 2004, 2008; Pickering & Garrod, 2004).

Additionally, in performing monologic tasks, the learner has to plan and generate a series of utterances by themselves, while in performing interactive tasks, higher-order message planning as well as lower-order utterance planning is distributed among the communicators because dialogue is a joint activity (e.g., Clark, 1996). More specifically, in dialogue, interactive alignment at various linguistic levels (e.g., situation model, semantic, and syntactic levels) can be expected, where layers of linguistic representations become and remain active when the speech comprehension system processes the lin-

<table>
<thead>
<tr>
<th>Measure</th>
<th>df</th>
<th>2</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Difficulty</td>
<td>1</td>
<td>0.16</td>
<td>.69</td>
</tr>
<tr>
<td>Frustration</td>
<td>1</td>
<td>1.05</td>
<td>.31</td>
</tr>
<tr>
<td>Confidence</td>
<td>1</td>
<td>0.41</td>
<td>.53</td>
</tr>
<tr>
<td>Interest</td>
<td>1</td>
<td>1.26</td>
<td>.26</td>
</tr>
<tr>
<td>Motivation</td>
<td>1</td>
<td>5.76</td>
<td>.02</td>
</tr>
</tbody>
</table>
guistic input supplied by the interlocutor (e.g., Kormos, 2006).

Furthermore, in monologue, the learner has to produce a series of long utterances (e.g., Tavakoli & Foster, 2008), which frequently entails speech planning while speaking or pausing whereas in dialogue, elliptical and fragmentary utterances are preferred and efficient means of communication due to the common ground are established between the communicators (e.g., Garrod & Pickering, 2004).

In terms of communication strategies, monologue also has to be carried out without strategic assistance from the interlocutor (e.g., Dörnyei & Kormos, 1998). These factors all imply that generally speaking, language processing on monologic tasks tends to be difficult and frustrating and the learner confidence may be lowered, especially when the learner lacks L2 knowledge, and the speech production system is not fully automatized (Poulisse, 1997). These conversational task demands might have overridden the predicted effects of increasing task complexity along the two dimensions especially in the case of Japanese learners, who are often claimed to lack L2 communicative opportunities.

On the other hand, Hypotheses 4, 5, 9, and 10 were confirmed. The results were compatible with Robinson (2001b) and Gilabert (2005). These results were rather welcomed. It would pose more difficult problems if the learner loses interest in, and motivation for, task performance when task complexity is increased.

References


Cadierno, T., & Robinson, P. (2009). Language typology, task complexity and the


ishikawa, T. (2007). The effect of increasing task complexity along the [+Here-and-


Appendix A. Picture cartoon A used in the present study
Appendix B. Picture cartoon A used in the present study