
**PERSONAL INVESTMENT IN L2 TASK DESIGN AND LEARNING:
A CASE STUDY OF TWO JAPANESE LEARNERS OF ENGLISH**

*Craig Lambert and Danny Minn
The University of Kitakyushu, Fukuoka, Japan*

The present study explores ways in which personal investment might be operationalized as a condition in L2 task design, how this condition might affect L2 learning, and how the effects of the condition might be effectively measured in future empirical research. Two intermediate-level Japanese females completed three task types (instructional, narrative, opinion) in two conditions (personal investment & teacher investment). The personal investment (PI) condition required them to supply the content and resource materials on which the pedagogic tasks operated, whereas in the teacher investment (TI) condition these materials were supplied as part of the task as set. Three types of data were then triangulated in order to explore ways in which this condition may affect L2 performance and metacognition on tasks: oral performances, post-performance protocols, and questionnaires. Results indicate that performances in the PI condition contained more elaboration and more associations with existing knowledge. The condition may thus provide one means of controlling the ‘need’ component of Laufer and Hulstijn’s (2001) Involvement Load Hypothesis. The study explores the potential of tasks in L2 acquisition, providing both a heuristic and initial hypotheses for future study of relationships between this task condition, L2 performance and metacognition.

Key words: Task Design; Motivation; Vocabulary; Involvement Load Hypothesis

Este estudio explora los modos en los que la “inversión personal” podría funcionar como una condición en el diseño de tareas en una L2, cómo afectaría esta opción al aprendizaje de la L2, y cómo los efectos de la

opción podrían medirse de forma efectiva en investigaciones empíricas futuras. Dos informantes femeninas japonesas de nivel intermedio realizaron tres tipos de tareas (instrucción, narración y opinión) en dos opciones (“inversión personal” e “inversión del profesor”). La opción “inversión personal” (IP) requería que ellas proporcionasen el contenido y los materiales didácticos sobre las que se realizaban las tareas pedagógicas, mientras que en la opción “inversión del profesor” (IP) estos materiales eran proporcionados como parte de la tarea. Se triangularon tres tipos de datos obtenidos con el fin de explorar los modos en los que estas opciones podrían afectar el rendimiento en la L2 y la metacognición en las tareas: representaciones orales, protocolos posteriores a las representaciones y cuestionarios. Los resultados indican que el rendimiento en la opción PI era más elaborada y había más asociaciones con el conocimiento previo. La opción puede así proporcionar un modo de controlar el “componente necesidad” de la “Involvement Load Hypothesis” de Laufer & Hulstijn (2001). El estudio explora el potencial de las tareas en la adquisición de una L2, proporcionando tanto una hipótesis heurística como inicial para futuros estudios de las relaciones entre estas opciones de tareas, el rendimiento en la L2 y la metacognición.

Palabras clave: Diseño de tareas, motivación, vocabulario, Involvement Load Hypothesis

1. Introduction

In most task-based approaches to L2 learning, it is generally argued that the communicative demands of the tasks themselves are capable of pushing L2 development. In other words, as the complexity of tasks increases, the developmental complexity of the language they require will be forced to increase as well if the tasks are to be successfully completed. While there will not, of course, be a one-to-one relationship between task complexity and L2 complexity, there will inevitably be considerable overlap, particularly at the beginning and lower-intermediate levels. Due to this overlap, L2 learning will be able to proceed effectively as a product of both the actual demands of the tasks that learners need to complete and of their own internal syllabuses,

personalities and preferences. Instruction can thus be integrated with both practice and acquisition rather than serving as preparatory step as it has in traditional approaches.

The tasks used in such an approach generally operate on a planned diversion in the information held by learners and usually approximate to some degree a real-world task that learners have to complete. The need to share information requires learners to communicate functionally in the L2, and the real-world connection allows them to acquire task-specific language and skills. Yule (1997) proposes a practical typology of communication tasks that consists of four general stages which he argues place progressively higher discourse demands on the speaker: (1) descriptive tasks, (2) instructional tasks, (3) narrative tasks and (4) opinion tasks. Although new approaches are emerging (e.g., Robinson, 2007; Skehan, 1998), Yule's typology provides a principled basis for sequencing tasks developmentally.

In terms of the measurement of learning, Skehan (1996, 1998) posits fluency, accuracy and complexity as goals for task-based L2 instruction. As attention is essential for learning, and the capacity of attention is limited, L2 learners will not be able to communicate in real time and acquire new L2 forms simultaneously. Skehan hypothesizes that a trade-off effect will occur between the complexity and fluency of speech in performance, and that this trade-off effect might be task-induced. In other words, task design factors that can be shown to induce more fluent production will do so at the expense of complexity, whereas those that induce more complex production will consequently be less fluent. A balanced approach to task-based L2 instruction will consist of sequences of tasks designed to provide constant cycles that alternatively emphasize complexity (i.e., destabilization and restructuring of the developing language system) and fluency in using new language once acquired. One purpose of task-based learning research will be to determine how task design might support such *dual-mode* processing.

In spite of the opportunities that well-designed and well-sequenced tasks provide for balanced L2 acquisition, experienced teachers will understand the pedagogic importance of the difference between the performances of learners who are only willing to do what they have to do to

complete communication tasks and those who are willing to do everything they can to complete them. Laufer and Hulstijn (2001) and Hulstijn and Laufer (2001) review the importance of learner involvement in research on learning, arguing that increased involvement will improve L2 learners' retention of the language that they process on tasks through elaboration and associations with existing knowledge. They propose a construct for task-induced involvement that consists of three components: need, search and evaluation. The combined effect of these three factors will constitute learners' level of involvement in a task and consequent retention of the language used. This theory is termed the *Involvement Load Hypothesis*. Need is the motivational component of the involvement construct and is most relevant to the current study. It is hypothesized to have two levels: *moderate* (imposed by an external agent) and *strong* (self-imposed by the learners). Tasks that induce a *strong* level of need are argued to increase learners' involvement and thus improve retention of the language used on tasks.

The personal investment (PI) condition investigated in the present study may represent one way of designing tasks to induce a *strong* as opposed to a *moderate* level of need. The PI condition required learners to supply the content and resource materials on which the pedagogic tasks operated, whereas in the teacher investment (TI) condition these materials were supplied as part of the task as set. The complexity, interactive demands, and discourse demands of the tasks in the respective conditions were ostensibly the same.

2. Methodology

This case study triangulates three sources of data on participants' task performances in order to explore ways in which the PI and TI conditions may affect L2 performance and metacognition on tasks: (1) participants' on-task production, (2) post-performance verbal protocols on what participants were thinking about while completing the tasks, and (3) questionnaires on

participants' motivation, anxiety and personal responses to the task conditions.

Participants

Both participants in the case study were Japanese females. They were both majoring in English and in their second year of the program when the study was conducted. One was 20 years old and had begun the program directly out of high-school; the other was in her mid-thirties and had worked for several years as an English teacher and tour guide before returning to university to pursue a degree in English. They were both communicatively competent in English and able to participate relatively fluently in everyday social interactions. They both had completed homestays in English-speaking countries. They could also understand and express their ideas during transactional tasks on less familiar topics without too much struggling for vocabulary and forms of expression. They were both high-achievers in their courses in the program, but neither had a clear idea about why she was studying English or what she would do after she graduated from the program. They were thus relatively typical of the high-proficiency, high-motivation sub-population of English majors at universities in Japan.

The study was conducted at a public university in Southern Japan. Approximately 6000 learners were enrolled in the university. Approximately 500 of these were majoring in English. The university had a high reputation in the area for its English program, attracting students from as far north as Hiroshima and as far south as Okinawa. In selecting participants for the study, announcements were made to all second-year English majors that volunteers were being sought for a research study on teaching oral English. Participants would be paid 1000 Japanese yen per hour (about seven Euros) for the time they contributed. They would be required to attend four sessions of approximately two hours each. During each session, they would perform four speaking tasks similar to the ones that they were used to performing in class, and these performances would be recorded on both audiotape and videotape. Following each session, they would complete questionnaires, watch a video of their own performances, and discuss it with one of the researchers. It was stressed that participants would need to be confident in

their ability to communicate in English as all of the tasks, materials, and discussions would be in English. Two learners volunteered to participate in the study.

Procedures

Each of the four data collection sessions followed the same basic procedures. They were conducted in the university's recording studio. Microphones and a video camera were set up before each session. Seats were arranged so that participants faced each other and could speak in a normal voice but were too far apart to share printed material. The camera was set at a slightly elevated angle so that participants would be able to recall what was on the desk in front of them as they watched the video. In each of the four recording sessions, the participants performed four versions of one task type (see Materials below). There were two versions of each task in the TI condition and two versions in the PI condition. In the TI condition this allowed each participant to perform both the primary speaker (information holder) role and secondary speaker (information receiver) role. In the PI condition this allowed each participant to supply the content and resource materials on which one task operated (and thus to be the primary speaker or information holder) and to discuss the content and resource materials that had been supplied by her partner on another task (and thus to be the secondary speaker or information receiver). The order in which the tasks were completed was counter-balanced to control for performance effects on condition and role.

After each recording session, both participants completed three questionnaires and the post-performance protocols. The protocol sessions were all recorded on audiotape. The protocols sought to determine what learners had been thinking about as they completed each task. Participants were asked to watch the videotape of their performances immediately after each recording session. They were given the remote control and asked to pause the videotape whenever they could recall something that they were thinking about while completing the task and explain it. In cases where a significant pause took place in the performance and the participant did not pause the video to comment, the researcher elicited a comment by signaling

the participant to pause the tape and asking her to explain what she was thinking about at that point. While one participant completed the protocols, the other completed the questionnaires. The order was alternated with the ordering of the tasks for each of the sessions. Each participant completed two motivation questionnaires (one for the TI versions of the task and the other for the PI versions) and two anxiety questionnaires (likewise, one for the TI versions of the tasks and the other for the PI versions). They also completed an open-response questionnaire eliciting their feedback on the tasks in each condition (see Materials below).

During the first protocol session, it was determined that the participants had not understood what was expected of them while performing the tasks in the first recording session. They had seemed afraid to speak and only spoke in simple, carefully constructed sentences. During the protocols, they explained that, in spite of directions to the contrary, they had wanted to speak “naturally” but were unsure if it was permissible during the recording. From the second recording session onward, they understood the instructions and spoke freely during each task.

Materials

The study was originally intended to consist of four task types based on Yule’s (1997) typology: descriptive, instructional, narrative and opinion. Four versions of each were created: two TI versions and two PI, requiring participants to perform both the information-giver and the information-receiver roles in each condition. However, the four descriptive tasks used in the first recording session were eliminated from further analysis when it was determined that participants had not understood the instructions. The remaining twelve performances of three task types (instructional, narrative and opinion) thus form the database for the present study (see Lambert, 2004, for additional examples of tasks in both the PI and TI conditions).

Instructional Task: The TI versions of this task required each participant to explain how to solve a problem based on an eight-frame picture sequence. The first problem was explaining how a farmer could move his chicken, wheat and fox across a river in a boat which could hold

only himself and one other thing at a time without leaving either the chicken and the wheat together alone (as the chicken would eat the wheat) or the fox and the chicken together alone (as the fox would eat the chicken). The eight pictures illustrated each step. The second problem involved three rings (small, medium and large) and three poles (left, middle and right). The rings were stacked largest to smallest on the left pole. The task was to explain how to move the rings from the left pole to the right pole without moving more than one ring at a time or without having a larger ring on top of a smaller ring. Again, the eight pictures illustrated each step.

The PI versions of the instructional task required each participant to explain how to do or make something that was interesting to her personally and which she would like to introduce to her partner. Eight blank picture boxes were provided and participants were given ten minutes to silently plan the content of their task and develop the resource materials by drawing pictures to illustrate each step in the blank boxes. They were not allowed to write down any language that they would use to complete the task. The researcher confirmed that these directions had been followed before the tasks were performed. One of the participants chose to explain how to make a plum tea drink that she had recently learned and believed to be healthy, and the other chose to explain how to make a cocktail that she had recently been trained to make at work.

Narrative Task: The TI versions of this task required each participant to tell a story based on a four-frame picture sequence. The stories were taken from practice versions of an English proficiency test developed in Japan (EIKEN). Japanese picture strips were chosen as they were more intuitive for the participants and controlled for culturally unfamiliar meanings and situations. The first story involved a child who was sitting next to her mother on a train and became preoccupied with something outside. While she was looking out the window, her mother offered her seat to an elderly woman. In the end, the child began to speak to the elderly woman, mistakenly assuming that it is her mother and became embarrassed when she realized that it was not. The second story involved a couple on a group gondola ride with a professional singer. One of the passengers suddenly stood up and began to

sing himself to the dismay and embarrassment of the others. In the end, he was knocked off the boat as it went under a bridge.

The PI versions of the narrative task required each participant to share a past experience that she felt would be interesting to her partner. They were given ten minutes to silently plan the content of their task and develop the resource materials by drawing pictures to illustrate the main events of the story into four blank picture boxes. They were not allowed to write down any language that they would use to complete the task. The researcher confirmed that these directions had been followed before the tasks were performed. One of the participants chose to tell a story about losing her handbag during her arrival for a homestay in Australia and the implications that it had for her and her new host family. The other chose to tell the story of how she had misplaced her bicycle at the station that afternoon and her embarrassment after mistakenly reporting it stolen.

Opinion Task: The TI versions of this task required each participant to explain an unfortunate scenario involving five characters. Each was given ten minutes to read her scenario and write key word notes that she would use to explain the scenario to her partner. They then had to discuss their opinions and rank the five characters from the most to the least guilty. The first scenario was adopted from Duff (1986) and involved a man who had had his leg amputated as a result of a complex series of events in which all of the characters were implicated to some degree. The second scenario was created to parallel the first in structure and complexity. It involved a death in a traffic accident.

The PI versions of the opinion task required each participant to create an unfortunate scenario that she thought would be interesting to discuss with her partner. Participants were given ten minutes to silently plan the content of their tasks by writing scenarios of their own and the resource materials of the task by writing key word notes that they would use to explain the scenarios to each other. The researcher confirmed that these directions had been followed before the tasks were performed. One participant's scenario involved a mugging that took place after a family quarrel connected with the father's loss of employment. The other

participant's scenario involved an act of vengeance resulting from a complicated set of romantic relationships between two couples.

Questionnaires: Finally, three questionnaires were used in the study. These questionnaires were adopted from Jacob (1996). The first two measured motivation and anxiety, respectively, during participants' performance of the tasks in each of the conditions (see Appendix 1). The third questionnaire elicited participants' feedback and opinions. It consisted of five open-ended questions asking them to compare their personal responses to the tasks in each condition. The questions asked which versions of the task they enjoyed (PI or TI), which versions they felt were more helpful for improving their English (PI or TI), which versions were more difficult to complete (PI or TI), the extent to which they felt they could confidently carry out the versions of the tasks in the PI condition in the future, and finally the extent to which they felt they could confidently carry out the versions of the task in the TI condition in the future. The conditions were not labeled as such on the questionnaires, but simply referred to as 'the tasks on Form 1' and 'the tasks on Form 2.'

Analysis

The twelve task performances (four versions of each of the three task types) were transcribed by one of the researchers and verified by the other. On reviewing the transcriptions, it was found that participants had not followed the directions regarding the use of key word notes during the performances of the opinion tasks. This resulted in performances that consisted of two very different tasks. In the first, participants read the scenarios verbatim. In the second, they discussed their opinions and ranked the characters. As the data for the first task consisted of a verbatim reading of a written text and was readily separable from the discussion task, this part of the data was eliminated from the subsequent analysis. Each participant turn was then divided into AS-Units (analysis of speech units, Foster, et al., 2000) against which complexity, fluency, accuracy and voluntary production were calculated.

AS-Units: The guidelines of Foster et al. (2000) for a ‘level two’ analysis were followed closely in dividing the discourse into AS-units. Independent clauses and subordinate clauses, as well as independent sub-clausal units that could be elaborated into full clauses, were counted as AS-units. Minor utterances that did not add referential meaning to the discourse, such as echo responses or non-elicited feedback to show comprehension or agreement (e.g., uh-huh, yeah, right, okay), were not coded. In more difficult cases, pauses and intonation contours were taken into consideration, the governing principle being whether a piece of discourse constituted a single chunk of micro-planning.

Complexity: This was the ratio of clauses (s-nodes) to AS-units. The guidelines provided by Foster et al. (2000) were followed closely in coding clauses. Clauses were tensed or untensed verbs that were not functioning as noun phrases. Thus, non-finite verbs were only coded when accompanied by a subject, object, complement or adverbial to verify their clausal status. The number of clauses was divided by the number of AS-units produced by each participant on each task (see Table 1).

Fluency: This was the ratio of dysfluency markers (pauses and repetitions) to AS-units. Dysfluent pauses were those in which participants seemed to be searching for language or which otherwise seemed due to a deficiency in English ability. Pauses that set off planned chunks or which were used for rhetorical effect were not coded as dysfluencies. AS-unit breaks were thus never double coded as dysfluent pauses. Dysfluent repetitions were cases in which participants repeated a word or phrase until they could put together the necessary language to continue. Thus, a series of uncompleted false starts followed in the end by an accurate AS-unit was coded as a number of dysfluent repetitions within one error-free AS-unit. Repetitions used for functional purposes within the discourse context (such as when a speaker wanted to stress something or make sure that the listener understood) were not coded as dysfluencies. The number of dysfluencies was divided by the number of AS-units produced by each participant on each task (see Table 1).

Accuracy: This was the percentage of error-free AS-units. The number of error-free AS-units was divided by the total number of AS-units produced by each participant on each task (see Table 1).

Voluntary Production: This was the percentage of non-obligatory AS-units. AS-units that did not function to complete an essential step of the task as set, but rather functioned to elaborate the discourse based on personal interest or associations with existing knowledge, were coded as voluntary (see Appendix 2 and Lambert, 2001, for examples). The number of voluntary AS-units was divided by the total number of AS-units produced by each participant on each task (see Table 1).

Both researchers independently coded all the task performances for AS-units, complexity, fluency, accuracy, and voluntary production. The coding of each variable were then compared and all differences were resolved resulting in 100% inter-rater agreement on the five variables.

Lexical Types: This was the number of distinct vocabulary items. All of the performances were analyzed using AntConc 3.2.0w concordance software (Anthony, 2006). The software automatically removed all punctuation and capitalization. Each of the twelve transcripts was sorted for speaker so that the total size of the lexicon employed by each participant during each performance could be determined (see Table 1). After running the concordance software, proper names and inflectional variants of the same word were excluded to arrive at the total number of word types used by each participant on each task.

Post-Performance Protocols: After participants' comments about what they were thinking while completing the tasks were transcribed, they were grouped into five categories created to account for the data: (1) thoughts connected with participants' real-world experiences (or associations with existing knowledge); (2) thoughts connected with the demands of the tasks and how to complete them; (3) thoughts connected with strategies for communicating more effectively or efficiently; (4) thoughts connected with the meaning, selection or use of vocabulary items; and (5) thoughts connected with the grammatical forms required to express ideas. The number

of comments in each category was then tallied for condition and task type (see Table 3).

Questionnaires: In the case of the motivation and anxiety questionnaires (see Appendix 1), participants' responses to each item were converted to a five-point scale in which "strongly agree" was allotted five points and "strongly disagree" was allotted one point. Their answers to the initial questions about the general constructs on a ten-point scale were halved in each case to match the aforementioned five-point scales. All of the questions on the motivation questionnaire were positive, whereas all of the items on the anxiety questionnaire, with the exception of Item 5, were negative. Responses to Item 5 were inverted as higher confidence indicates lower anxiety. Participants' responses on all items for motivation and all items for anxiety were then averaged to arrive at a general measure of each participant's level of motivation and anxiety on each task type and in each condition (see Table 4). The third questionnaire elicited participants' feedback on the task conditions. In the case of the questions on the enjoyment, usefulness and difficulty of the tasks, responses in which participants indicated a definite opinion about the tasks in one condition over the other, as well as responses in which no definite opinion was expressed either way, were identified and tallied (see Table 5). In the case of the questions about their confidence to complete each task, positive and negative responses were identified and tallied (see Table 5).

3. Results

On-Task Production: The observed differences in the language that participants produced while completing the tasks in each of the conditions (PI and TI) are summarized in Table 1. Consistent differences were observed both in the amount of voluntary production and the number of lexical types that the participants employed. These two participants produced a considerably greater proportion of voluntary AS-units on all three of the task types in the PI condition than they did on those in the TI condition. They

also employed a considerably larger vocabulary on the tasks in the PI condition than in the TI condition.

| Task | Condition | Lexical Types | Voluntary | Accuracy | Dysfluency | Complexity |
|---------------|-----------|---------------|------------|------------|-------------|-------------|
| Instructional | PI | 191 | 34% | 68% | 1.28 | 1.15 |
| | TI | 135 | 04% | 61% | 1.28 | 1.04 |
| Narrative | PI | 166 | 54% | 81% | 1.30 | 1.44 |
| | TI | 158 | 34% | 64% | 1.30 | 1.32 |
| Opinion | PI | 246 | 13% | 76% | 1.24 | 1.42 |
| | TI | 173 | 05% | 82% | 1.86 | 1.55 |
| Overall | PI | 603 | 34% | 75% | 1.27 | 1.34 |
| | TI | 466 | 14% | 69% | 1.48 | 1.30 |

There were no clear differences observed in the accuracy, fluency or complexity of the participants' speech between conditions. Although participants' production was somewhat more accurate and complex on the instructional and narrative tasks in the PI condition than in the TI condition, the opposite was observed for their production on the opinion tasks. Their fluency only differed between conditions on the opinion tasks. It was identical in both conditions on the instructional and the narrative tasks.

An interesting secondary observation was a trade-off between fluency and complexity-accuracy across tasks (instructional, narrative, and opinion). Table 2 summarizes the observed differences in participants' speech with conditions combined.

| Task | Lexical Types | Voluntary | Accuracy | Dysfluency | Complexity |
|---------------|---------------|-----------|----------|------------|------------|
| Instructional | 326 | 19% | 65% | 1.28 | 1.10 |
| Narrative | 324 | 44% | 73% | 1.30 | 1.38 |
| Opinion | 419 | 09% | 79% | 1.55 | 1.49 |

As the discourse demands of the tasks increased from instructional to narrative to opinion (Yule, 1997), participants' production progressively increased in complexity and accuracy while decreasing in fluency. No systematic differences in voluntary production or size of lexicon were observed across tasks.

Post-Performance Protocols: The distributions of learners' thoughts across tasks and conditions as revealed by their comments during the post-performance protocols are summarized in Table 3.

Table 3: Differences Observed in Metacognition between Tasks and Conditions

| Comment Types | Real-World Connections | | Task Demands | | Comm. Strategies | | Vocabulary Selection | | Language Form | |
|---------------|------------------------|----|--------------|----|------------------|----|----------------------|----|---------------|----|
| | PI | TI | PI | TI | PI | TI | PI | TI | PI | TI |
| Instruction | 7 | 0 | 0 | 10 | 3 | 8 | 3 | 6 | 0 | 1 |
| Narrative | 10 | 7 | 3 | 6 | 1 | 2 | 1 | 4 | 0 | 1 |
| Opinion | 14 | 11 | 27 | 7 | 4 | 5 | 1 | 3 | 1 | 1 |
| Overall | 31 | 18 | 30 | 23 | 8 | 15 | 5 | 13 | 1 | 3 |

Participants' comments reflected considerably more associations with existing knowledge on all three tasks in PI condition than they did in the TI condition. By contrast, they consistently reported more thoughts connected with communication strategies, vocabulary usage and language form in the TI condition than in the PI condition. Comments concerning thoughts about the demands of the tasks, however, seemed to be more connected with the nature of the tasks themselves than with the condition. Participants reported thinking more about task demands on the TI versions of both the instructional and narrative tasks, but during the opinion task they reported thinking more about task demands on the PI versions.

Questionnaires: The results of the motivation and anxiety questionnaires are summarized in Table 4. The results for motivation indicate that both participants generally perceived the versions of tasks in the PI condition to be considerably more motivating for them than the versions in the TI condition. The exception was the instructional tasks which they perceived as equally motivating in both conditions. By contrast, both participants felt that the versions of the tasks in the TI condition made them considerably more anxious than the versions in the PI condition. The exception was Participant 1 who indicated that all versions of the instructional task made her equally anxious.

| Instructional Tasks | Motivation | | Anxiety | |
|-------------------------|-------------|-------------|-------------|-------------|
| | PI | TI | PI | TI |
| Participant 1 | 3.79 | 3.79 | 1.92 | 1.92 |
| Participant 2 | 3.79 | 3.79 | 2.00 | 2.42 |
| Totals | 3.79 | 3.79 | 1.96 | 2.17 |
| Narrative Tasks | PI | TI | PI | TI |
| Participant 1 | 4.79 | 4.57 | 0.92 | 1.75 |
| Participant 2 | 3.93 | 3.50 | 1.92 | 2.33 |
| Totals | 4.36 | 4.04 | 1.42 | 2.04 |
| Opinion Tasks | PI | TI | PI | TI |
| Participant 1 | 4.07 | 3.21 | 2.00 | 3.92 |
| Participant 2 | 3.86 | 3.64 | 1.92 | 2.92 |
| Totals | 3.97 | 3.43 | 1.96 | 3.42 |
| Overall | PI | TI | PI | TI |
| Participant 1 | 4.22 | 3.86 | 1.61 | 2.53 |
| Participant 2 | 3.86 | 3.64 | 1.95 | 2.56 |
| Condition Totals | 4.04 | 3.75 | 1.78 | 2.54 |

Finally, Table 5 summarizes the results of the participant feedback questionnaire. The overall results indicate that they slightly favored the tasks in the PI condition as being more enjoyable, more useful for improving English, less difficult, and generating slightly more confidence in their ability to perform them in the future than those in the TI condition. However, these differences were nominal, and they were inconsistent across task types. Furthermore, there were a large proportion of instances in which participants' responses did not express any perceived difference (ND) between the tasks in either condition.

| Question Types | Enjoyment | | | Usefulness | | | Difficulty | | | Confidence PI | | Confidence TI | |
|----------------|-----------|----------|----------|------------|----------|----------|------------|----------|----------|---------------|----------|---------------|----------|
| | PI | TI | ND | PI | TI | ND | PI | TI | ND | Yes | No | Yes | No |
| Instruction | 1 | 1 | 0 | 0 | 0 | 2 | 0 | 1 | 1 | 2 | 0 | 1 | 1 |
| Narrative | 2 | 0 | 0 | 2 | 0 | 0 | 0 | 1 | 1 | 2 | 0 | 2 | 0 |
| Opinion | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 |
| Overall | 3 | 2 | 1 | 2 | 1 | 3 | 1 | 2 | 3 | 5 | 1 | 4 | 2 |

4. Discussion

Overall, the results of the study indicate that these two participants were more involved in the tasks in the PI condition than they were in those in the TI condition. On the PI versions of the tasks, they went beyond the demands of the task as set, producing more voluntary AS-units and elaborating on the content discussed (see the voluntary exchanges in Appendix 2, for example). The post-performance protocols also revealed that participants made more associations with existing knowledge on the PI versions of the tasks than on the TI versions. In addition, they employed a larger range of vocabulary during their performance of the tasks in the PI condition than during those in the TI condition. Allowing learners to supply the content and resource materials on which tasks operate, tailoring them to their needs and interests, may thus constitute one form of task-induced involvement, improving learners' retention of the language that they process on tasks through increased elaboration and associations with existing knowledge. In other words, the PI condition may provide one means of controlling the 'need' component of learner-involvement in the *Involvement Load Hypothesis* (Laufer & Hulstijn, 2001; Hulstijn & Laufer, 2001).

A secondary observation was the systematic differences in fluency, complexity and accuracy that were observed across task types (see Table 2). As tasks increased in discourse complexity from instructional to narrative to opinion (Yule, 1997), participants' production progressively increased in complexity and accuracy while decreasing in fluency. This observed trade-off effect is consistent with Skehan's (1996, 1998) model of task-induced *dual-mode* processing. In the present study, the difference observed between the PI and TI conditions seem as if they could have been independent of these variables, having potentially complementary effects on L2 acquisition.

These initial hypotheses are still data-based speculations, however. Before any generalizations can be made regarding a relationship between the PI condition and learner involvement, relationships between variables, or differences between research and classroom contexts, inferential empirical work with larger numbers of learners will be required. The present case

study nevertheless seems to justify such efforts and provides a heuristic for both operationalizing the PI condition in L2 task design and for triangulating data to verify its affect on performance and metacognition.

References

- Anthony, L. (2006). *AntConc 3.2.0w*. Free concordance software [Online] Available: http://www.antlab.sci.waseda.ac.jp/antconc_index.html.
- Duff, P. (1986). Another look at interlanguage talk: taking task to task. In R. Day (Ed.). *Talking to Learn* (pp. 147-181). Rowley: Newbury House.
- Foster, P., Tonkyn, A., Wigglesworth, G. (2000). Measuring the spoken language: a unit for all reasons. *Applied Linguistics* 21(3), 354-375.
- Hulstijn, J. & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning* 51(3): 559-558.
- Jacob, A. (1996). *Anxiety and motivation during second language task performance in Singaporean schools*. Unpublished M.A. Dissertation, National University of Singapore/RELC, Singapore.
- Lambert, C. (2001). Task sequencing and affective performance variables. *Kitakyushu University Faculty of Foreign Studies Bulletin* 103: 97-175.
- Lambert, C. (2004). Reverse-engineering communication tasks. *ELT Journal* 58(1): 18-27
- Laufer, B. & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: the construct of task-induced involvement. *Applied Linguistics*, 22(1), 1-26.
- Robinson, P. (2007). Criteria for classifying and sequencing pedagogic tasks. In Maria del Pilar Garcia Mayo (Ed.). *Investigating Tasks in Formal Language Learning* (pp. 7-26). Clevedon: Multilingual Matters.
- Skehan, P. (1996). A framework for the implementation of task-based learning. *Applied Linguistics* 17(1), 38-62.

Skehan, P. (1998). *A Cognitive Approach to Language Learning*. Oxford: Oxford University Press.

Yule, G. (1997). *Referential Communication Tasks*. Mahwah, NJ: Lawrence Erlbaum.

Appendix 1: Questionnaires (adopted from Jacob, 1996)

MOTIVATION: When you are motivated, you want to do something very much. Please indicate your level of motivation while completing this task. '1' should be used to indicate the lowest level of motivation, and '10' should be used to indicate the highest level.

1 2 3 4 5 6 7 8 9 10

Six statements on motivation during this task are provided below. Indicate the extent to which you agree with each statement by circling the appropriate response.

1. I like the way that this task is designed.

Strongly Agree Agree Neutral Disagree Strongly Disagree

2. Doing this task can help me improve my English.

Strongly Agree Agree Neutral Disagree Strongly Disagree

3. I was very involved in arriving at the answer for this task.

Strongly Agree Agree Neutral Disagree Strongly Disagree

4. My partner was very involved in arriving at the answer for this task.

Strongly Agree Agree Neutral Disagree Strongly Disagree

5. Sharing information with my partners was useful for completing this task.

Strongly Agree Agree Neutral Disagree Strongly Disagree

6. I think doing this task can help me outside the classroom as well.

Strongly Agree Agree Neutral Disagree Strongly Disagree

ANXIETY: When you are anxious, you feel worried or nervous. Please indicate your level of anxiety while completing this task. '1' should be used to indicate the lowest level of anxiety, and '10' should be used to indicate the highest level.

1 2 3 4 5 6 7 8 9 10

Five statements on anxiety during this task are provided below. Indicate the extent to which you agree with each statement by circling the appropriate response.

1. I felt nervous when I was doing the task.

Strongly Agree Agree Neutral Disagree Strongly Disagree

2. I was worried that I would not be able to express my ideas clearly.

Strongly Agree Agree Neutral Disagree Strongly Disagree

3. I was worried that my partner was not giving me enough information.

Strongly Agree Agree Neutral Disagree Strongly Disagree

4. I was anxious that it was too difficult to arrive at the answer.

Strongly Agree Agree Neutral Disagree Strongly Disagree

5. I am confident that I can do another task like this well.

Strongly Agree Agree Neutral Disagree Strongly Disagree

Appendix 2: Voluntary on-task production

A transcript of a narrative task performance in the PI condition is provided below. Voluntary AS-units are *italicized* and boxed.

A: My story is about / my first home stay in Australia.

B: *You went to Australia?*

A: *Yeah.*

B: *When when did you go?*

A: *Yeah, the last when I was / Sydney, I went to Sydney / when I was 18 // Yeahm any years ago.*

// Anyway / I arrived at the Sydney airport. // Then first, I had to go to / the / English school.

B: Yeah.

A: So I took a taxi. Then I arrived at / the / English school. // Then I went into / the room. // Then I realized / that I lost my handbag.

B: Yeah?

A: It's / I'm not sure / that / I left at the airport or / I left it in / in the taxi or / maybe I was / I had stolen / my bag was stolen. // I'm not sure.

B: *Yeah, but did you have money?*

A: *Yeah.*

// I had the passport / with me, but in my handbag / there was a traveler's checks/ and some money/ and the key / of the suitcase.

B: Oh!

A: Then / I went to / the / the house of the home stay. // Then I / told everything / to my host / mother and father. // So I asked / my host father / to / break / the suitcase / because I have / I had to open.

B: Yeah.

A: And the next day, I had to go to the Tokyo Bank / to / to fill in the paper / then / I need / some money / the money back, the traveler's checks

B: Yeah.

A: *So it was a miserable / a miserable / beginning / of my home stay.*

B: *So you didn't find your handbag?*

A: *Un -un, no.*

B: *You... oh!*

A: *Somewhere / in Sydney.*

B: *Still?*

A: *Yeah.// Maybe.*

B: *That's funny.*

Received 15 June 2007: Revised version received 1 October 2007