THE JOURNAL OF ASIA TEFL Vol. 5, No. 3, pp. 199-223, Autumn 2008

Task-based Strategy Assessment: The Effect of Task Difficulty on Listening Strategy Use of Advanced EFL Learners

Mahmood Reza Atai
Tarbiat Moallem University, Iran

Seyyed Mohammad Reza Hashemi
Tarbiat Moallem University, Iran

Recently there has been heated debate on the nature of strategy assessment in task-absent situations where it might lead to learners' over- or underreporting of the strategies they actually use. More specifically, it is argued that that learners need to be involved in the real process of completing a language task just before they are asked to report the strategies they use. The present study, thus, was aimed at investigating the effect of presence of tasks and task difficulty on advanced students' reports of their listening strategy use. Results from the analysis of the data measured by the Listening Strategy Questionnaire and collected from 62 Advanced students, in three task conditions, indicated that (1) task condition (No Task, Easy Task, Difficult Task) has an effect on advanced EFL learners' listening strategy use, and (2) considering each strategy subcategory under study, the subjects reported different frequencies of strategy use across the above task conditions. This was especially true for cognitive and memory strategies, whereas for compensation strategies the difference in frequency of strategy use across the three conditions was not statistically significant. The findings promise some implications for developing context-specific task design frameworks, and task-based strategy instruction and evaluation.

Key words: task-based, strategy assessment, memory strategies, cognitive strategies, compensation strategies

INTRODUCTION

Among a good number of determining factors related to the process of language learning, learning strategies seem to have gained a significant amount of importance in the past two decades (Oxford, 1990; Cohen, 1998; Oxford, 2001). Consequently, strategy instruction and assessment have gained attention in the corresponding avenue of literature (Oxford, 1990; Chamot, Barnhardt, El-Diary, & Robins, 1996; Cohen, 1998, 2003, to cite a few). Strategy assessment, in particular, has been considered as an area in need of further research (see Oxford, Cho, Leung, & Kim, 2004).

Traditionally, researchers assessed language learners' strategy use through questionnaires (e.g. Oxford's SILL, 1990), requiring learners to provide hypothetical answers to items. However, this way of eliciting learners' strategy use is criticized for its inconsistency since, as Cohen (2003) argues, it may cause learners' under- or over-reporting a set of strategies, which are not actually used when the learners are involved in completing tasks. One approach toward eliciting a more valid account of strategy use, then, is the so-called task-based strategy assessment (Oxford et al., 2004).

From among the variety of interrelated factors affecting task-based strategy use and assessment in second language acquisition, task difficulty seems to be of considerable significance (Robinson, 2005; Oxford et al., 2004). However, even though robust literature is available on areas such as task-based teaching (see Bygate, Skehan, & Swain, 2001), and employing strategies to accomplish tasks (O'Malley & Chamot, 1990), few studies have specifically addressed "task-based strategy assessment" (Ikeda & Takeuchi, 2000; Oxford et al., 2004), which is the focus of the present study.

This study may be considered an extension of the above-mentioned line of research. However, the focus of the present study is on listening comprehension strategy use of advanced EFL learners, considering the subcategories of

direct strategies including memory, compensation and cognitive strategies (Oxford, 1990). The purpose was to find out whether the learners use different task-specific strategies in different situations and whether frequency of strategy use by learners is significantly different across different task conditions. The significance of studies such as the present one lies in the claim that they would pave the way, in theory and practice, for linking research-based assessment methods to task design frameworks.

TASK-BASED STRATEGY ASSESSMENT: AN OVERVIEW OF THEORY AND RESEARCH

Assessment is an inseparable part of any strategy-based language program (Cohen, 1998; Oxford, 2001). However, it seems that it is hardly ever possible to have an exact and perfectly thorough assessment of the strategies a learner uses in a particular situation: "Since language learning strategies are generally internal and mentalistic processes, certain research may fail to reveal adequately which strategies learners apply" (Cohen & Scott, 1996, p. 90).

Among a good number of techniques for eliciting learner response for investigating strategy use, questionnaires seem to be most popular. What is questionable, though, is whether using questionnaires per se help teachers come to reliable judgments about how frequently learners use language learning strategies. In other words, it is worth investigating whether learners' reports of their strategy use are affected by the demand of the context or the type of task they are completing.

Cohen (2003) and Oxford et al. (2004) argue that learners' reports of their strategy use are, to a great extent, affected by the type of task they are doing. These scholars, hence, have attempted to shed light on a new avenue of research that is built on the premise that learners' strategy use should be assessed through more reliable procedures. One such current line of practice is linking strategies and tasks together to run research into task-based strategy assessment.

Task Difficulty

The issue of task difficulty, as Nunan asserts, is "of central importance to researchers, curriculum developers, syllabus designers, materials writers, and classroom teachers, and it is not surprising that it has been the subject of considerable research" (2004, p. 72). The reason why degree of difficulty of language tasks is critical in L2 task design is "if tasks are too easy they will present no challenge, and are not likely to extend any other goals of restructuring, accuracy, or fluency in any effective way" (Skehan, 1996, p. 53). On the other hand, tasks which are too difficult "are likely to overemphasize fluency" (Skehan, ibid).

However, as Nunan (2004, p. 72) states, despite much research in the area, "researchers have, in fact, only begun to scratch the surface, and there is, as yet, no objective method for determining task complexity/difficulty." To Skehan (1996), task difficulty is especially important because it presents a challenge for course designers.

Task difficulty is actually a determining factor that is geared, closely, to engineering design factors in the course of developing a task as a reliable consciousness-raising instrument for assessing learner strategies. Task difficulty, thus, has become the focus of a number of studies so that the nature of task complexity with regard to pedagogical and psychological factors involved would be subject to more rigorous analysis. Ellis (2003), for instance, offers a list of criteria related to task complexity. Ellis's (2003) list of criteria accounting for task complexity include input, conditions, processes, and outcomes. Along similar lines, Robinson (2005) offers a triadic conceptual framework for task complexity. Robinson (ibid) claims that using this framework enables us to analyze complex classroom learning and testing in a manageable way. His framework allows interaction among elements of complexity, difficulty, and condition.

As for analyzing difficulty of skills-oriented tasks, particularly listening task difficulty, which is the focal point in the present study, research has introduced factors and variables affecting task design. Slatyer, Brindley, &

Wiggleworth (2000), for example, examined the role of some key variables in listening task design. They found that there exists a complex interaction between the text and other task components. The difficulty factors in Slatyer et al. (2000) were not classified a priori; rather, these factors were identified through post hoc analysis.

In their article on task difficulty in ESL listening assessment, Brindley and Slatyer (2002) consider some factors underlying task difficulty. These factors, having been identified through previous research (see Buck, 2001), include the nature of input (i.e. speech rate, length of passage, syntactic complexity, vocabulary, discourse structure, noise level, accent, register, propositional density, and amount of redundancy), the nature of assessment task, the listener factors, etc. (Brindley & Slatyer, 2002). The notion of task difficulty, however, would prove to be of more complicated nature when it is linked to strategies.

Linking Tasks and Strategies

As a good number of scholars (Skehan, 1998; Cohen, 2003; Oxford et al., 2004) argue, linking strategies to tasks can lead to more accurate assessment of strategies of language learning and language use. Oxford et al. (2004) assert that difficulty of task is among the most significant factors closely related to L2 performance. However, as stated earlier, a few studies investigated task difficulty and its possible effect on learners' strategy use (Cohen, 2003). Cho et al. (2001, cited in Hsiao & Oxford, 2002) examined the effect of presence or absence of an actual task on students' self-reports of their strategy use through a reading strategy questionnaire. The results indicated that in the task-present condition some strategies were reported with different frequencies compared with an earlier administration of the same questionnaire in the No Task condition. However, the frequencies reported for most of the strategies did not change across task-present or task-absent conditions. Similarly, Ikeda and Takeuchi (2000) investigated the effect of presence or absence of tasks on the learners' reports of their reading strategy use. They

also examined how task difficulty caused variations in types and frequencies of strategies used by the learners. Their study showed that high proficiency and low proficiency learners over-reported their use of strategies in the No Task condition compared to task-present conditions and that task difficulty had a significant effect on the types and frequencies of reported strategy use.

Vandergrift (2002) investigated the development of metacognition in L2 listening comprehension by asking a group of beginning-level French students to complete some listening comprehension tasks and reflective exercises, using the instruments that engaged the students in the prediction and evaluation of the process. The results showed that the subjects' reflection on the process of listening raised their awareness of the process of listening and helped them understand strategies involved in successful completion of the L2 listening tasks. Vandergrift (2002) further maintains that "reflection on the successful completion of a listening task through consciousness-raising and subsequent deployment of appropriate planning, monitoring and evaluation strategies can build student motivation for L2 listening and learning, and, in addition, develop constructive attributional beliefs about the efficacy of these strategies" (Vandergrift, 2002, p. 115).

Oxford et al. (2004) worked on task-specific strategy use of the learners. The researchers used Flesch Readability measure in order to estimate the degree of difficulty of their tasks. Since their study was exploratory in nature, they chose 0.10 as the level of significance. Generally speaking, the findings showed that high proficiency learners reported more frequent use of metacognitive (top-down) strategies than low proficiency ones in the No Task and Easy Task conditions; however, in the difficult task condition, low proficiency learners had more difficulty with the difficult reading passage which, in turn, led them to use various types of strategies and with greater frequency than high proficiency learners did (Oxford et al., 2004).

As the above review of the relevant empirical studies may indicate, investigating and assessing learners' strategy use may prove more valid and reliable if we consider as many underlying factors as possible.

Research Questions

In the present study, care was taken to consider the relationship between tasks and strategies in the course of assessing learners' listening comprehension strategies. To this end, the following research questions were formulated:

(1) Does task condition (No Task, Easy Task, Difficult Task) have any significant effect on high proficiency language learners' listening strategy use? (2) Considering each strategy subcategory under study (i.e. memory, compensation, cognitive), do high proficiency learners report different frequencies of use across each of the above task conditions?

METHOD

Participants

Out of a total of 86 students who were initially planned to take part in the study, 62 could make it all the way through the project. The subjects were chosen from a pool of available TOEFL classes for high proficiency candidates. After the proficiency test was administered, 17 scores turned out to be ±1 standard deviation below and above the mean (mean = 505.91, SD = 27.5 3), so these students were not included in the study. Also, seven other students were not considered as participants because of being absent from class during the three-session phase of data collection. Consequently, 62 students (30 female, 32 male) participated in the study. The subjects had an age range of 16 to 30 and were advanced language learners studying English in "Listening for TOEFL" classes at the Iran Language Institute (ILI), Tehran, Iran.

Instrumentation and Materials

The following instruments were used in this study.

The Proficiency Test

First, a retired version of the TOEFL (adopted from Philips, 2001), containing 50 listening, 40 structure and written expressions, and 50 reading items, was used as a measure to ascertain the homogeneity of the subjects' proficiency level prior to the study. Also, it provided a criterion against which the validity indices for the two listening comprehension tasks were estimated. The reliability of the TOEFL test, estimated using KR-21 formula, was .84, which proved to be satisfactory.

The Listening Tasks

Two listening comprehension tasks, task A (the easy task) and task B (the difficult task) were used in the task-present conditions. The easy listening text was an old story taken from a listening series called "Active Listening: Building Skills for Understanding," by Helgesen and Brown (1994). The difficult listening text was adopted from a taped book summary: "Arc of Ambition" by Champy and Nohria (2000). Also, 10 True/False items (see Appendix A) were developed for each task.

The validity indices for the easy and difficult tasks were .71 and .77 respectively. The indices were estimated through correlating the task scores with the TOEFL test scores. Also, the reliability indices, estimated against K-R 21 formula, were .61 for the easy task and .64 for the difficult task.

To ensure that the listening tasks were statistically different in terms of their difficulty level, the researchers administered the tasks to a 16-student class prior to the target group administration. Then the scores were analyzed using a paired t-test. According to the results of the t-test ($T_{observed}$ = 17.816 > $T_{critical}$ = 2.602, with df =15 at .01 significance level) the difference was statistically meaningful.

The difficulty level of each task, following Brindly and Slatyer (2002), was also scrutinized in terms of input characteristics such as rate of speech, lexical complexity and text difficulty. The above input characteristics were

analyzed using the following measures:

(1) Rate of speech (wpm): When choosing the task, care was taken to pick out tasks with moderately different rates of speech. However, although the difficult listening task was faster than the easy one, the difference was not in line with the reported rates for easy and difficult tasks. So the speed of the difficult task was increased using sound editing software. The rate for the easy task was 123 wpm (slow), and the initial rate of speech for the difficult task was 129 wpm, which was increased to 166 wpm, which is considered to be normal for native speakers (Griffiths, 1990).

(2) Lexical complexity: The Cobuild Dictionary for Advanced Learners was used to distinguish between high frequency (easy) and low frequency (difficult) words. Information on the frequency of words in Cobuild is given using a five-level frequency scale, ranging from the most frequent (5 diamonds) to the least (1 diamond). According to Cobuild, dictionary bands 1 and 2 include more advanced vocabulary whereas bands 4 and 5 are very common words that account for about 75% of all English usage. Table 1 shows the results of the comparison of the easy and difficult tasks according to their lexical difficulty.

(3) Text difficulty (Flesch's formula): Difficulty level of the text (input) was estimated in terms of readability using Flesch's readability formula. The indices were estimated using Microsoft Word. For the easy task Flesch Reading Ease was 83.1, and for the difficult task it was 44.7.

Another measure for judging the difficulty level of the two listening tasks was expert judgment. So five experts (PhD holders in TEFL) were asked to evaluate the tasks and rate them easy or difficult in comparisons with each other. Among the five of them, four rated task A easy and task B difficult. One of the experts, however, rated both tasks easy.

TABLE 1
Summary of Difficulty Estimates for the Listening tasks

Input Characteristics	Task A (Easy)	Task B (Difficult)
Lexical Difficulty	5: 86 %	5: 58.86 %
(frequency percentage:	4: 5.8 %	4: 11.70 %
5=The most frequent,	3: 5.8 %	3: 12.41 %

1=the least frequent)	2: 1.07 % 1: 1.33 %	2: 6.4 % 1: 10.63 %
Flesch Readability	83.1	44.7
Rate of Speech	123	166

The Listening Strategy Questionnaire

The last data collection instrument was a listening strategy use inventory which was developed and validated by Afsarnia (2000). Afsarnia developed his questionnaire based on previously standardized models (Oxford's SILL, 1990; Bacon, 1992; O'Malley & Chamot, 1995). Afsarnia (2000) used 1389 subjects to develop the questionnaire. First 624 ILI intermediate students in Shiraz self-reported their listening comprehension strategies in two consecutive sessions. Then the Farsi version of the Listening Comprehension Strategy Inventory was administered to 676 ILI students for validation purposes. The estimated reliability indices of the Farsi questionnaire were reported as follows: test-retest alpha= .88 (n=96), test-retest correlation coefficient= .78 (n=96), standard item alpha= .88 (n=96), and K-R 21= .86 (n=676). Also, the questionnaire items revealed a 70 % match with Oxford' (1990) SILL, a 75% match with Bacon's (1992) inventory and an 80% match with O'Malley and Chamot's (1995) model. A principal varimax rotation factor analysis was also conducted to ensure the construct validity of the Farsi questionnaire.

The questionnaire was used to collect data about students' use of direct strategies. Therefore, only 32 (out of 67) questions were taken from the above-mentioned questionnaire. To ensure the internal consistency of the questionnaire, the coefficient α was calculated for this study (α=.68).

The questionnaire consisted of questions on direct listening strategies: memory strategies (1-10), compensation strategies (11-16), and cognitive strategies (17-32). The answers for frequency of strategies were scaled from 1 to 5 with 1 indicating "never or almost never true of me," and 5 "always or almost always true of me." Because of the unequal number of questions in the original questionnaire, the modified questionnaire also contained unequal number of questions for different categories.

Pilot Study

A pilot study was designed and conducted in the summer and fall of 2005 in order to examine the appropriateness and difficulty level of the tasks. Different criteria were used in doing so. The first criterion for judging the appropriateness and difficulty level of the listening tasks was expert judgment. To this end, five experts (PhD holders) were asked to comment on task type and appropriateness of the items, considering their difficulty level, and to rate the tasks easy or difficult in comparison with each other. Four of the experts rated task A easy and task B difficult. However, one of them rated both tasks easy. Their feedback indicated that task A could be called easy in comparison to task B. Based on the experts' further comments, some of the items were also modified (for task A, items 1, 6, and 7 were made easier; and for task B, items 3, 4, 8 and 10 were made more difficult). Later, a group of 20 ILI advanced students were given the tasks to complete. They were also asked to rate each task easy, average or difficult in comparison with the other. As a result, 16 students rated task A easy, and task B difficult. Three of the students believed that the tasks were of average difficulty level and one of them rated both tasks easy.

The implication of the pilot phase for the study was that researchers modified the tasks according to the feedback they received. It also helped the researchers develop a better understanding of the process of conducting the study and take into account the probable problems, such as students being absent, reports being made too quickly, lack of time, teachers' reactions, etc.

Validation Procedure

To probe the psychometric values for the listening comprehension tasks, the researchres chose a group of 16 ILI advanced students, highly similar to the target group, to take the listening tasks along with the TOEFL test. The scores of the listening tasks, then, were correlated with those of the proficiency test. The correlation indices obtained for tasks A and B were .71

and .77 respectively. The indices imply that the tasks benefited from satisfactory criterion validity. Reliability estimates for tasks A and B were checked against Kuder-Richardson formula 21 (K-R 21). The indices, .61 for task A and .64 for task B, seemed acceptable for this study. Also, to ensure the internal consistency of the listening comprehension questionnaire, Cronbach's alpha coefficient analysis was conducted on the results of the Listening Strategy Questionnaire which was given to 62 students. According to the results, the reliability coefficient α was .62 which proved quite acceptable for this study.

Procedures

For the purpose of investigating the effect of presence and difficulty of tasks on advanced students' listening strategy use, the researchers collected the data during three sessions, with a one-week interval between each two. In the first session, the students were asked to fill out the Listening Strategy Questionnaire without doing any listening task. This condition was called "No Task Condition" (NTC). In the second session, after one week, the students were given a listening task, task A (the easy task), followed by the questionnaire. First, they had to listen to the tape and do task A. Then, immediately after doing the task, they were asked to fill out the questionnaire with respect to the strategies they used during the task. This condition was called "Easy Task Condition" (ETC). Finally, in the third session, again after one week, the students were assigned to do the difficult task (task B), and were asked to fill out the questionnaire right after they were through with the task. This condition was called "Difficult Task Condition" (DTC). In the ETC and DTC, just after completing the task, the students were instructed to answer the questions according to the strategies they had just used during each listening task.

As one of the reviewers has rightly mentioned, there could be some data contamination as far as the familiarity effect would concern, thus affecting the students' reports of their strategy use. However, as the students were exposed to one task condition on each occasion, familiarity effect would

hardly influence their reports of strategy use.

RESULTS

Data analysis included repeated measures analysis of variance (ANOVA) followed by post hoc Scheffé test. The first null hypothesis of this study stating that task condition—NTC, ETC, DTC—has no effect on high proficiency language learners' listening comprehension strategy use, was probed through a repeated measures ANOVA and the means of total questionnaire rates were compared across the three task conditions. Table 2 presents descriptive data of the rates across the three task conditions. What might be interesting at first glance is that the reported frequency of strategy use in the No Task condition (mean=103.81, SD=11.24) was higher than that in the Easy Task condition (mean= 96.4, SD= 12.66) and the Difficult Task condition (mean= 88.94, SD= 15.34). As it is clear, the reported frequency in DTC was the lowest.

TABLE 2
Means and Standard Deviations of Total Frequencies Across the Task Conditions

	No Task	Easy Task	Difficult Task
N	62	62	62
Mean	103.81	96.40	88.94
Std. Deviation	11.24	12.66	15.35

Next, a repeated measures ANOVA was used to test whether these differences were statistically significant. The results of the repeated measures ANOVA on the overall reported mean frequency of strategy use (Table 3) indicated significant differences across the task conditions. According to Table 3, the difference between the mean rates was statistically significant, $F(2,122)= 24.78$, $p< .000$ or $F_{observed} = 24.78 > F_{critical} = 4.98$ with the significance level of .01. Therefore, the first null hypothesis is rejected, indicating that task difficulty has an effect on advanced students' reports of their listening strategy use.

TABLE 3
Results of ANOVA on Overall Reported Frequency

Source	Sum of Squares	df	Mean Square	F	Sig.
Within	66855.559	2	3427.780	24.787	.000*
Error	16871.108	122	138.288		

* Correlation is significant at the 0.01 level (2-tailed)

Also, a Scheffé test was conducted to locate the exact differences between the means. Table 4 shows the results. As shown in Table 4, the students' overall listening strategy use in the No Task condition, significantly differs from that in the Easy Task and Difficult Task conditions and the difference lies between each two of the conditions. So in each task condition, the mean rate significantly differs from the other two conditions, i.e. students' reports of frequency of the overall listening comprehension strategies they use differ significantly from condition to condition (No Task, Easy Task, Difficult Task).

TABLE 4
Results of the Scheffé Test on Overall Reported Frequency

(I) FACTOR1	(J) FACTOR1	Mean Difference (I-J)	Std. Error	Sig.
Φ	A	7.403	1.528	.000*
	B	14.871	2.359	.000*
A	Φ	-7.403	1.528	.000*
	B	7.468	2.342	.002*
B	Φ	-14.871	2.359	.000*
	A	-7.468	2.342	.002*

* Correlation is significant at the 0.01 level (2-tailed).
Φ: No Task A: Easy Task B: Difficult Task

To examine the second research question (Considering each strategy subcategory under study, do high proficiency learners report different frequencies of strategy use across the three task conditions?), the researchers divided the question into three more specific questions regarding the three strategy subcategories, i.e. memory, cognitive, and compensation. Then each of the 62 questionnaires was divided into three parts, and frequency counts for each subcategory were done separately. For each subcategory there were three sets of frequency counts, each corresponding to one of the task

conditions. Subsequently, the mean frequencies were compared across the three task conditions through repeated measures ANOVA. For the strategy subcategories (memory, compensation, cognitive), the means and standard deviations are shown in Table 5.

For memory strategies, as shown in Table 5, the mean frequencies are different across the three task conditions, with NTC having the highest mean and DTC the lowest. The results of repeated measures ANOVA (Table 6) showed that the differences among the mean frequencies of memory strategies across the three tasks conditions were statistically meaningful, $F(2,122)= 6.039$, $p < .003$ or $F_{observed} = 6.039 > F_{critical} = 4.98$ with the significance level of .01. The results of the Scheffé test showed that the significant difference in the use of memory strategies was only between the No Task condition and the Difficult Task condition at .01 level of significance. Table 7 depicts the results of the Scheffé test for the strategy subcategories.

TABLE 5
Means and Standard Deviations for Frequency of Memory, Compensation and Cognitive Strategies Across the Task Condition

Subcategories		No Task	Easy Task	Difficult Task
Memory	N	62	62	62
	Mean	32.40	31.37	28.90
	Std. Deviation	5.68	8.12	5.57
Cognitive	N	62	62	62
	Mean	18.65	18.23	17.63
	Std. Deviation	3.47	3.67	4.05
Memory	N	62	62	62
	Mean	52.76	46.81	42.40
	Std. Deviation	7.07	7.68	10.20

TABLE 6
Results of ANOVA on Reported Frequency of Memory Strategies, Compensation and (c) Cognitive Strategies

Subcategories	Source	Sum of Squares	df	Mean Square	F	Sig.
Memory	Within	66855.559	2	200.522	6.039	.003*
	Error	16871.108	122	33.205		
Compensation	Within	32.333	2	16.167	1.663	.194
	Error	1186.333	122	9.727		

Cognitive	Within	3348.677	2	1674.339	6.039	.000*
	Error	6481.323	122	53.126		

* Correlation is significant at the 0.01 level (2-tailed).

For the second subcategory, i.e. compensation strategies, means and standard deviations for NTC and ETC were quite close (M_{NTC} = 18.65, SD_{NTC} =3.47; M_{ETC} = 18.23, SD_{ETC} = 3.67). Mean and standard deviation in DTC were 17.63 and 4.05 respectively. The results of the repeated measures ANOVA for this subcategory (Table 6) revealed that there was no significant difference in the students' compensation strategy use across the three task conditions ($F_{observed}$ = 1.663 < $F_{critical}$ = 4.98, with the significance level of .01). This shows that high proficiency learners did not report different frequencies of compensation strategy use across the three task conditions.

TABLE 7
Results of the Scheffé Test on Reported Frequency of Memory and Cognitive Strategies

	(I) FACTOR1	(J) FACTOR1	Mean Difference (I-J)	Std. Error	Sig.
Memory	Φ	A	1.032	.984	.298
		B	3.500	.912	.000*
	A	Φ	-1.032	.984	.298
		B	2.468	1.189	.042
	B	Φ	-3.500	.912	.000*
		A	-2.468	1.189	.042
Cognitive	Φ	A	5.952	.910	.000*
		B	10.355	1.476	.000*
	A	Φ	-5.952	.910	.000*
		B	4.403	1.461	.004*
	B	Φ	-10.355	1.476	.000*
		A	-4.403	1.461	.004*

* Correlation is significant at the 0.01 level (2-tailed).
Φ: No Task A: Easy Task B: Difficult Task

Finally, the third subcategory (cognitive strategies) was shown to be used by the students with significantly different frequencies across the three tasks conditions (ETC, DTC, NTC). Like memory strategies, the mean rates were descending across the task conditions. Table 5 depicts means and standard

deviations of frequencies across the tasks. The results of the repeated measures ANOVA (Table 6) showed that the differences in the strategy frequencies across the three task conditions were significantly meaningful, $F(2, 122)= 31.517$, $p< .000$ or $F_{observed} = 31.517 > F_{critical} = 4.98$. The results of the Scheffé test revealed that the difference lay between mean frequencies of each pair of the three task conditions. That is to say that the significant difference in reported frequencies of cognitive strategies is between NTC and ETC, NTC and DTC, and ETC and DTC at .01 significance level (Table 7).

DISCUSSION

The first research question was: Does task condition (No Task, Easy Task, Difficult Task) have any significant effect on high proficiency language learners' listening strategy use? The results showed that the students reported different frequencies of strategies across the three task conditions. This is quite similar to Ikeda and Takeuchi's (2000) finding which showed that high proficiency and low proficiency learners reported more frequent use of strategies in the No Task condition, compared to Easy and Difficult conditions. The results are in phase with Oxford et al.'s (2004) findings where high proficiency learners reported more frequent use of strategies in the No Task condition than in the Easy and Difficult task conditions. Considering the mean frequencies (Table 5), it is possible to speculate that the students over-reported their strategy use frequencies in the No Task condition. The empirical evidence from this study also provides support for Cohen's (2003) suggestion about learners' over-reporting or under-reporting their strategy use in task-absent conditions. More importantly, the findings support the idea that a good number of strategy assessment studies (e.g., Shmais, 2003), in which students were asked to fill out questionnaires without having to go through an actual task, might be proved to be questionable if the studies were to be run for a second time in task-present settings.

For the second research question, the results showed that high proficiency learners reported different frequencies of memory and cognitive strategy use across the three task conditions. The results, however, revealed no significant difference for compensation strategies across the conditions. One reason why the differences in using cognitive strategies in different task conditions and memory strategies across the No Task and the Difficult Task conditions were significantly meaningful might be the fact that the nature and kind of input can influence the process of completing a task (Skehan, 1996; Robinson, 2005). In the No Task condition, the students had to fill out the questionnaire without first completing a real listening task. They had to recite their previous experience of doing listening tasks as a criterion to make generalizations of the strategies they might have used when doing the tasks. This lack of input, then, could lead to over- or under-reporting of the strategies. However, in the task-present conditions, where the difficulty was mostly manipulated through the input (i.e. the listening text), students could use memory and cognitive strategies in a far more specific manner. This means that they could narrow down their choices to a number of highly specific strategies most relevant to the type of the input in each task. In other words, they did not have to make their choices from a very large repertoire of hypothetical strategies; rather, they could easily choose among a limited number of task-specific strategies. It is noteworthy that in using memory strategies, the students did not report statistically meaningful differences between NTC and ETC as well as ETC and DTC. The only significant difference lay between the No Task condition and the Difficult Task condition. This, in some way, shows that for the Difficult Task condition the students were more selective in their use of memory strategies and the strategies they used were significantly affected by the high difficulty level of the task in comparison to No Task and Easy Task conditions.

The results about compensation strategies turned out to be unexpected. In fact, before data analysis, taking into account the findings in Oxford and Ehrman (1995), it was the researchers' intuition that students might use compensation strategies more frequently than other types in the task-present

conditions, especially in the Difficult Task condition. But the results proved the opposite. The mean frequencies of compensation strategies were very close and relatively low for NTC (mean= 18.65), ETC (18.23), and DTC (17.63) with quite constant and low standard deviations (3.44, 3.67, and 4.05). Virtual, untrue responses may be one reason for low frequency of compensation strategy use among the students. As a matter of fact, these students might consider reporting of such strategies embarrassing or belittling for advanced language learners. The researchers arrived at such a conclusion from a question asked by one of the students, while filling out the questionnaire: "Do you really think that we use such strategies [pointing to the questions in the subcategory of compensation strategies]? This is a TOEFL class, not an elementary class." Very interestingly, this can be linked to what Oxford and Ehrman (1995) found about the relationship between language learning strategies and language learners' ego boundaries. Oxford and Ehrman (1995) found that compensation strategy use had a wide range of significant correlations with thinner ego boundaries, and compensation strategy users tended to be highly flexible. Of course, one question asked by one participant should not be generalized to be a reliable result, but the question, considering the previous findings (Oxford & Ehrman, 1995), seems to be worth further research.

In such studies, when addressing more deep-seated levels of analysis, quantitative data may also reveal a number of practical hints for linking high frequency strategies to designing tasks for specific purposes. For instance, data analysis may possibly show that for a certain task type, students report very high frequencies of one or more strategies. This, then, can be a good start for considering that strategy as one representative element of the process of task completion and for designing similar tasks more insightfully. In this study, for example, picturing the scene was reported to be of utmost frequency for the easy task, which was a simple short story. However, using lexical clues and background knowledge proved to have the highest frequency for the difficult task, which was an authentic book summary.

CONCLUSION

After Ikeda and Takeuchi (2000) and Oxford et al. (2004), the results of this study confirm that presence and absence of tasks as well as their degree of difficulty can have an effect on the students' reports of their strategy use. Apparently, students' use of strategies and their reported frequencies seem to be affected by a variety of task characteristics including task difficulty. If task characteristics (cultural load, demand of the process, task setting, instruction, etc.) can affect the type and frequencies of strategies students use, then it would be very important to research different aspects of the issue in detail.

In summary, such studies promise numerous implications for both pedagogical and theoretical aspects of language teaching. Theoretically speaking, the findings help us develop a more accurate framework for designing tasks and use them more carefully in the architecture of our strategy-based courses. In practice, however, the results might affect the course of language learning task design, the nature of language learning strategy training in L2 classrooms, and also the approaches to strategy assessment in different language teaching/learning contexts. In task-based language teaching programs, for example, the task designer/teacher would take into account frequently used strategies with easy and difficult tasks to modify the difficulty level of tasks and do more accurate sequencing and grading. It is also possible to link strategy training and task-based instruction when developing the syllabus.

Finally, there is room for further research into the present topic, considering key issues such as task factors other than task difficulty (e.g., task setting, task type, authenticity, input genre), learner factors and output, assessment methods alternatives, etc. Also, task designers might be eager to reverse the process through conducting what the researchers call 'strategy-based task evaluation' to open new horizons of specific purpose task design.

THE AUTHORS

Mahmood Reza Atai is assistant professor of TEFL at Tehran Tarbiat Moallem University. His current research interests include teaching language skills and components and ESAP. His recent publications include *Content schemata, linguistic simplification, and EFL learners' comprehension and recall* with M.H. Keshavarz & H. Ahmadi (2007), *On the Effect of Text Authenticity & Genre on EFL Learners' Performance in C-Tests* with M. Soleimany (2008), and *A Cross-Cultural Genre Study of Hedging Devices in Discussion Section of Applied linguistics Research Articles* with L. Sadr (2008).
Email: atai@saba.tmu.ac.ir

Seyyed Mohammad Reza Hashemi is Deputy Director for Research & Planning at the Iran Language Institute. He has an MA in TEFL from Khatam University. He is currently a Ph.D. candidate in TEFL at Tehran Tarbiat Moallem University. He has taught English courses at colleges and universities. His areas of research interest are task-based instruction, materials production and teacher education.
Email: hashemi_ili@yahoo.com

REFERENCES

Afsarnia, M. (2000). *Strategy inventory for listening comprehension. Its psychometric characteristics, factor analysis, and the interrelated effects of gender, proficiency level, listening comprehension strategies and listening comprehension scores in authentic listening comprehension texts and tasks.* Unpublished MA thesis. Shiraz University: Shiraz.

Bacon, S. M. (1992). The relationship between gender, comprehension, process strategies, and cognitive affective response and FL listening. *Modern Language Journal*, *76*, 160-177.

Brindley, G., & Slatyer, H. (2002). Exploring task difficulty in ESL listening assessment. *Language Testing*, 369-395.

Buck, G. (2001). *Assessing listening*. Cambridge: Cambridge University Press.

Bygate, M., Skehan, P., & Swain, M. (eds). (2001). *Researching pedagogic tasks, second language learning, teaching, and testing.* Harlow: Longman.

Chamot, A. U., Barnhardt, S., El-Diary, P. B., & Robins, J. (1996). Methods for teaching learner strategies in the foreign language classroom. In R. Oxford (Ed.), *Language learning strategies around the world: Cross-cultural perspectives* (pp. 157-187). Manoa, HI: University of Hawai'i Press.

Champy, J., & Nohria, N. (2000). *Arc of ambition: Defining the leadership journey*. Perseus Books.

Cho, Y., Kim, H. J., & Leung, S. (2001). *The effect of presence and difficulty of task on strategy use*. Paper presented at the Annual Conference of American Association for Applied Linguistics, St. Louis, MO.

Cohen, A., & Scott, K. (1996). A synthesis of approaches to assessing language learning strategies. In R. Oxford (Ed.), *Language learning strategies around the world: Cross-cultural perspectives* (pp. 89-106). Manoa, HI: University of Hawai'i Press.

Cohen, A. (2003). The learners' side of foreign language learning: Where do styles, strategies, and tasks meet? *IRAL*, *41*, 279-291.

Cohen, A. (1998). *Strategies for learning and using a second language*. Harlow, Essex: Longman.

Cobuild English Dictionary for Advanced Learners (2001). Glascow: HarperCollins Publishers.

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.

Griffiths, R. (1990). Speech rate and narrative spoken comprehension: A preliminary study in time benefit analysis. *Language Learning*, *40*, 311-336.

Hsiao, T-Y., & Oxford, R. L. (2002). Comparing theories of language learning strategies: A confirmatory factor analysis. *Modern Language Journal, 86*, 360-383.

Helgesen, M., & Brown, S. (1995). *Active listening: Introducing skills for understanding.* Cambridge: Cambridge University Press.

Ikeda, M., & Takeuchi, O. (2000). Tasks and strategy use. *JACET bulletin, 31*, 21-32.

Nunan, D. (2004). *Task-based language teaching*. Cambridge: Cambridge University Press.

O'Malley, J. M., & Chamot, A. U. (1995). A factor-analytic study of language learning strategies use: Interpretation from information processing theory and social psychology. *Modern Language Journal*, *77*, 11-22.

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

New York: Newbury House.

Oxford, R. L. (2001). Language learning styles and strategies. In M. Celce-Murcia (Ed.), *Teaching English as a second or foreign language* (pp. 359-366). Boston, MA: Heinle & Heinle.

Oxford, R. L., Cho, Y., Leung, S., & Kim, H. (2004). Effect of the presence and difficulty of task on strategy use: An exploratory study. *IRAL, 42,* 1-47.

Oxford, R. L., & Ehrman, M. E. (1995). Adults' language learning strategies in an intensive foreign language program in the united states. *System*, *23*(3), 359-386.

Philips, D. (2001). *Longman complete course for the TOEFL.* Harlow Essex: Longman.

Robinson, P. (2005). Cognitive complexity and task sequencing: studies in a componential framework for second language task design. *IRAL*, *43*, 1-32.

Shmais, W. A. (2003). Language learning strategy use in Palestine. *TESL-EJ*, *7*(2), A-3, 2-13.

Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics*, *17*, 38-62.

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

Slatyer, H., Brindley, G., & Wiggleworth, G. (2000). *Task difficulty in ESL listening assessment.* Paper presented at the 22nd Language Testing Research Colloquium, Vancouver.

Vandergrift, L. (2002). It was nice to see that our predictions were wrong: Developing metacognition in L2 listening comprehension. *Canadian Modern Language Review, 58*(4), 555-575.

APPENDIX A.
The listening tasks

Task A. Listen to a short story called "The Medicine Pipe." Check *True* or *False*. Then fill out the questionnaire. Your answers to the items on the questionnaire should be based on the listening strategies you use when you are doing this task. So you'd better be conscious of the strategies you will use.

	True	False
1. Two old men walked in a big garden.	☐	☐
2. The men were lost in the snow.	☐	☐
3. It was a summer day.	☐	☐
4. They saw an old woman who was carrying a bunch of flowers.	☐	☐
5. One of the men ran toward the woman.	☐	☐
6. The woman had some food with her.	☐	☐
7. The woman changed into a buffalo.	☐	☐
8. The men ran away from the buffalo.	☐	☐
9. Both men could find their way back to their camp.	☐	☐
10. The buffalo changed into a pipe.	☐	☐

Task B. Listen to a book summary. The book is called "The Arc of Ambition." Check *True* or *False*. Then fill out the questionnaire. Your answers to the items on the questionnaire should be based on the listening strategies you use when you are doing this task. So you'd better be conscious of the strategies you will use.

	True	False
1. Ambitious ideas help people finalize all their spectacular plans.	☐	☐
2. Ambition is the only thing a financial dynasty needs to acquire.	☐	☐
3. According to the authors of the book, abusing ambition is invaluable.	☐	☐
4. The arc of ambition, as stated by the authors, is a feature	☐	☐

that illustrates how a success's life changes course.		
5. Seeing what others do not see is a stage where ambition increases in a hit-and-miss way.	☐	☐
6. For Del the curve rose at an early age.	☐	☐
7. The curve of the arc of ambition rises swiftly for different individuals.	☐	☐
8. Ambition is closely related to how old a person is.	☐	☐
9. Following one's ambition might sometimes lead to precarious situations.	☐	☐
10. Ambitious people cannot bring their simple ideas to life.	☐	☐