The Effect of Processing Instruction and Output-Based Instruction on the Interpretation and Production of English Causatives

Parviz Birjandi [*]
Professor of TEFL, *Islamic Azad University, Science and Research Branch, Tehran, Iran*
&
Jamileh Rahemi [†]
PhD Student in TEFL, Islamic Azad University, Science and Research Branch, Tehran,Iran

Abstract
This study was intended to compare processing instruction (PI), an input-based approach to L2 grammar instruction developed by VanPatten (1996), with an output-oriented type of instruction (OI) to assess their relative effects on learners' ability to interpret and produce English causatives. A pretest and posttest (immediate and delayed effects) design was used. 151 university students from four intact classes were randomly assigned to three treatment groups of PI, OI, and EI (Explicit-information-only) and one uninstructed control group (C). Students were assessed on interpretation and controlled written production tasks at the sentence level. Within-group comparisons indicated that the three instructional options, as compared to the control group, resulted in some kind of knowledge gain in both interpretation and production tasks, but the gains were not equal. The results of between-group comparisons contradicted VanPatten's claims about the superiority of PI over OI. While PI and OI were equally better than EI on interpretation tasks, OI group outperformed both PI and EI on production tasks. No significant difference was found between PI and EI on production tasks. The same results were obtained after a one-month interval, reflecting the durability of the instructional effects on the interpretation and production of the target structure.

[*] *E-mail address*: birjandi@srbiau.ac.ir

[†] *E-mail address*: Rahemi@srbiau.ac.ir

Any correspondence should be done to this e-mail: jamilehr@yahoo.com
Address: Islamic Azad University, Science and Research Branch, Pounak Square, Tehran, Iran

Keywords: Input processing theory; Output hypothesis; Processing instruction; Output-based instruction; Interpretation; Production; English causatives

Introduction

Today there is a wide consensus that learners' exposure to input plays an essential role in second language acquisition so that it seems rather impossible to conceive learning a new language without considering a role for input in some form or other. However, what is still disputable is how input can better be transformed into intake (Qin, 2008). VanPatten's (1996) model of input processing (IP) is mainly concerned with this debated issue. Motivated by the perspectives in cognitive psychology, VanPatten postulates that learners can attend to only a limited amount of incoming input at a given time; this is why form and meaning may compete for attention resources during input processing, and since the communicative goal of learners is to understand the content of messages rather than understanding how that message is encoded, learners tend to process input for content (meaning) before they process it for the code (form). He identifies three sets of acquisitional processes named as 'input, intake, and developing system' which are responsible for taking linguistic data in the input, converting it to intake, and making the intake available to the developing system. For VanPatten, what learners do with the input during comprehension determines how intake is derived, and IP is specifically concerned with those psycholinguistic strategies by which learners derive intake from input. Focusing on the strategies and constraints that govern learners' input processing, VanPatten holds that L2 learners may use some default strategies which inhibit them from processing input efficiently. Thus, changing the way learners process input should affect the quality and/or the quantity of intake, and consequently affect the developing system. The IP assumptions are manifested in two basic principles of the theory.

1. The Primacy of Meaning Principle: L2 learners process input for meaning before processing it for form.
2. The First Noun Principle: Learners tend to assign the role of subject or agent to the first (pro) noun they encounter in a sentence. (see VanPatten, 2002b for the corollaries of the principles)

As Gass and Selinker (2008) state, the two principles are to account for how acquisition takes place- from meaning (principle 1) to form (principle 2). For VanPatten (2002a), "input provides the data, IP makes (certain) data available for acquisition, other internal mechanisms accommodate data into the system …, and

output helps learners become communicators and, again, may help them become better processors of input" (p. 762).

The assumptions led to the development of a new grammar instructional approach known as 'processing instruction' (PI) which mainly aims at helping learners to readjust their inefficient processing strategies. Unlike output-based instruction which emphasizes grammar rules and oral/written production practice, the purpose of PI is to alter how learners process the input and to make them develop better form-meaning mappings which result in a grammatically richer intake. VanPatten (2002b) assumes that an explicit type of instruction with a focus on learners' processing strategies is more effective than approaches which require learners to produce language too prematurely.

A considerable body of research (e.g., Allen, 2000; Bentai, 2001, 2004, and 2005; Collentine, 1998; DeKeyser & Sokalski, 1996; Erlam, 2003; Nagata, 1998a, 1998b; Qin, 2008; Radwan, 2009; and Toth, 2006) has been conducted to evaluate the efficiency of PI by comparing it with other types of instruction, including output-based instruction. Not all the studies did produce convincing evidence for the advantage of PI over OI. The following section presents a brief review of some of the related studies which are more relevant to the present study.

Related Empirical Studies

In their original study, VanPatten and Cadierno (1993) compared PI with a traditional type of output-oriented instruction which involved grammar explanations followed by output practice. The results indicated that the PI group outperformed the OI group on the tasks measuring the interpretation of Spanish clitic object pronouns, and equally well on production tasks, though they never produced the grammar feature during the instruction phase. They concluded that instruction is seemingly more beneficial if it is directed at how learners perceive and process input rather than focusing on practice via output. In another study on Spanish past tense verb morphology, Cadierno (1995) obtained the same result: PI was superior to OI in interpretation tasks and equally well in production tasks. Benati's (2001) research also provided some supporting evidence that PI has more positive and more durable effects on the acquisition of Italian verbal morphology than OI. The findings of a subsequent study by Benati (2004) strengthened the evidence regarding the positive effects of structured input practice in PI with a different structure (gender agreement in Italian adjectives) and a more spontaneous

and communicative task. Similarly, in his next experimental study on the acquisition of English past simple tense, Benati (2005) found that PI group outperformed both the traditional and the meaning-based OI groups in the interpretation test while the three groups improved equally well on their production test. In a recent research by VanPatten et al. (2009) the overall superiority of PI was found over dictogloss tasks too.

In contrast to the studies with favorable results for PI, some other studies have failed to produce convincing evidence for the superiority of PI over OI. For example, DeKeyser and Sokalski's study (1996) on Spanish clitic direct object pronoun and conditional structures indicated that input practice worked better for comprehension skills but output practice was better for improving production skills. In addition, the relative effectiveness of production versus comprehension practice depends on the complexity of the target structure and on the delay between practice and testing. Toth (2006) also suggested a higher role for output in acquisition involving attention and metalinguistic analyses of L2 structures. Allen (2000) did not find any advantages for PI over OI group in how they interpreted the French causative, while OI proved to be more effective in enabling learners to produce the form. In a study comparing the effects of the two approaches on the students' ability to comprehend and produce direct object pronouns in French, Erlam (2003) also found greater gains for OI group. Celik-Yazici's study (2007) showed no significant difference between PI and OI in comprehension and production tasks as far as the development of English WH-questions by Turkish EFL learners was concerned. Nagata (1998a, 1998b) concluded that the input processing instruction is not necessarily more effective than production practice on the ground that in both studies output-focused groups outperformed the input-focused groups in the production, and they performed equally well in the comprehension of the structures.

Purposes of the Study and Research Questions

Given the conflicting findings of the research into the efficiency of PI approach, the present study was planned (1) to compare the relative effectiveness of processing instruction and a meaning-based type of output instruction on the Iranian L2 learners' ability to interpret and produce sentences containing English causatives, and (2) to see which of the two instructional approaches might produce more persistent or durable results. It is noteworthy that, to the researchers' knowledge, no research has been done to date into the effectiveness of the two

instructional options on the interpretation and production ability of Iranian learners of English. Although the overall design of the current study was similar to the previous research reviewed above, it was not an exact replication of the studies because the following differences: (a) it used a larger sample size; (b) two new types of assessment tasks (a translation task in the interpretation test and a paraphrasing task in the production test) were added to the typical test tasks used in the previous studies (see VanPatten, 2002a for typical samples). The task type was not a variable of the study, though; and (c) beside OI and PI, the study included one more treatment group namely 'explicit-information-only' group (EI) who received the same explicit information about the target structure that the other two treatment groups did but without getting involved in doing any types of follow-up activities. The aim was to ensure that the advantages of instruction, if any, were attributed to the kind(s) of activities used in their classes and not to the explicit information given before the activities. Only in a few studies, including VanPatten and Oikennon (1996), EI was included in the design of the research. To achieve the aims of the present study, the following research questions were developed:

1. Are there any significant differences among the study groups regarding their performance on the immediate post-test measuring the interpretation of English causatives?
2. Are there any significant differences among the study groups regarding their performance on the delayed post-test measuring the interpretation of English causatives?
3. Are there any significant differences among the study groups regarding their performance on the immediate post-test measuring the production of English causatives?
4. Are there any significant differences among the study groups regarding their performance on the delayed post-test measuring the production of English causatives?

Method

Participants

169 students from four intact classes at Islamic Azad University, Naragh Branch, participated in the study. The participants had enrolled in 'General English' course as a required part of the university curriculum. The four classes were then randomly assigned to three treatment groups and one control group in order to reduce the effects of extraneous variables and selection bias. It is necessary to mention that only 151 students (81 males and 70 females) were chosen from the

original pool as the main subjects of the study. The final sample included the students who (a) were of similar level of language proficiency, measured by the proficiency test of PET, (b) had little or no previous knowledge of the target structure, measured by a grammar pretest, (c) had no exposure to the target form outside the class during the treatment period, examined by a brief background questionnaire, (c) attended all training, treatment, and assessment sessions, and (d) completed all the assessment measures. The final sizes of the groups were as follows: PI (35: 18 males and 17 females), OI (40: 23 males and 17 females), EI (40: 21 males and 19 females), and C (36: 19 males and 17 females). The age range was from 18 to 24. It is noteworthy that neither age nor gender was a variable of the study.

Target Structure

The targeted forms were the English causative verbs to which the participants had not been exposed before. They were delimited to the periphrastic causatives *have* and *get* both in active and passive causative sentences, i.e. [NP1-HAVE-NP2-STEM- (NP3)], [NP1-HAVE-NP2- P.P of VERB], [NP1-GET-NP2-toINF-(NP3)], and [NP1-GET-NP2- P.P of VERB]. The theoretical and pedagogical reasons for choosing the target structure were as follows: First of all, causative structure was deemed a good target for the first-noun principle of IP theory (VanPatten, 2002a); It is argued that learners usually tend to use a default processing strategy that assigns the role of subject to the first noun or phrase they see or hear in the input containing a causative verb, and PI can help learners adjust this false strategy. For example, in the sentence 'my teacher had me work hard' the tendency would be to interpret that it was 'teacher', and not 'I' who worked hard. Second, English causative structure is among the most problematic constructions for Iranian students both to interpret and to produce. This difficulty might be due to the fact that the causative structure is different between English and Persian languages. Persian verbs are usually causativised by affixing the causative infix /*än/. T*he infix changes an intransitive verb to a causative transitive one. For example, the intransitive verb *Khandidan* (to laugh) is simply changed to transitive verb of *Khandandan* (to make somebody laugh) by inserting the morpheme /*a:n/* into the word (Bateni, 1384). This strategy is called morphological causativisation which mostly occurs in languages such as Persian which are more inflectional than agglutinative. In English, in contrast, the lexical causativisation and periphrastic causativisation are more common (Lotfi, 2008). Another reason why interpreting and producing accurate causative sentences seem demanding for learners is that the

first meanings of 'have and get' that they usually encounter and learn are 'to possess' and 'to receive' respectively, so many Iranian learners have problem with identifying their causative usages. Finally, producing accurate causative sentences requires learners to be aware of the correct form of the verb following each causative verb too.

Procedures

Phase 1. Developing and Trialing Treatment Packs

Two separate packs of PI and OI were produced based on VanPatten's (2002a) guidelines and the samples available in the literature. Since the instructional packs (as well as the assessment measures) included some picture-cued activities, an artist was asked to draw all the pictures required for the instruction and assessment tasks. The activities in PI and OI packs were developed at the sentence-level. As Swain (1998) states, pushing learners to produce at the sentence level is sufficient for learning to occur. Marsden (2006) also reminds that IP model is concerned with sentence-level processing with the aim of altering interpretation strategies. The instruction packs were balanced in terms of the explicit information about the structure, the vocabulary, the total amount of time devoted to explicit instruction and follow-up practice, and the number of activities. Thus, OI and PI packs differed only as to the type of the follow-up activities. Besides, in order to minimize the lexical load of the activities for the participants, the vocabulary items were chosen from those in their high school books.

PI materials in the present study included the three essential components that typical PI (VanPatten, 1996, 2002a) is often associated with: (1) a script of metalinguistic information aimed at introducing the causative form explicitly and briefly; (2) an explicit reminder of typical inefficient input processing strategy (i.e. *the first noun principle* in the present study) that learners usually use in interpreting or producing causative sentences; and (3) structured input activities prompting the learners to make form–meaning mappings. They included 20 pictorial and non-pictorial *referential* activities and 10 non-pictorial *affective* tasks presented in written and oral modes. Referential activities are defined by VanPatten (2002a) as "those for which there is a right or wrong answer and for which the learner must rely on the targeted grammatical form to get meaning" (p. 766); affective tasks, aimed at providing more exemplars of the target form in the input, were designed to engage learners in processing information about the real world and to push them to process (not produce) the information presented in the input containing

causative verbs by expressing their personal opinions about the sentences (see Appendix A for some samples).

OI pack was comprised of two components: (1) the same script of metalinguistic information about the causative structure in the PI pack, and (2) the meaning-based production activities which included 30 pictorial and non-pictorial sentences requiring the participants to use the new form and complete the written tasks. The activities were designed in a way that required learners to attend both to the form and to the meaning and make a connection between them; otherwise, they could not complete the tasks correctly (Appendix B). Meaning-based output activities were used to respond to the criticisms made about the studies in which mechanical production activities were employed (see Erlam, 2003 and Toth, 2006 for more details). It needs to be noted that at the beginning a far larger number of OI and PI activities were developed than what was really needed. After both packs of instructional materials were prepared, four Iranian English teachers who were experienced in teaching the target structure were asked to read the packs and make comments on the form, the content, the number of activities, etc. In addition, both packs were tested on two 'General English' classes (78 students); Piloting the packs was done in order to (a) determine the possible practical problems in applying the instructional treatments in EFL classes of Iran, (b) determine the time required for completing the tasks in large classes, (c) examine the quality of the drawings and their relevance to the sentences, and (d) to receive the students' feedback regarding the quality and quantity of the tasks. The fruitful views of the teachers and the students' feedback resulted in identifying some practical difficulties, decreasing the number of activities (30 for each pack), and changing or omitting some of the drawings and ambiguous sentences. PI and IO packs were then ready to be used in the main study.

Phase 2. Developing and Piloting Assessment Measure

The original test consisted of 46 pictorial and non- pictorial items, all in written modality. It was aimed at assessing the participants' interpretation ability (23 items) and their production knowledge (23 items). Pictures were used to contextualize the items and to restrict the descriptions produced by participants for more objective scoring. For the sake of validating the content of the test and ensuring the adequacy, appropriateness, and efficiency of the test instructions, pictures, timing, wording, and scoring procedures, the original pool of items was reviewed by a panel of six Iranian university faculty members and three high school teachers who

had many years of experience in teaching English grammar. As a result of the panel review, several items, instructions, and pictures were either eliminated or modified. Besides, two English native speakers were asked to read the items and omit or modify the items that seemed odd or sounded unnatural. The remaining 32 items (15 interpretation and 17 production items) were then trialed on a sample of 130 students enrolling in 'General English' course to examine the item characteristics. The item facility and the item discrimination of the interpretation items as well as the choice distribution of their options were examined. The analyses resulted in elimination of three problematic interpretation items and changing some choices. Due to the different nature of production items, the comments made by the panel group and a number of students who were interviewed after taking the tests were used as the criteria for choosing the appropriate items on production section of the test. Their views resulted in omitting nine production items too. The final distribution of the test items was as follows: The interpretation section of the test included 22 pictorial and non-pictorial items, 10 of which were distracters, while the production section included 16 written controlled production items, 8 of which were distracters (Appendix C). As for the reliability of the test, an internal-consistency method (Cronbach's alpha) was used. The reliability estimates for interpretation and production sections of the test were 0.83 and 0.80 respectively.

Following VanPatten and Cadierno (1993), another similar version of the test was created to use a split-block design in test administration and to control for the test learning effects. The two versions (A & B) were exactly the same in terms of the format, test tasks, the overall length, the number of target and distracter items, the instructions, the vocabulary, and timing. Item sequences were, however, reshuffled and the content words were substituted for other equally familiar words.

Phase 3. Main Study

The main study phase lasted a whole semester (12 weeks) including the time required for administering the assessment measures as well as the training sessions. The instruction and the administration of the assessment measures in all the groups were done by one of the researchers of the study in order to control for the teacher variable and to avoid the possible threats to the internal validity of the study. A brief background questionnaire as well as the proficiency test of PET was administered within the first two weeks of the term to ensure the homogeneity of the participants (169 subjects) before the commencement of the study treatments.

(It is necessary to mention that the speaking section of the proficiency test was not administered due to practical limitations). The mean (M= 27.91) and standard deviation (SD= 4.80) of the PET were used as the criteria for choosing the participants. To reduce the effect of selection bias, the subjects who scored higher than M-plus-one SD and those whose scores fell below M-minus-one SD were considered as high and low proficient respectively, and were thus excluded from the final data pool. Only the data belonging to the students who scored within the range of M-plus-one SD and M-minus-one SD (i.e., between the scores of 23.11 and 32.77) were included in the final data pool.

In the third week of the semester, the pretest was administered to examine the participants' interpretation and production knowledge of the target structure before the treatment commencement, and consequently, to take care of the internal validity of the results. Following Toth (2006), the students who scored above the arbitrary cut-off score of 50% on each section of the pretest were eliminated from the final analysis. The production part of the test was administered before the interpretation section in order to minimize the possibility that production items would serve as input for interpretation tasks and would then contaminate the results.

Since the participants were not familiar with the new approaches, the two treatment groups of OI and PI were trained to do the tasks designed for them before the commencement of the treatment phase. The training lasted two weeks, and the structures were some grammar features other than the target ones (the simple, progressive, and perfect modes of the present and past tenses. The training was followed by the main treatment phase that took two weeks. In order to be consistent with VanPatten's (1996) guidelines that one thing should be presented and practiced at a time, in the first week the causative structure containing the verb 'have' in two modes of active and passive was introduced and practiced; the following week, which started with a 15-min warm-up consisting of a short review of the verb 'have', was devoted to teaching and practicing the causative verb 'get'. An attempt was made to equalize the following factors for the groups in order to prevent their possible confounding effects on the results of the study and to ensure that any possible gains made by the learners were related only to the differences in the follow-up activities they received after getting the same explicit information about the target forms: (a) none of the groups was aware of the purposes of the measurements and instructional treatments; they were told that all the tests and the

tasks were parts of the term schedule so their active participation was required; (b) all instruction and assessment took place in the students' regular class hours; (c) all the worksheets were collected after each treatment session and no homework was assigned to the participants during the treatment period; (d) during the study, the participants received no instruction on the target features other than that given in the class; and (e) due to the low language proficiency of the participants, the instruction was given in Persian language to ensure the subjects' understanding.

Thus, the three treatment groups first received the same explicit information about the target forms. Then the PI and OI groups worked on the specific tasks of PI and OI packs designed for them. At no time during instruction, were the PI learners allowed to produce the target feature, while OI group was pushed to use the newly taught forms immediately after receiving the relevant explicit information. In contrast to the PI group who was informed about where to focus their processing effort, the OI learners were not given any information on the typical processing problems. The EI group was given just the same brief explicit information about the structure with no opportunity to practice the newly taught features. The control group took all the tests administered to the other groups, but their whole class time was spent on the instruction targeting the development of reading comprehension skills, with no reference to the target grammar items.

To investigate the short-term effects of the instructional treatments, the first posttest was administered the following week after the completion of instruction. Using the split-block design, the delayed post-test was administered after four weeks to see whether the advantage(s) of each instructional approach within the treatment groups, if any, would maintain within a one-month interval or not. During the interval, the classes continued working on a reading comprehension book with no reference to grammar forms.

Results

For the statistical analyses, only the target items were scored. The scoring procedure for the interpretation section (12 items) was simple; one point was assigned for each correct answer, and a zero point for each incorrect response. The scoring procedure for the production part (8 items) was more complex; A zero point was assigned to a fully incorrect response, and one point and a half was awarded to a fully correct answer which met the following criteria: the correct position of the causative verb, the inclusion and correct position of the object, the

correct form of the main verb, and the rest of sentence were assigned 0.25, 0.5, 0.5, and 0.25 respectively. This analytical scoring scale helped to reveal the intermediate effect of instruction (VanPatten & Cadierno, 1993) and to ensure the objectivity of the scoring procedure for the written production items. However, in order to achieve the inter-rater reliability of the production part of the tests, two raters scored the production items based on the same analytical scale. Coefficient alpha was computed for the three production scores: 0.98, 0.99, and 0.99 for the pretest, immediate and delayed posttests respectively. To confirm the normality of the distribution, the Kolmogorov-Smirnov test was run. As shown in Table 1, the p-values for the three test administrations are greater than 0.05 which means that the sample was drawn from a normal distribution.

Table 1
One-Sample Kolmogorov-Smirnov Test

		Pretest	Post-test 1	Post-test 2
N		151	151	151
Normal Parameters(a,b)	Mean	3,1291	6,4669	7,3576
	SD	1,65576	3,52830	3,94858
Most Extreme Differences	Absolute	,136	,098	,079
	Positive	,136	,098	,079
	Negative	-,073	-,047	-,061
Kolmogorov-Smirnov Z		1,677	1,208	,972
Asymp. Sig. (2-tailed)		,07	,108	,301

Then the subjects' scores on both sections of the pretest were submitted into separate one-way ANOVA analyses (significance level was set at .05). The results revealed no significant differences among the four classes before instruction, reflecting that the differences among the groups on the posttests were not due to their prior knowledge of the target structure: F (3,147)=1.32, p= 0.269 for the interpretation section and F (3,147)=0.62, p=0.602 for the production section.

In order to answer the first research question, first the within-group changes over time were analyzed. Table 2 displays the group means and standard deviations of interpretation section over the three administrations as well as the results of paired-samples t-tests. The results indicate that although the control group showed some minor changes from the pretest to the follow-up posttests on the interpretation tasks, the changes were not high enough to make significant

differences in the group performance over time. In contrast, the three treatment groups improved significantly from the pretest to the first posttest as far as their interpretation ability was concerned: t (35)=-9.67, $p<.05$ for the PI, t (40)= -7.57, $p<.05$ for the OI, and t (40)=-5.93 , $p<.05$ for the EI group. The means also showed that the PI and OI groups made higher gains on posttest 1 as compared to the EI group. In addition, although the treatment groups showed a decline on posttest 2, their performance was still significantly higher than that of the pretest.

Table 2
Descriptive statistics and paired-samples t-tests for interpretation section

		Pretest		Post test1		Post test2		Pre-Post1	Post1-Post2	Pre-Post2
Group	n	M	SD	M	SD	M	SD	t/ρ	t/ρ	t/ρ
PI	35	2,34	1,39	6,00	1,89	5,57	1,92	-9,67 0,000	1,106 0,276	-8,69 0,000
OI	40	2,82	1,48	6,12	2,30	5,80	2,31	-7,57 0,000	0,77 0,443	-6,05 0,000
EI	40	2,30	1,26	4,60	2,15	4,10	2,30	-5,93 0,000	1,24 0,221	-4,18 0,000
C	36	2,69	1,45	2,47	1,46	2,61	1,43	1,03 0,308	-0,824 0,416	0,337 0,738

Although the results of within-group comparisons indicated that the instructional options were all effective in improving learners' ability to interpret the target form, it was not clear what effect there might be for each instructional treatment when the three groups were compared with each other on each single test. Accordingly, in order to determine whether there was/were any particular type(s) of instruction which could improve the learners' ability to interpret (and to produce) the causative sentences more effectively than others, between-group comparisons (see Table 3) were carried out by submitting the subjects' scores on the first posttest into one-way ANOVA (significance level was set at .05).

Table 3
The results of ANOVAs for the interpretation and production sections

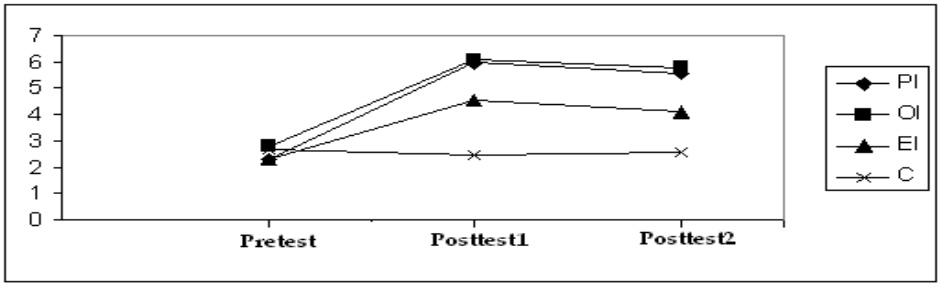
	Pretest	Post test1		Post test2	
Test	F/p	F/p	μ2	F/p	μ2
Interpretation Items	1,32 0,269	26,57 0,000	0,35	19,32 0,000	0,28
Production Items	0,622 0,602	22,82 0,000	0,31	33,69 0,000	0,40

As it is shown in Table 3, there was a significant difference in how the four groups interpreted the causative sentences on posttest 1: F (3, 147) =26.57, p<0.05. Using the formula suggested in Dornyei (2007), effect sizes were computed manually to provide information about the magnitude of the observed differences. The figure (eta squared= 0.35) indicates that not only the differences were significant, but also the magnitude was large enough to be really meaningful. Thus, the answer to research question one is *Yes*, i.e., there were significant differences among the groups as for their performance on the immediate posttest measuring the interpretation of the causative structure. Post hoc Tukey HSD tests indicated no significant differences among OI and PI, reflecting that the two types of instruction improved the interpretation ability of the learners equally well. However, they were significantly better than both EI group, who was given no follow-up activities, and the no-instruction control group. The performance of the four groups on the interpretation tasks can be summarized as follows: PI=OI> EI >C (with > meaning 'better than', and = indicating no significant differences).

As for the second research question, i.e., the durability of the instruction effects over one month, the results of within-group comparisons (see Table 2) indicated that there were no significant difference between the mean scores of the treatment groups on the posttest 1 and posttest 2 in the interpretation section [t (35)=1.10, $p<.05$ for the PI, t (40)= 0.77, $p<.05$ for the OI, and t (40)=1.24 , $p<.05$ for the EI], while the mean scores on the pretest were significantly different from the posttest 2 (35)=-8.69, $p<.05$ for the PI, t (40)=-6.05, $p<.05$ for the OI, and t (40)=-4.18 , $p<.05$ for the EI]. In other words, the three treatment groups maintained their improved ability to interpret English causatives not only from the immediate posttest to the delayed posttest, but also from the pretest to the delayed posttest. In order to examine the differences among the groups in the effect of instruction after a one-month interval, the delayed posttest scores of the interpretation section were

submitted into one-way ANOVAs too. The results (Table 3) again revealed a significant difference in how the four groups interpreted the causative sentences on the delayed posttest: F (3/147) =19.327, p= 0.000; the eta squared (0.28) also suggested considerable effect sizes for the result. Thus, the answer to the second research question is also *positive,* i.e., there were significant differences among the groups as for their performance on the delayed posttest measuring the interpretation of the target structure, which suggest that the positive effects for instruction on the interpretation tasks were retained equally well over a one-month interval by the instructional groups. Post hoc Tukey HSD tests also showed exactly the same results emerged from the analysis of the immediate posttest scores in the interpretation section: PI=OI> EI >C. In other words, the advantages of the PI and OI groups over EI in interpretation tasks held over time. Visual representations of the gains and longer-term effects for the interpretation tasks across the pretest and the follow-up tests are shown in Figure 1.

Figure 1
Interpretation group means



The third research question was aimed at examining the groups' performance on production tasks. The means as well as the results of paired-samples t-tests conducted on the treatment groups' pretest and immediate posttest scores in the production section (see Table 4) suggest that although the control group showed some minor changes from pretest to the two follow-up posttests, the changes were not high enough to make significant differences in the group performance over time. However, the three treatment groups improved significantly from pretest to posttest 1: t (35)=-9.13, $p<.05$ for the PI, t (40)= -7.86, $p<.05$ for the OI, and t (40)=-8.22 , $p<.05$ for the EI group. The means also showed that the OI group made higher gains on the first posttest than the PI and EI. In addition, although the

treatment groups showed a decline on posttest 2, their performance was still significantly higher than that of the pretest.

Table 4
Descriptive statistics and paired-sample t-tests for production tasks

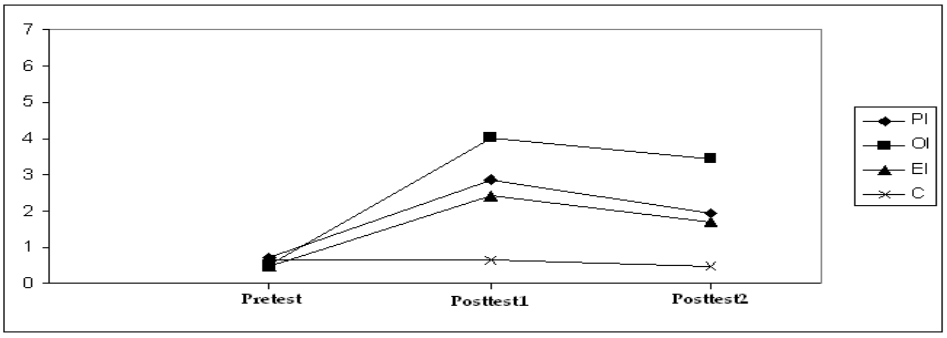
		Pretest		Post test1		Post test2		Pre-Post1	Post1-Post2	Pre-Post2
Group	n	M	SD	M	SD	M	SD	t/p	t/p	t/p
PI	35	0,71	0,97	2,87	1,62	1,94	1,31	-9,13 0,000	4,002 0,000	-5,332 0,000
OI	40	0,52	0,83	4,03	2,65	3,46	1,67	-7,86 0,000	1,31 0,197	-9,92 0,000
EI	40	0,48	0,70	2,42	1,51	1,7	1,32	-8,22 0,000	2,87 0,000	-6,01 0,000
C	36	0,63	0,70	0,66	0,72	0,47	0,64	-2,06 0,838	1,50 0,142	1,43 0,160

Then in order to determine whether there were any advantages of one type of instruction over the other in helping the learners produce English causative sentences, their scores on the first posttest in the production section were submitted into one-way ANOVA (significance level was set at .05). The results (see Table 3) revealed a significant difference in how accurately the groups produced the causative sentences: F (3, 147)=22.826, p= 0.000; The computed effect sizes (0.31) ensured the meaningfulness of the magnitude of the differences. Accordingly, the answer to the third research question is *Yes*, i.e., there were significant differences among the four groups regarding their performance on the immediate production posttest. The results of Post hoc Tukey HSD tests were, however, different from those of the interpretation section: the three types of instruction were better than no instruction (C); the OI group outperformed the PI class; no significant difference was found in how the PI group and the EI group produced the causative sentences. The performance of the groups on the production tasks of immediate posttest are summarized as follows: OI> PI=EI> C.

As for the last research question, the results of within-group comparisons (see Table 4) showed that the immediate improvement in producing causative sentences on posttest 1 was sustained equally well over a one-month interval on posttest 2 only by the OI group (t (39)=1.31, p<.05), while the improvement in the production ability of PI (t (34)= 4.002, p<.05) and EI (t (39)= 2.87, p<.05) declined

significantly between the first and the second posttests. However, no significant differences were found between the treatment groups' pretest and the delayed posttest scores. In other words, none of the treatment groups returned to the same level of performance it showed on the pretest, which suggests the durability of the effects of the three instructional options on the production ability of learners. In order to investigate the differences in the effects of instruction on the groups' production ability after a one-month interval, the group scores of the delayed posttest in the production section were submitted into one-way ANOVA too. The results (Table 3) revealed a significant difference in how the four groups produced the causative sentences: F (3, 147) =33.695, p= 0.000; the eta squared figures (0.40) also suggest considerable effect sizes for the result. Accordingly, the answer to research question 4 is *Yes*, i.e., there were significant differences among the groups as for their performance on the delayed posttest measuring the production of the target structure, and the positive effects for instruction on the production tests were retained equally well over a month by the instructional groups. The results of Post hoc Tukey HSD tests were similar to those of the immediate production posttest: OI > PI=EI> C. The visual representations of the gains and longer-term effects for the production tasks across the pretest and the follow-up tests are shown in Figure 2.

Figure 2
Production group means



Discussion

The results of the study can be summarized as follows: first, it appears that all the treatment groups made gains from the pretest to posttest 1 in both sections of interpretation and production, but the gains were not equal. While in the interpretation section PI and OI gains were equally greater than EI gains, on the

production tasks of posttest 1 it was OI that outperformed both PI and EI. However, it seems noteworthy that the gains made by the treatment groups of OI and PI on both sections of the test, though statistically different from and higher than EI and C groups, were not that dramatic. A simple glance at the means of the groups over three administrations (Tables 2 and 4) shows that neither OI nor PI approach could improve the learners' ability of interpretation and production as highly as it was expected. This might be related to the low proficiency of the subjects evident from their scores on the proficiency test of PET, or even to the novelty of approaches to the subjects in spite of getting training for two weeks. The minimal changes might also imply that, as Ellis (2006) emphasizes, use should be made of both input-based and output-based instructional approaches to achieve more efficient results.

Second, the learning gains on the immediate posttest of interpretation were maintained on the delayed posttest for all the treatment groups. However, the improved ability to produce the causatives reflected on the immediate posttest remained constant between the first and the second posttest only in the OI group. Third, although the instructional effects seemed to atrophy between the immediate and the delayed posttests for the treatment groups, none of the groups returned to the same level of performance it showed on the pretest in either interpretation or production section.

The finding that none of the treatment groups returned to the same level of performance observed before instruction implies the beneficial effects of instruction, though minimal, regardless of the type of instruction. However, the lower gains made by the EI group in comparison to the other two treatment groups with input-based or output-based activities might imply the necessity of engaging learners in doing some sort of follow-up activities after introducing a grammar form. That is, providing just explicit information about a grammar form does not help learners achieve an acceptable level of accuracy especially in production tasks. This conclusion might have an implication for those L2 teachers who do not spend enough time on involving their students in doing follow-up practice.

The study also led to findings about the relative effectiveness of input-based and output-based activities on the interpretation and production abilities of learners, some of which concur while others contradict VanPatten's (1996) claims about the superiority of PI over OI. The finding which supported VanPatten's

claims in PI model was that PI group improved significantly not only in interpretation but also in production of the causative structure. While PI improvement in the interpretation test was expected because of the role that structured input activities might have played in helping the subjects to readjust their default processing strategies, their significant improvement on production test was interesting, given that they were never allowed to produce the causative form during the instruction phase. The finding is in line with VanPatten's (2002a) argument that PI can help L2 learners modify their underlying system because it can maximize learners' intake by pushing them away from incorrect processing input and by building a form-meaning connection, which results in improving the accuracy of both comprehension and production of grammar features.

However, PI instructional benefits should be expounded with reservation because, as the within-group comparisons showed, the learning gains made by the PI learners on the interpretation tasks were also observed in the OI and EI groups despite the fact that they were not involved in processing activities during the treatment phase. Besides, in this study, the OI option could affect the learners' grammar knowledge equally well. Between-group comparisons (similar to Allen, 2000; Erlam, 2003; Nagata, 1998a, 1998b, and Toth, 2006) also showed no advantages for PI over OI in improving the subjects' interpretation ability. This is in contradiction with Lee and VanPatten (2003) who viewed the role of output practice only in the way of helping with fluency and accuracy in production, not as responsible for getting the grammar into the learners' head. The finding that the OI group outperformed PI class in producing the causative sentences may foster the conclusion that the claimed PI benefits in production might not be generalized to all the grammatical structures or to all EFL contexts.

The lack of superiority of PI over OI in this study was more obvious in the comparison made between the learning gains of the two groups and the instructional benefits of EI class on the immediate and the delayed posttests of production (OI>PI=EI). This might imply that whether the learners are provided with PI activities or just with a brief grammar explanation with no follow-up practice, the result would be the same as far as the learners' production ability is concerned. This finding might lead to the wrong conclusion that spending time and energy on preparing PI materials and applying them in L2 classes, especially the large ones, might be a waste of time. However, the minimal gains of EI as compared with those of OI (both in interpretation and production tasks) and PI (in

interpretation tasks) implies that after presenting a new form, learners should be engaged in doing follow-up activities, especially the output-based activities, to get more efficient results.

While in the studies of VanPatten and Cadierno (1993), Cheng (1995, in VanPatten, 2002a), Tanaka (1996, in Ellis, 1999) the gains of PI and OI were maintained over the month-long post-testing phase both in interpretation and production tests, in this study PI and OI (as well as EI) maintained their learning gains from posttest1 to posttest 2 only on the interpretation section. However, on the production tasks, it was just OI that maintained its gains in producing causative sentences over a month. The superiority of OI over PI in improving the learners' production might be related to the functions that Swain (2000) lists for output. That is, the feedback the learners of this study received after doing the output activities might have helped them to compare their own production with the correct model provided by the teacher, and through this comparison they might have noticed the gap in their linguistic knowledge. Since PI learners were not producing anything during the instructional phase, providing feedback did not help them much because they did not have the chance to notice the gap in their developing system. The learners' output in the OI class might have also acted like 'auto input' (Ellis, 2003) for them and hence helped the acquisition process.

It is necessary to remind that the different results obtained in this study should not be considered as evidence for problems with the IP theory or with the relevant work in the literature. Given that the study was not an exact replication of the original VanPatten and Cadierno' study (1993), any differences in the findings of this study might be related to the materials used during the instruction, the assessment measures, the procedures, the kind of sampling, and many other unknown factors. Therefore, caution needs to be exercised when interpreting these results.

Limitations of the Study

The present study contained some limitations that one should bear in mind in order to avoid undue generalizations: The first weakness was related to the non-random nature of subject selection; since the participants of the study were assigned to classes by the university registration office, the random selection was out of the researchers' control. Besides, even though effect size calculations reinforced the significance of the principal findings, the lack of control over variables such as

individual differences might mean that the results of the present study should not be generalized with confidence beyond the sampling frame of Islamic Azad University of Naragh. The third limitation was that because of the administration constraints, including time restrictions, only one grammatical structure was examined; other research could seek further evidence for the interaction of structure with instruction by studying other structures. The next obvious limitation was that the content and form that learners were required to produce were controlled (not free) in output instructional tasks as well as in production tests both to ensure the participants' use of target forms in their outputs as well as helping to score the written production more objectively. The advantages might be a limitation of the study in that the participants were not provided with opportunities for more natural and communicative use of the target forms. Finally, the learners in this study were adult university students with a rather low level of proficiency; thus it remains to see whether the gained results would hold the same for learners representing other ages or proficiency levels.

Conclusion

Based on the results discussed above, it can be concluded that the type of follow-up activities plays an important role in improving learners' interpretation and production abilities and in the durability of their progress. It seems that the often-claimed instructional benefits of PI in changing L2 learners' underlying system should be accepted with reservation because in this study the OI group improved equally well in interpretation tests and outperformed the PI in production tasks. At the theoretical level, the results of the study may contribute directly to the ongoing discussion on the roles of input and output in SLA. At the pedagogical level, it contributes to the pedagogical research in application of output-based and input-based approaches to grammar instruction by providing empirical evidence for a significant role for output-based instruction. The findings might also be used by material developers and syllabus designers who may find it necessary or at least helpful to incorporate some of the suggested activities in the study into the materials or syllabi aimed at developing English grammar knowledge.

Acknowledgements

We are grateful to all our university colleagues and high school teachers who read and commented on earlier versions of the instructional packs and the assessment measures. Their great contribution to developing and validating the instruments is deeply acknowledged. Also, we owe enormous thanks to Mrs. Jane Modarresi and

Mr. Tim Mackay, the two native speakers of English, for their insightful comments on the assessment measures. Special thanks are due to all the students at the University of Naragh who took part either in the pilot or the main stage of the study.

Received 23 July, 2009
Accepted 2 August, 2009

References

Allen, L. Q. (2000). Form-meaning connections and the French causative: An experiment in processing instruction, *Studies in Second Language Acquisition*, 22, 69-84.

Benati, A. (2001). A comparative study of the effects of processing instruction and output-based instruction on the acquisition of the Italian future tense, *Language Teaching Research*, 5(2), 95-127.

Benati, A. (2004). The effects of processing instruction and its components on the acquisition of gender agreement in Italian, *Language Awareness*, 13(2), 67 – 80.

Benati, A. (2005). The effects of processing instruction, traditional instruction, and meaning-output instruction on the acquisition of the English past simple tense, *Language Teaching Research*, 9(1), 67-93.

Cadierno, T. (1995). Formal instruction from a processing perspective: An investigation into the Spanish past tense*, The Modern Language Journal,* 79*,* 179-193.

Celik-Yazici, I. (2007). A study of the effects of processing instruction on the development of English WH-questions by Turkish EFL learners, *Interlinguistica,* 17*,* 250-260.

Collentine, J. (1998). Processing instruction and the subjunctive, *Hispania*, 81, 576-587.

DeKeyser, R. & Sokalski, K. (1996). The differential role of comprehension and production practice, *Language Learning*, 46, 613- 642.

Dornyei, Z. (2007). *Research methods in applied linguistics.* Oxford: OUP.

Ellis, R. (1999). Input-based approaches to teaching grammar. A review of classroom-oriented research, *Annual Review of Applied Linguistics*, 19, 64-80.

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

Ellis, R. (2006). Current issues in the teaching of grammar: An SLA perspective, *TESOL Quarterly*, 40(1), 83-107.

Erlam, R. (2003). Evaluating the relative effectiveness of structured-input and output-based instruction in foreign language learning: Results from an experimental study, *Studies in Second Language Acquisition*, 25, 559-582.

Gass, S. M. & Selinker, L. (2008). *Second Language acquisition: An introductory Course (*3rd ed.). NY: Routledge.

Lee, J. & VanPatten, B. (2003). *Making communicative language teaching happen*. New York: McGraw-Hill.

Lotfi, A. R. (2008).Causative constructions in Modern Persian, *California Linguistic Notes,* XXXIII (2), 1-31.

Marsden, E. (2006). Exploring input processing in the classrooms: An experimental comparison of processing instruction and enriched input, *Language Learning,* 56(3), 507-566.

Nagata, N. (1998a). Input vs. output practice in educational software for second language acquisition, *Language Learning & Technology,* 1(2), 23-40.

Nagata, N. (1998b). The relative effectiveness of production and comprehension practice in second language acquisition, *Computer Assisted Language learning*, 11(2), 153-177.

Qin, J. (2008). The effect of processing instruction and dictogloss tasks on acquisition of the English passive voice, *Language Teaching Research,* 12 (1). 61-82.

Radwan, A. A. (2009). Input processing instruction and traditional output practice instruction: Effects on the acquisition of Arabic morphology, *The Asian EFL Journal Quarterly*, 11(3), 267-298.

Swain, M. (1998). Focus on form through conscious reflection. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp.64-81). Cambridge: CUP.

Swain, M. (2000). The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. In J. Lantolf (Eds.), *Sociocultural theory and second language learning* (pp. 97-114). Oxford: OUP.

Toth, P. D. (2006). Processing instruction and a role for output in second language acquisition, *Language learning*, 56 (2), 319-385.

VanPatten, B. (1996). *Input processing and grammar instruction: Theory and research.* Norwood, NJ: Ablex.

VanPatten, B. (2002a). Processing Instruction: An update, *Language Learning,* 52*,* 755-803.

VanPatten, B. (2002b). Processing instruction, prior awareness, and the nature of second language acquisition: A (partial) response to Batstone, *Language Awareness,* 11(4), 240-258.

VanPatten, B. & Cadierno, T. (1993). Explicit instruction and input processing, *Studies in second language acquisition*, 15, 225-243.

VanPatten, B. & Oikkenon, S. (1996). Explanation vs. structured input in processing instruction, *Studies in Second Language Acquisition,* 18, 495–510.

VanPatten, B., Inclezan, D., Salazar, H. & Farley, P. A. (2009). Processing Instruction and Dictogloss: A Study on Object Pronouns and Word Order in Spanish, *Foreign Language Annals,* 42(3), 557-575.

منبع فارسی:

باطنی ، محمدرضا. (۱۳۸٤).توصیف ساختمان دستوری زبان فارسی بربنیاد یک نظریه عمومی . انتشارات امیرکبیر .

Appendix A

*The Script of Explicit Grammar Information about the Causative Verb 'Have'
(Translations are provided for the reader only)

In the sentence *Mary polished the shoes*, 'Mary' is the person who did the act of polishing. Now if you want to express the idea that Mary required, caused, or wanted another person to polish the shoes for her, you can use the English causative structure. The causative meaning can be conveyed using one of the causative verbs such as the verb 'have'; the structure expresses the idea that a person wants something to be done for him. A causative sentence can be in two forms of active and passive depending on whether the real doer of the action is known or not.

Active form: subject + Have + Person + Simple Form of Verb + Object
E.g. Mary had her brother polish the shoes.

Passive form: Subject + Have + object + Past Participle of Verb
E.g. Mary had the shoes polished.

In the passive construction above, Mina arranged for something to be done by a third person, but that person is not known, while in the active construction the person who did the act of polishing for her is known.

- The following information was provided **ONLY** for PI group.

1. NOTE that the verb 'Have' can be used to convey a meaning different from the usual ones such as 'to posses or to eat'. When 'have' is used in the causative meaning, it expresses the idea of someone causing something to take place. Thus, YOU MUST pay careful attention to the form of a sentence containing 'have' to see if it implies the causative meaning or not.
2. DO NOT take the first noun in the causative sentence as the doer of the act; the first noun refers to a person who wants another person to do the act for him. The real doer of the act is known in active construction; it is located immediately after the causative verb. The doer of the act is not known in passive construction so you can not find it in the sentence. The noun following the verb 'have' is the object of the act not the subject.

*Samples of Referential Activities for the Causative Verb 'Have'

Read the following sentences which contain the verb 'have'. Circle around 'Yes' if the verb is used as the causative verb, and mark 'No' if it implies meaning other than causative.

1. Did you have a lot of rain in your city last winter? **Yes** **No**
2. The teacher had Mary type her projects. **Yes** **No**
3. They have not painted their house for 7 years **Yes** **No**

(5 more activity items of the same type)

You are going to hear some sentences in English which contain the causative 'have'. Listen carefully and circle around 'a' if the causative sentences are active, and choose if they are in passive form.

Teacher's script:

1. The police had the man tell them the name of his friends.
   a. active b. passive
2. She is having her hair dyed.
   a. active b. passive

(4 more activity items of the same type)

Look at each picture carefully. Then select the sentence that goes with it.

1. a. Mrs. Pearson and John are cleaning the yard.
   b. Mrs. Pearson is having John clean the yard.

(5 more pictorial items for each causative verb in active and passive modes)

* Samples of Affective Activities for the Causative Verb 'have'

Read the following ideas given by some people. Decide whether you agree or disagree with the suggested ideas.

1. Parents should have their children sleep early during the school time.
   a. agree b. disagree
2. Women have the power to have their husbands stop smoking.
   a. agree b. disagree

(3 more activity items of the same type)

You are going to hear some sentences. Listen carefully, and circle around 'True' if they are true about you or your family too, and mark 'False' if they are not.

Teacher's script:

1. His parents have him buy the bread everyday. a. True b. False
2. I always have my clothes ironed. a. True b. False

(3 more activity items of the same type)

Appendix B

* Samples of Meaning-based Output Activities for the Causative Verb 'have'

Use the causative verb 'have' and answer questions both in active and passive modes.

1. Is she going to make her new dress herself?
   No, she…….
   No, she ……

(9 more activity items of the same type)

Read the information given in the chart below. Match the relevant information. Then use the verb 'have' and make 10 complete, meaningful, and grammatically correct sentences. Add your words if necessary. The first one is done for you.

Causer	Doer	Act
Parents	his patients	do the chores
Teachers	children	fix the car
A doctor	a mechanic	type their projects
You	students	take their medicine on time.
……..	……	……

E.g. 1. Parents should have their children do chores.

Look at each picture. Use the causatives verb 'have' and the cue word/s given for each picture. Then write a sentence that expresses the meaning implied in the picture.

1. Jane could not repair her shoes, so she …….

(repair)

(9 more activity items of the same type)

Appendix C

*Samples of Interpretation Test Items

PART 1: Read the sentences. Select the drawing which shows the meaning conveyed in each sentence. Mark the correct letter 'a' or 'b' in you answer sheet.

1. Mrs. Brown usually has a cake made for her son's birthday.

a.

b.

PART 2: Read each sentence. Choose the correct answer. Mark the correct letter 'a, b, or c' in your answer sheet.

1. The new teacher gets the students to work hard. The sentence says……
   a. the students work hard
   b. the new teacher works hard
   c. the students have a hardworking teacher
2. Mina has had her sister's dress ironed. Who has ironed the dress?
   a. Mina's sister
   b. Mina
   c. someone else

PART 3: Read each sentence carefully. Determine the message expressed by each sentence by selecting the correct translation. Mark the correct letter 'a, b, or c' on your answer sheet.

I'm going to get a new picture drawn for my room.

a. تصميم دارم براي اتاقم سفارش نقاشي يك تابلو جديد را بدهم.
b. تصميم دارم براي اتاقم يك تابلو جديد نقاشي خريداري كنم.
c. تصميم دارم براي اتاقم يك تابلو جديد نقاشي كنم.

*Sample of Production Test Items

PART 1: Look at each picture. Use the cue word/s given for each picture and complete the sentence in a way to describe it in the best way. Add your own words if required. Make any necessary changes. Write only the missing parts in your answer sheet.

The basket was very heavy, so the old lady …….it.

(carry/got)

PART 2: Here are some sentences. For each item, complete the sentence B so that it means or implies the same meaning as the first sentence/s. Write only the missing parts in the answer sheet.

The teacher wanted him to type the project, but he asked me to do it for him.

A: **Did he type the project himself?**

B: **No, he got ……**