Iranian Journal of Applied Linguistics (*IJAL*)

Vol. 20, No. 1, March 2017, 81-111

Measuring the Effectiveness of Explicit and Implicit Instruction through Explicit and Implicit Measures

Sasan Baleghizadeh*, *Shahid Beheshti University,Tehran, Iran*
Ali Derakhshesh, *Shahid Beheshti University,Tehran, Iran*

Abstract

Many studies have examined the effect of different approaches to teaching grammar including explicit and implicit instruction. However, research in this area is limited in a number of respects. One such limitation pertains to the issue of construct validity of the measures, i.e. the knowledge developed through implicit instruction has been measured through instruments which favor the recipients of explicit instruction. The present study expands on the previous studies by exploring the effectiveness of explicit and implicit instructions through administering a timed GJT and an untimed GJT. Data were collected from three different groups: (a) the explicit group was presented with rules of verb complementation, (b) the implicit group received visually enhanced texts, and (c) the control group received no instruction of the target feature. The results of the mixed between-within subjects ANOVA test revealed that the learners in explicit group outperformed the learners in implicit and control groups in post-tests and delayed post-tests even in the case of the implicit measure. Furthermore, the results indicated the more durable effects of explicit instruction compared with implicit instruction. Taken together, the study provided evidence for the efficiency of explicit teaching compared to implicit teaching in the context of L2 development.

Keyword: Explicit instruction; Implicit instruction; Explicit testing; Implicit testing

Article Information:

Received: 21 August 2016 **Revised: 12 February 2017** **Accepted: 21February 2017**

Corresponding author: Department of English Language, Shahid Beheshti University, Tehran, Iran
Email address: sasanbaleghizadeh@yahoo.com

[DOI: 10.29252/ijal.20.1.81]

1. Introduction

A survey of different models of language competence reveals the importance of grammar component as it appears in all the models (e.g., Bachman, 1990; Canale & Swain, 1980; Celce-Murcia, Dörnyei, & Thurrell, 1995). As a result, much research has been conducted on how to teach this component and more specifically, on one of its constituents, that is, syntax (henceforth referred to as grammar). While there are many strategies and methods that are used currently in teaching grammar, they could be categorized under two super-ordinate terms: explicit and implicit approaches. The current cumulative research findings suggest the superiority of explicit instruction (EI) over implicit instruction (II) (Alanen, 1995; De Graaff, 1997; De la Fuente, 2009; Hernández, 2008; Rosa & Leow, 2004; Yoshimi, 2001). However, in their meta-analysis, Norris and Ortega (2000) claimed that the value of EI over II should be interpreted with caution and tempered by the methodological factors. Foremost among these factors is the choice of outcome measure. They rightly commented that most of these studies failed to include a measure of implicit knowledge and this may explain why the groups which received EI outperformed those which received II. Admitting the lack of valid measures of second language (L2) implicit and explicit knowledge, attempts most predominantly led by Ellis (2005) and his colleagues (1998, 2009) were made to develop and validate explicit and implicit measures. Since then, these instruments have been employed in the studies of EI and II (e.g., Akakura, 2012; Hernandez, 2011; Philp, 2009). However, the bulk of research in this area is small yet and for that reason alone, it would seem to be worth continuing to ask questions about explicit/implicit grammar instruction. Employing appropriate explicit and implicit measures, the present study is after (a) finding how successful

[DOI: 10.29252/ijal.20.1.81]

EI is compared to II in promoting knowledge on verb complementation, and (b) showing whether the effect of EI or II is more durable.

2. Review of the Related Literature

2.1. Explicit vs. implicit instruction

Based on a number of criteria, different taxonomies of instructional activities have been proposed (e.g., Doughty & Williams 1998; Ellis, 2001; Williams 2005). However, since the psycholinguistic and practical validity of these taxonomies have not been fully demonstrated, activities have been arranged along the continuum of 'explicitness' (De Graaf & Housen, 2009). Instruction is implicit if it '… is directed at enabling learners to infer rules without awareness' (Ellis, 2008, p. 438). Therefore, II is portrayed by the absence of rule presentation or instruction in the hope that learners would process the input to find out if the data could be described with a rule (Hulstijn, 2005). This type of instruction suggests learners be exposed to exemplars in a meaning-focused and comprehensible context in the hope of inferring patterns. EI, on the other hand, '… insists upon the value of deliberate study of a grammar rule' (Scott, 1990, p. 779). DeKeyser (1995) stated EI occurs when 'some sort of rule is being thought about during the learning process' (p. 380). This means learners are aware of what is being taught to them and are encouraged to develop metalinguistic knowledge.

There have been a number of intervention studies on the effectiveness of EI and II (Loewen, Erlam, & Ellis, 2009; Radwan, 2005; Rosa & Leow, 2004; Varnosfadrani & Basturkmen, 2009; Zhou, 2010). Predominantly, these studies have shown the relative effects of EI in contrast with II. In a study which comprised four learning conditions, Radwan (2005) explored the effects of degrees of explicitness on the acquisition of English dative alternation. The four study groups were a textual enhancement condition, a

[DOI: 10.29252/ijal.20.1.81]

rule-oriented condition, a content-oriented condition, and a control group. As test instruments, the researcher used a grammaticality judgment task, a preference task, and a controlled writing task. The results of the study showed that rule-oriented condition outperformed other groups in immediate and delayed post-tests. This, he contends, gives support to the fact that less explicit forms of instruction like input enhancement are insufficient, at least in the case of dative alternation, to change learners' interlanguage. In another study, Alanen (1995) examined the effect of EI and textual enhancement on the acquisition of locative suffixes and patterns of consonant change in Finnish. The participants were assigned to one of the four conditions: control, textual enhancement, explicit rule presentation, and explicit rule presentation plus textual enhancement. She found that although treatments affected the learning process, as demonstrated by learners' think-aloud and performance on test instruments, the groups who were exposed to explicit rule instruction performed better than the other two groups. Alanen concluded that perceptual salience of input spurred learners to process it cognitively; however, the method may not have been perceptually salient enough to focus their attention on the instructed form.

The number of studies which have investigated the longitudinal effect of EI and II is limited (e.g., Klapper & Rees, 2003; Morgan-Short, Sanz, Steinhauer, & Ullman, 2010, 2012; Spada & Lightbown, 1993; Tode, 2007; White, 1991). In one longitudinal study, Tode (2007) examined the effect of EI and II on beginners with respect to the learning of the copula *be*. To this end, 89 Japanese participants were divided into three groups. EI consisted of an explanation phase, an identification phase, and a writing phase. II was given in the form of presenting English exemplars with their translation to learners and then asking them to memorize the sentences. The control group was given instruction on the auxiliary *can*. Five post-tests were given over a period of six months in the form of sentence completion items. Data

[DOI: 10.29252/ijal.20.1.81]

analysis revealed that learners who received EI made short term gains, while learners in II did not. In addition, the author did not find a significant difference between the implicit group and the control group. Moreover, the results of post-test analyses showed that the effect of EI was not durable. Tode suggested in order to obtain the durable effect of EI, supplementary material which creates more exposure to the target feature, induces the noticing of the target structure, and gives learners the opportunity to contrast target forms with other forms, should be supplied to learners.

2.2. Explicit vs. implicit testing

Douglas (1998) argued that language testing and second language acquisition (SLA) studies are related to each other as practitioners in these two fields use elicitation devices to make inferences about learners' state of interlanguage. Though it is a prevalent belief that in order to make justified inferences we need to make sure that there is a link between the elicitation methods and the construct under investigation, that is, tests should have construct validity, researchers have still failed to consider the (construct) validity of the testing instruments in some areas (Douglas, 2001). For example, in case of explicit and implicit teaching, lack of attendance to construct validity was apparent until recently in the absence of appropriate implicit assessment instruments capable of measuring of implicit knowledge.

While it is true that finding pure and sensitive measures tapping exclusively into implicit and explicit sources of knowledge is hard (DeKeyser, 2003, 2009), Ellis (2005) conducted a study, which originated in his earlier study with Han (1998), to design and validate a battery of English language tests which would provide separate measures of these two types of knowledge. These tests included: (a) Oral narrative test; (b) Imitation test;

[DOI: 10.29252/ijal.20.1.81]

(c) Timed GJT; (d) Untimed GJT; and (e) Metalinguistic knowledge test. Results of factor analysis showed that the first three tests loaded heavily on the implicit knowledge factor, whereas the last two tests loaded heavily on the explicit knowledge factor. Following this line of research, other studies have further confirmed these results even with other languages (Bowles, 2011; Ellis & Loewen, 2007; Ellis, Loewen, Edler, Erlam, Philp, & Reinders, 2009, Godfroid, Loewen, Jung, Park, Gass, & Ellis, 2015; Zhang, 2015). These studies demonstrated that the kind of knowledge GJTs prompt learners to tap into depends on whether the test is timed or untimed. Thus, time pressure plays a pivotal role with reference to implicit and explicit knowledge.

In addition to examining the role of time pressure, Gutiérrez (2013) scrutinized the role of stimulus type (i.e., grammatical vs. ungrammatical). The results demonstrated that the grammatical items of both timed and untimed GJTs loaded on the construct of implicit language knowledge, and the ungrammatical items of both GJTs and metalinguistic knowledge test loaded on explicit language knowledge. Although ungrammatical items in Gutiérrez (2013), regardless of time pressure, loaded on the construct of explicit knowledge, the correlations between ungrammatical items of untimed GJT and items of metalinguistic knowledge test – identified as a measure of explicit knowledge (Suzuki & De Keyser, 2015; Vafaee, Suzuki, & Kachisnke, 2017) – were higher than the correlations between ungrammatical items of timed GJT and those of metalinguistic knowledge test, indicating ungrammatical items of untimed GJT are a purer measure of explicit knowledge. Formerly, Ellis (2005) had used only the ungrammatical scores of untimed GJT in his analysis. The results showed that ungrammatical scores of untimed GJT loaded heavily on the explicit construct. Consequently, Ellis and Loewen (2007), in a follow-up study, used only the ungrammatical scores of the untimed GJT. Their results

[DOI: 10.29252/ijal.20.1.81]

confirmed the findings in Ellis (2005) regarding the loading of the ungrammatical scores of untimed GJT. Then, it can be concluded that ungrammatical scores of the untimed GJT are better measures of explicit knowledge. Overall, these findings suggest that along with time pressure, stimulus type also influences the type of knowledge that learners draw on.

Given that Gutiérrez's (2013) finding regarding the ungrammatical items of timed GJTs has not been widely supported in other studies yet, the researchers decided to follow the general trend considering (a) timed GJT as a measure of implicit knowledge, (b) untimed GJT as a measure of explicit knowledge, and (c) ungrammatical items of untimed GJT as a purer measure of explicit knowledge.

3. The Study

Although SLA researchers have responded to Norris and Ortega's call for more measures of unanalyzed knowledge, more empirical studies are needed to explore whether the use of these enhanced measures leads to different results of the effectiveness of EI and II. Also, as Spada and Tomita (2010) asserted, there is a need for studies which include delayed post-tests in their design. The present study attempts to add insights to the literature by investigating the effect of two types of input-based instruction on the acquisition of verb complementation measured through a measure of implicit knowledge and a measure of explicit knowledge. The research addresses the following research question:

Which type of instruction – explicit or implicit – is more effective over time when the effects of instruction are measured in terms of

1) a measure of explicit language knowledge, and its purer sub-set, that is, ungrammatical items?

2) a measure of implicit language knowledge?

[DOI: 10.29252/ijal.20.1.81]

4. Method

4.1. Participants

The present study was conducted in a university context. The participants were 65 freshmen (18 to 20 years old), both male and female, majoring in psychology. The first language of all participants was Persian and they were completing a 4-credit general English course at the time of the study. They came from three intact classes: one group was given EI (N = 19), another group received II (N = 25), and the third group was not provided with instructional treatment (control group, N = 21). In Iran, students study English at secondary school for four years. A structural syllabus serves as the basis for the design of these general courses and the learners receive focus-on-forms instruction through a combination of the principles of Audiolingual and Grammar-translation methods. Therefore, students mostly develop knowledge of language usage than use. For admission to university, students have to take a National University Entrance Exam based on the syllabi they have studied at secondary school. The English section of the exam comprises twenty-five discrete-point multiple-choice items. The background questionnaire showed that on average the participants had answered 25% of the items correctly in the English section of the exam. Also, the main course-book assigned by the university English department for these students was a reading book at lower-intermediate level. Therefore, the researchers were almost sure that the participants of the study were fairly homogeneous in terms of language proficiency.

[DOI: 10.29252/ijal.20.1.81]

4.2. Instructional material

The instructional materials included six texts accompanied by some true/false items. The texts, seeded with fourteen verbs and their respective complements, were developed for the sake of the study. Verb complementation was selected because (a) its complexity suited the proficiency level of the participants, (b) few studies have considered it as the focus of instruction, and (c) it was not part of the participants' earlier syllabi in secondary school. The chosen verbs were those which permit either infinitive-type or gerund-type complement.

In order to make the job of developing texts more practical, the verbs were randomly divided into two groups and each set was used in a different text. This means two texts for each session and since there were three treatment sessions, a total of six texts were developed. Therefore, in each session, the participants encountered seven verb complements in one of the texts and the other seven ones in the second text. The texts were developed by Ph.D. students of English Literature and Teaching English as a Foreign Language. They were told to use simple vocabulary and structures, and to use each verb once in the text. The written texts were further revised by the researchers. The texts were 260 words long on average and using the Fry Graph readability formula, the difficulty levels of texts were calculated to be 6 which means readable by most adults.

4.3. Measures of instructional outcome

The participants completed a timed grammaticality judgment test (GJT) and an untimed GJT test during three test sessions. The GJTs in this study were designed based on the principles described in Ellis (2005). Timed GJTs require participants to indicate in a certain period of time if each sentence is

[DOI: 10.29252/ijal.20.1.81]

grammatically correct or incorrect. Such tests are predicted to encourage learners to use their intuition and since learners are pressed for time, they have little opportunity to use their meta-linguistic knowledge. Untimed GJTs follow a procedure like that of timed GJTs with the exception that participants can take up as much time as they need. Such tests encourage a high degree of awareness and provide enough time for examinees to access their meta-linguistic knowledge. The design features of these two tests are shown in the following table.

Table 1

Design features of timed GJT and untimed GJT (adopted and adapted from Ellis, 2009)

Criterion	**Timed GJT**	**Untimed GJT**
Degree of awareness	Feel	Rule
Time available	Pressured	Unpressured
Focus of attention	Form	Form
Utility of knowledge of meta-language	No	Yes

The timed GJT in this study was a computer-delivered test which consisted of 28 sentences (mean length = 10.48 words, *SD* = 1.75). Of these, seven items contained verb complementation in a grammatically correct context and seven items contained verb complementation in a grammatically incorrect context. The remaining sentences targeted other grammatical structures, evenly distributed between grammatical and ungrammatical. Sentences were randomly arranged to create three versions of the test for pre-test, post-test and delayed post-test. In order to determine the time limit of the test items, each sentence should have been judged by native speakers. However, since the authors had no access to native speakers in the EFL

[DOI: 10.29252/ijal.20.1.81]

context, the items were trialed on ten Ph.D. students who had been constantly exposed to English for at least fifteen years. Regarding the easy nature of the test items, it was supposed that the reaction time of the Ph.D. candidates would be similar to that of native speakers. Following Loewen (2009), after recording the reaction time of Ph.D. student for each item, the median response time was calculated and then an additional 20% was added to account for the slower processing speed of the learners. Therefore, the time allowed for judging each item ranged from 3.6 to 7.5 seconds.

The untimed GJT contained the same set of sentences in the timed GJT. However, it was a pen-and-paper test. The participants were instructed to indicate their answers by checking the relevant box. Again, the items of the same test were scrambled to create three versions for the three test sessions.

As for scoring, both tests were scored dichotomously. If an item were judged correctly one point would be awarded and if it were judged incorrectly no point would be awarded. Besides, items which were left blank were considered as incorrect.

4.4. Data collection procedure

One week prior to the intervention, the participants were given a background questionnaire and then the pre-tests, that is, the timed GJT and the untimed GJT. The order of test presentation was the timed GJT followed by the untimed GJT in order to prevent instrument effect. The participants were informed that timed GJT items would appear on the screen with different lengths of time, within which they should indicate their answers by pressing a key. In case of untimed GJT, they were told to take their time answering the items. Following the pre-test session, instruction was provided during the regular class time by the second researcher, over the

[DOI: 10.29252/ijal.20.1.81]

next three weeks each lasting 25 minutes according to the following procedure.

Explicit Group. The participants were allowed to read the two texts on their own for 10 minutes. Then, they received EI in form of a hand-out describing the pattern governing the use of the learning targets accompanied by the instructor's explanation. Next, the instructor worked with the participants on the vocabulary and structures, while directing the learners' attention to the verbs and their respective complements.

Implicit Group. The participants in this group received the same texts with typographical enhancements made to the target feature. This was achieved by underlining and increasing the font size of the verbs and using bold typeface and larger font size for the verb complements. These techniques are purported to present the target feature implicitly through increasing its perceptual salience. The participants spent 10 minutes reading the texts silently, and then the teacher worked with students on their comprehension of the texts without giving any EI or indicating what to look out for.

Control Group. As for the control group, the participants received the same texts with no modifications or explanations. After the texts were distributed among the participants, they read them individually for 10 minutes. Next, the instructor helped them with difficult words and probable ambiguities for the sake of better understanding. Then, the subjects answered true/false items concerning the content of the reading passages.

As can be seen, instructions were input-based and neither of the groups had production practice. The day after the third treatment session, the timed GJT and the untimed GJT were given to the participants as the first post-

[DOI: 10.29252/ijal.20.1.81]

test. Four weeks later, the delayed post-test was administered to assess the long term effects of the instructions.

5. Results and Discussion

The test scores were analyzed with a mixed between-within subjects ANOVA with *test time* (pre-, post-, delayed post-tests) as within-participants factor and *instructional group* (control, explicit, implicit) as between-participants factor. In this section, the score analyses of untimed GJT, ungrammatical items of untimed GJT, and timed GJT are presented. In all three cases, homogeneity of variance was checked through Levene's Test, homogeneity of covariance was checked through Box's Test, and sphericity was checked through Mauchly's Test. In these tests, the assumption is violated if the significance values are smaller than .05, .001, and .05, respectively. The examination of Levene's Test table revealed that the assumption was violated in post- and delayed post-tests results of timed GJT and thus, the more stringent significance level of .01 was set for evaluating the results (Pallant, 2016). Also, Box's M test had *p*-values which were larger than .001, suggesting that the assumption of covariance homogeneity was met. Finally, Mauchly's Test revealed that the sphericity assumption was violated for untimed GJT data and therefore, the Greenhouse-Geisser correction for degrees of freedom was used (Tabachnick & Fidell, 2013).

Untimed GJT. The descriptive statistics on the untimed GJT scores (see Table 2) revealed a rise in accuracy for all groups from pre-test to post-test but a fall in delayed post-test.

[DOI: 10.29252/ijal.20.1.81]

Table 2

Descriptive statistics for untimed GJT scores

Group	Test	N	Minimum	Maximum	Mean	Std. Deviation
Control	Pre-test	21	3.00	10.00	6.214	1.536
	Post-test	21	2.00	11.00	7.219	1.886
	Delayed post-test	21	4.00	10.00	6.761	1.612
Explicit	Pre-test	19	4.00	8.00	5.588	1.335
	Post-test	19	5.00	13.00	8.716	2.676
	Delayed post-test	19	4.00	13.00	8.526	2.517
Implicit	Pre-test	25	3.00	10.00	5.510	1.607
	Post-test	25	3.00	12.00	6.850	1.921
	Delayed post-test	25	3.00	12.00	6.280	2.310

The analysis revealed that the main effects due to instructional group ($F(2,62) = 3.85$, $p < .05$, $\eta^2 = .10$) and test time ($F(1.80,112.02) = 22.21$, $p < .05$, $\eta^2 = .26$) were significant. As a result, we can claim that instruction created positive effect on the improvement of learners' knowledge of verb complementation. The results further showed that the three test times significantly differed in instructional groups. Additionally, the interaction between instructional group and test time was significant ($F(3.61,114.12) = 5.06$, $p < .05$, $\eta^2 = .13$).

The results of the Bonferroni test on the pre-test showed insignificant difference between the three groups. Furthermore, the analysis of the post-test results found significant differences between explicit group and control group ($p < .05$), explicit group and implicit group ($p < .05$), but not implicit group and control group ($p = 1.0$). On the delayed post-test, the pairwise comparisons again found significant differences between explicit group and

[DOI: 10.29252/ijal.20.1.81]

implicit group ($p < .05$), explicit group and control group ($p < .05$), but not implicit and control groups ($p = .6$).

Also, pairwise comparisons were made on the results of each group on the pre-, post- and delayed post-tests. The test scores of the control group did not differ significantly from pre-test to post- ($p = .1$) and delayed post-tests ($p = 1.0$). On the other hand, the analysis of the scores for the explicit group revealed a significant difference between pre-test to post- and delayed post-tests (for both $p < .05$). Finally, in the implicit group, the examinees' scores in pre-test differed significantly from post-test ($p < .05$) but not delayed post-test ($p = .5$).

Untimed GJT ungrammatical items. In accordance with the findings of previous research (Ellis, 2005; Ellis & Loewen, 2007), which found the ungrammatical subset of untimed GJT items to be a better measure of explicit knowledge, separate scores for these items were calculated. The descriptive results of these items are presented in Table 3.

Table 3

Descriptive statistics for the scores of untimed GJT ungrammatical items

Group	Test	N	Minimum	Maximum	Mean	Std. Deviation
Control	Pre-test	21	.00	6.00	2.786	1.410
	Post-test	21	.00	7.00	2.875	1.816
	Delayed post-test	21	.00	5.00	2.032	1.352
Explicit	Pre-test	19	.00	4.00	2.000	1.290
	Post-test	19	2.00	7.00	4.010	1.512
	Delayed post-test	19	1.00	7.00	3.474	1.916
Implicit	Pre-test	25	.00	5.00	1.910	1.290
	Post-test	25	.00	7.00	2.105	1.825
	Delayed post-test	25	.00	6.00	2.110	1.732

[DOI: 10.29252/ijal.20.1.81]

The results of mixed between-within subjects ANOVA printout revealed significant results for the main effect of test time, Wilks' Lambda = .15, $F(2,61) = 5.43$, $p < .05$, $\eta^2 = .15$, the main effect of instructional group, $F(2,62) = 4.63$, $p < .05$, $\eta^2 = .12$, and the interaction between the main factors, $F(4,122) = 4.70$, $p < .05$, $\eta^2 = .13$.

Results of the Bonferroni on the pre-, post- and delayed post-tests showed the following. No significant differences were found between the three groups on the pre-test. On the post-test, significant differences were shown between explicit and control groups ($p < .05$) and explicit and implicit groups ($p < .05$), but not implicit and control groups ($p = .4$). Similarly, the results of the pairwise comparisons on the delayed post-test showed there was a significant difference between explicit and control groups ($p < .05$) and also explicit and implicit groups ($p < .05$) but not implicit and control groups ($p = 1.0$).

Additionally, pairwise comparisons of each group on the three tests revealed that scores in control group did not change significantly from pre-test to post-test and delayed post-test ($p = 1.0$ and $p = .2$, respectively). In a similar vein, scores in implicit group did not change significantly ($p = 1.0$ for both post- and delayed post-tests). But, the results of explicit group in post-test and delayed post-test were significantly higher than pre-test. The difference between pre-test and post-test was significant at $p < .05$ and the difference between pre-test and delayed post-test was significant at $p < .05$, too.

Timed GJT. The descriptive statistics on the timed GJT scores showed the scores of pre-test increased over both post-tests for explicit group while the other two groups showed a rise in post-test and a decline in the delayed post-test.

[DOI: 10.29252/ijal.20.1.81]

Table 4

Descriptive statistics for timed GJT scores

Group	Test	N	Minimum	Maximum	Mean	Std. Deviation
Control	Pre-test	21	3.00	9.00	5.887	1.868
	Post-test	21	4.00	10.00	6.530	2.152
	Delayed post-test	21	4.00	11.00	6.467	1.660
Explicit	Pre-test	19	3.00	9.00	5.815	1.791
	Post-test	19	6.00	13.00	8.388	1.873
	Delayed post-test	19	4.00	13.00	8.717	2.758
Implicit	Pre-test	25	3.00	9.00	5.130	1.502
	Post-test	25	3.00	13.00	6.780	2.179
	Delayed post-test	25	3.00	11.00	6.375	1.851

The results indicated that verb complementation knowledge was promoted significantly as a function of test time, Wilk's Lambda = .65, $F(2, 61) = 16.32$, $p < .05$, $\eta^2 = .34$), and instructional group ($F(2,62) = 3.40$, $p < .05$, $\eta^2 = .10$). Also, there was a significant interaction between time and group ($F(4,122) = 2.93$, $p < .05$, $\eta^2 = .08$).

In order to detect which of the three groups differed from each other, pairwise comparisons were made. Pre-test scores showed no significant difference between any pair of the groups. Since the assumption of homogeneity of variance was violated in case of post- and delayed post-tests, significance level was set at .01. On the post-test, it was found that the difference between explicit group and control group was significant ($p < .01$), but the difference between explicit and implicit groups was not so ($p = .03$). It was also found that the difference between implicit and control groups was not significant ($p = 1.0$). In delayed post-test, there were two

[DOI: 10.29252/ijal.20.1.81]

significant differences: between explicit group and the other groups (for both $p < .01$).

On the basis of the results of ANOVA, Bonferroni tests were also conducted on the pre-, post- and delayed post-tests of each group. Accordingly, test scores of examinees in control group did not differ significantly from pre-test to post-test ($p = .8$) and delayed post-test ($p = .9$). On the other hand, the explicit group did significantly better on both post-test ($p < .05$) and delayed post-test ($p < .05$). The post-hoc analysis of implicit group scores revealed a significant difference between pre-test and post-test ($p < .05$), but not pre-test and delayed post-test ($p = .1$).

The research question addressed whether EI or II was more effective over time on a measure of explicit/implicit language knowledge. Regarding the untimed GJT, the analysis showed a substantial main effect for time. An investigation of the mean scores showed a general rise of the scores from pre-test to post-test for all three groups with the highest gain for the explicit group (from 5.58 to 8.71). Considering the results of the Bonferroni test, those who received explicit instruction could improve their score significantly from pre-test to post- and delayed post-tests; meanwhile, the effect of instruction for implicit group produced significantly higher results only in post-test. The scores of the participants in control group did not change significantly. These results indicated that in the case of using untimed GJT, EI yielded more durable effects. Also, the results of the mixed between-within subjects ANOVA on the untimed GJT scores showed the significant effect of instructional group. Further analysis showed that in post- and delayed post-tests, the students in the explicit group outperformed both implicit and control groups. The quantitative results also suggested that the performance of the implicit group did not differ significantly from that of the control group. Taken together, the results indicated that the effect of

[DOI: 10.29252/ijal.20.1.81]

EI was superior to II in promoting learners knowledge of verb complementation when measured through untimed GJT.

With respect to ungrammatical items of the untimed GJT, too, the effect of time was found to be significant. Post-hoc analysis of the ungrammatical sentences in the untimed GJT revealed a significant increase of scores only for explicit group from pre-test to post- and delayed post-tests; however, the implicit group did not improve any more than the control group. Given the fact that almost similar results were obtained from ungrammatical items of untimed GJT, the durability of EI is further corroborated. The differential results of the analyses observed between total test scores and ungrammatical items scores of implicit instruction group is quite logical by considering the fact that the ungrammatical items afford a better measure of explicit knowledge. Furthermore, the main effect comparing the experimental groups was significant. Post-hoc analyses of the results showed that the explicit group performed significantly better than the other groups on post- and delayed post-tests. On the other hand, the results for the implicit group showed that the group failed to outperform the control group in a significant way. These results further confirm that EI is more effective than II for the development of explicit knowledge on verb complementation in the case of using an explicit measure.

Finally, with respect to the implicit test, the results showed a significant main effect for time. Though in the beginning, the control group had the highest mean, over time they made little progress. Concerning the results of the instructional groups, both explicit and implicit groups made significant gains from pre-test to post-test; however, the learners in the explicit group, but not the implicit group, were able to maintain their progress over the weeks following the treatment. The analyses conducted on the results of the timed GJT also supported the effect of EI on the development of verb complementation. Additional information obtained from post-hoc tests

[DOI: 10.29252/ijal.20.1.81]

showed that learners in the explicit group displayed a greater performance than learners in the control group in post- and delayed post-tests. Further, the comparison of the scores revealed that those who received EI displayed a significantly better performance than those who received II only in delayed post-test. Lastly, the participants in the implicit group were unable to outperform those in the control group significantly. The results of the timed GJT were almost similar to those of the untimed GJT and ungrammatical items of the untimed GJT meaning that when either timed or untimed GJT is used, the effect of EI is superior to II in promoting learners knowledge of verb complementation.

The advantage of EI over II found in this study is in line with Spada and Tomita's (2010) meta-analysis in which they found that regardless of linguistic complexity, II does not appear to have as significant an effect as EI and that the effect of II on controlled and free outcome measures revealed small and medium effect sizes. The results of this study also accord with a number of studies that have found a positive effect for EI compared to II (Andringa, Glopper, & Hacquebord, 2011; De la Fuente, 2009; De Graaff, 1997; Hernández, 2011; Radwan, 2005; Rosa & Leow, 2004; Tode, 2007; Yoshimi, 2001). Our results are generally similar to those of Andringa, et al. (2011). In the case of one of their target structures, that is, degrees of comparison (DoC) in Dutch, they found an interaction between instruction and time on GJT: the explicit group had performed significantly better in both post-tests. The results of their implicit test of free written task also favored EI excluding those participants whose L1 expresses the DoC differently from the Dutch DoC. In a similar vein, Radwan (2005) found that compared to students who were assigned to textual enhancement condition, those who were presented with rule explanation displayed a greater performance in both the explicit test and the implicit test.

[DOI: 10.29252/ijal.20.1.81]

Another major finding is that the II did not lead to significantly greater knowledge development compared to the control group. This indicates that it might have been repeated exposure to the target structure that enabled the learners in these two groups to obtain almost similar improved scores. This result aligns with the conclusions of the previous studies that textual enhancement, as an implicit approach, has no effect on intake, acquisition, or comprehension (Alanen, 1995; Izumi, 2003; Leow, Egi, Nuevo, & Tsai, 2003; Radwan, 2005; Wong, 2003). This observation is consistent with Reinders and Ellis' (2009) results that overall input enhancement and enriched input condition had no effect on the development of learners' implicit and explicit knowledge as measured by timed and untimed GJTs, respectively. These findings could be interpreted in light of Ellis' (2004) proposal that performing on a GJT involves three processing operations: (1) semantic processing (i.e., recognizing sentence meaning); (2) noticing (i.e., looking for formal anomaly in the sentence); and (3) reflecting (i.e., identifying what the ill form is and why it is incorrect). If we agree with Ellis and Loewen (2007) that both semantic processing and noticing hinge upon implicit knowledge, it follows that when GJTs are time-constrained, the participants have enough time to access only their implicit knowledge to determine grammaticality of a given sentence, though this may not be sufficient to access their explicit knowledge to determine ungrammaticality of a given sentence. Formerly, Bialystock (1979) had expressed the same idea. Bialystock inferred from the result of her study that the participants make grammaticality judgment based on their implicit knowledge unless more detailed and fine-grained decision is to be made. That participants in the implicit group of this study could not make appropriate grammaticality judgments even in the case of timed GJT indicates that textual enhancement was not effective enough to induce changes in the participants'

[DOI: 10.29252/ijal.20.1.81]

interlanguage by forming implicit knowledge. The participants may have been unable to work out the rule.

The last finding of the present study documented that EI yielded more delayed effects than II. This finding as Ellis (1993) contended may relate to the nature of implicit knowledge. This type of knowledge is slow and laborious to form due to the extra time required for internal processing to convert input to implicit knowledge (Gass, 1997; Nassaji & Fotos, 2004; VanPatten, 1996). The length of II or the time between the last treatment session and post-test and also delayed post-test may not have been long enough for the internal processes to take place. Of course, this finding should be treated with caution because a closer look at the mean scores of the implicit and no instruction groups over time showed that generally learners in the former group made more progress, though not significantly, compared to the latter group.

6. Conclusion

The differential effects of EI and II on target forms have been widely investigated. Most of these studies, which have furnished the results in favor of EI, have been criticized on the grounds that their results are contaminated by the use of biased measures. This criticism has rendered the operationalization and measurement of implicit and explicit knowledge with the help of appropriate measures a hectic area of research. In the current project, despite the use of separate measures of explicit and implicit knowledge, the results still testify EI leads to more gains compared to II. Norris and Ortega (2000) complained about the inconsistency of instructional method operationalization. One way to surmount this problem is to compare the result of each instructional group with a control group. Even in the case of this study, the results revealed II improved learners'

[DOI: 10.29252/ijal.20.1.81]

performance only marginally compared to control group. Of course, it is premature to conclude that II of textual enhancement type is ineffective because as Lee and Huang (2008) mentioned in their meta-analysis, prior knowledge, learner proficiency, intensity of treatment, and developmental readiness are the intervening variables that might reduce the effect of visual input enhancement. Therefore, the textual enhancement format used in this study might come to be effective in another context. Based on the same line of argument, we can claim while the present study showed little effect of II, it does not rule out the possibility that longer instructional treatments might produce different results given that short-time instruction is inherently biased against implicit learning (Dekeyser, 2009) and this type of learning is a cumulative process. This is true especially in light of Gass's (1997) claim that repeated exposure to a specific target structure can have an impact on noticing and Ellis' (1993) assertion that the nature of implicit L2 learning is slow. Therefore, further studies should incorporate longer periods of treatment and more exposure to target features in order to better assess efficiency of II.

Finally, we wish to acknowledge two shortcomings of our study. First, the explicit group in this study received more exposure to the target verbs than the implicit group because of the additional handout. This might have confounded the results of the study. Second, compared to the other measures of implicit knowledge, namely oral production and elicited imitation, timed GJT is the least implicit one. Thus, the administration of other implicit measures might have tipped the scale less in favor of EI.

[DOI: 10.29252/ijal.20.1.81]

7. References

Akakura, M. (2012). Evaluating the effectiveness of explicit instruction on implicit and explicit L2 knowledge. *Language Teaching Research, 16*(1), 9-37.

Alanen, R. (1995). Input enhancement and rule presentation in second language acquisition. In R. Schmidt (Ed.), *Attention and awareness in foreign language acquisition* (pp.259-302). University of Hawaii, Honolulu.

Andringa, S., Glopper, K., & Hacquebord, H. (2011). Effect of explicit and instruction on free written response task performance. *Language Learning, 61*(3), 868-903.

Bachman, L. (1990). *Fundamental considerations in language testing*. New York: Oxford University Press.

Bialystok, E. (1979). Explicit and implicit judgments of L2 grammaticality. *Language Learning, 29*(1), 81-103.

Bowles, M. A. (2011). Measuring implicit and explicit linguistic knowledge: What can heritage language learners contribute? *Studies in Second Language Acquisition, 33*(2), 247–271.

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics, 1*(1), 1-47.

Celce-Murica, M., Dornyei, Z., & Thurrell, S. (1995). Communicative competence: A pedagogically motivated model with content specifications. *Issues in Applied Linguistics, 6*(2), 5–35.

De Graaff, R. (1997). The Experanto experiment: Effects of explicit instruction on second language acquisition. *Studies in Second Language Acquisition, 19*(2), 249-276.

[DOI: 10.29252/ijal.20.1.81]

De Graaff, R., & Housen, A. (2009). Investigating the Effects and Effectiveness of L2 Instruction. In M. H. Long, & C. J. Doughty (Eds.), *The Handbook of Language Teaching* (pp. 726-775). Wiley-Blackwell, Oxford, UK.

DeKeyser, R. (1995). Learning second language grammar rules: An experiment with a miniature linguistic system. *Studies in Second Language Acquisition, 17*(3), 379-410.

DeKeyser, R. M. (2003). Implicit and explicit learning. In C. Doughty, & M. Long (Eds.), *The Handbook of Second Language Acquisition* (pp. 313-349). Oxford: Blackwell.

De Keyser, R. M. (2009). Cognitive-psychological processes in second language learning. In M. Long, & C. Doughty (Eds.), *The handbook of language teaching* (pp. 119-138). Malden, MA: Blackwell.

De la Fuente, M. (2009). The role of pedagogical tasks and focus on form in acquisition of discourse markers by advanced learners. In R. P. Leow, H. Campos, & D. Lardiere (Eds.), *Little words: Their history, phonology, syntax, semantics, pragmatics, and acquisition* (pp. 211-221). Washington, DC: Georgetown University Press.

Douglas, D. (1998). Testing methods in context-based second language research. In L. Bachman, & A. Cohen (Eds.), *Interfaces between second language acquisition and language testing research* (pp. 141–155).Cambridge: Cambridge University Press.

Douglas, D. (2001). Performance consistency in second language acquisition and language testing: A conceptual gap. *Second Language Research, 17*(4), 442-456.

Doughty, C. J., & Williams, J. (1998). Pedagogical choices in focus on form acquisition. In C. J. Doughty & J. Williams (Eds.), *Focus on form in*

[DOI: 10.29252/ijal.20.1.81]

classroom second language acquisition (pp. 197-262). Cambridge: Cambridge University Press.

Ellis, R. (1993). The structural syllabus and second, *TESOL Quarterly 27*(1), 91-113.

Ellis, R. (2001). Introduction: Investigating form-focused instruction. *Language Learning, 51*, supplement 1, 1-46.

Ellis, R. (2004). The definition and measurement of explicit knowledge. *Language Learning, 54*(2), 227-275.

Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition, 27*(2), 141-172.

Ellis, R. (2008). Explicit form-focused instruction and second language acquisition. In B. Spolsky, & F. M. Hult (Eds.), *The handbook of educational linguistics* (pp. 437-455). Oxford: Blackwell.

Ellis, R. (2009). Implicit and explicit learning, knowledge and instruction. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philp, & H. Reinders (Eds.), *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 3-25). Bristol, UK: Multilingual Matters.

Ellis, R., & Loewen, S. (2007). Confirming the operational definitions of explicit and implicit knowledge in Ellis (2005): Responding to Isemonger. *Studies in Second Language Acquisition, 29*(1), 119-126.

Ellis, R., Loewen, S., Edler, C., Erlam, R., Philp, J., & Reinders, H. (Eds.). (2009). Implicit and explicit knowledge in second language learning, testing, and teaching. Bristol, UK: Multilingual Matters.

Gass, S. M. (1997). *Input, interaction, and the second language learner*. Mahwah, NJ: Lawrence Erlbaum.

Godfroid, A., Loewen, S., Jung, S., Park, J., Gass, S., & Ellis, R. (2015). Timed and untimed grammaticality judgments measure distinct types of knowledge. *Studies in Second Language Acquisition*, *37*(2), 269-297.

[DOI: 10.29252/ijal.20.1.81]

Gutiérrez, X. (2013). The construct validity of grammaticality judgment tests as measures of implicit and explicit knowledge. *Studies in Second Language Acquisition, 35*(3), 423-449.

Han, Y., & Ellis, R. (1998). Implicit knowledge, explicit knowledge and general language proficiency. *Language Teaching Research, 2*(1), 1-23.

Hernández, T. A. (2008). The effect of explicit instruction and input flood on students' use of discourse markers on a simulated oral proficiency interview. *Hispania, 91*, 665-675.

Hernández, T. A. (2011). Re-examining the role of explicit instruction and input flood on the acquisition of Spanish discourse markers. *Language Teaching Research, 15*(2), 159-182.

Hulstijn, J. H. (2005). Theoretical and empirical issues in the study of implicit and explicit second-language learning: Introduction. *Studies in Second Language Acquisition, 27*(2), 129-140.

Izumi, S. (2003). Visual input enhancement as focus on form. *Sophia Linguistica, 13*, 1-30.

Klapper, J., & Rees, J. (2003). Reviewing the case for explicit grammar instruction in the university foreign language learning context. *Language Teaching Research, 7*(3), 285-314.

Lee, S., & Huang, H. (2008). Visual input enhancement and grammar learning: A meta-analytic review. *Studies in Second Language Acquisition, 30*(3), 307-331.

Leow, R., Egi, T., Nuevo, A. M., & Tsai, Y.-C. (2003). The roles of textual enhancement and type of linguistic item in adult L2 learners' comprehension and intake. *Applied Language Learning 13*(2), 1-16.

Loewen, S. (2009). Grammaticality judgment tests and the measurement of implicit and explicit L2 knowledge. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philp, & H. Reinders (Eds.), *Implicit and explicit knowledge in*

[DOI: 10.29252/ijal.20.1.81]

second language learning, testing and teaching (pp. 94-112). Bristol, UK: Multilingual Matters.

Loewen, S., Erlam, R., & Ellis, R. (2009). The incidental acquisition of third person -s as implicit and explicit knowledge. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philp, & H. Reinders (Eds.), *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 262-280). Bristol, UK: Multilingual Matters.

Morgan-Short, K., Sanz, C., Steinhauer, K., & Ullman, M. T. (2010). Second language acquisition of gender agreement in explicit and implicit training conditions: An event-related potential study. *Language Learning, 60*(1), 154-193.

Morgan-Short, K., Sanz, C., Steinhauer, K., & Ullman, M. T. (2012). Explicit and implicit second language training differentially affect the achievement of native-like brain activation patterns. *Journal of Cognitive Neuroscience, 24*(4), 933-947.

Nassaji, H., & Fotos, S. (2004). Current developments in research on the teaching of grammar. *Annual Review of Applied Linguistics, 24*, 126-145.

Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning, 50*(3), 417-528.

Pallant, J. (2016). *SPSS survival manual: A step by step guide to data analysis* (6th Ed.). Open University Press: New York.

Philp, J. (2009). Pathways to proficiency: Learning experiences and attainment in implicit and explicit knowledge of English as a second language. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philp, & H. Reinders (Eds.), *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 167-193). Bristol, UK: Multilingual Matters.

[DOI: 10.29252/ijal.20.1.81]

Radwan, A. A. (2005). The effectiveness of explicit attention to form in language learning. *System, 33*(1), 69-87.

Reinders, H., & Ellis, R. (2009). The effects of two types of input on intake and the acquisition of implicit and explicit knowledge. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philp, & H. Reinders (Eds.), *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 281-302). Bristol, UK: Multilingual Matters.

Rosa, E. E., & Leow, R. P. (2004). Computerized task-based exposure, explicitness and type of feedback on Spanish L2 development. *Modern Language Journal, 88*(2), 192-217.

Scott, M. V. (1990). Explicit and implicit grammar teaching strategies: New empirical data. *The French Review, 63*(5), 779-789.

Spada, N., & Lightbown, P. M. (1993). Instruction and the development of questions in L2 classrooms. *Studies in Second Language Acquisition, 15*(2), 205-224.

Spada, N., & Tomita, Y. (2010). Interactions between type of instruction and type of language feature: A meta-analysis. *Language Learning, 60*(2), 263-308.

Suzuki, Y., & DeKeyser, R. (2015). Comparing elicited imitation and word monitoring as measures of implicit knowledge. *Language Learning, 65*(4), 860-895.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th Ed.). Boston: Pearson.

Tode, T. (2007). Durability problems with explicit instruction in an EFL context: The learning of the English copula be before and after the introduction of the auxiliary be. *Language Teaching Research, 11*(1), 11-30.

[DOI: 10.29252/ijal.20.1.81]

Vafaee, P., Suzuki, Y., & Kachisnke, I. (2017). Validating grammaticality judgment tests: Evidence from two new psycholinguistic measures. *Studies in Second Language Acquisition, 39*(1), 59-95.

VanPatten, B. (1996). *Input processing and grammar instruction in second language acquisition*. Norwood, NJ: Ablex.

Varnosfadrani, A. D., & Basturkmen, H. (2009). The effectiveness of implicit and explicit error correction on learners' performance. *System, 37*(1), 82-98.

White, L. (1991). Adverb placement in second language acquisition: Some effects of positive and negative evidence in the classroom. *Second Language Research, 7*(2), 133-161.

Williams, J. (2005). Form-focused instruction. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 671–92). Mahwah, NJ: Lawrence Erlbaum.

Wong, W. (2003). The effects of textual enhancement and simplified input on L2 comprehension and acquisition of non-meaningful grammatical form. *Applied Language Learning, 14*, 109-132.

Yoshimi, D. R. (2001). Explicit instruction and JFL learners' use of interactional discourse markers. In K. Rose, & G. Kasper (Eds.), *Pragmatics in language teaching* (pp. 223-244). Cambridge: Cambridge University Press.

Zhang, R. (2015). Measuring university-level L2 learners' implicit and explicit linguistic knowledge. *Studies in Second Language Acquisition, 37*(3), 457-486.

Zhuo, C. (2010). Explicit recast, implicit recast and the acquisition of English noun plural: A comparative study. *Chinese Journal of Applied Linguistics, 33*(6), 55-70.

[DOI: 10.29252/ijal.20.1.81]

Notes on Contributors:

Sasan Baleghizadeh is associate professor of TEFL at Shahid Beheshti University in Tehran, Iran, where he teaches courses in applied linguistics, syllabus design, and materials development. He is interested in investigating the role of interaction in English language teaching and issues related to materials development.

Ali Derakhshesh is currently a Ph.D. candidate of TEFL at Shahid Beheshti University. He has made several presentations at different national conferences, and has compiled a number of books and authored articles related to TEFL. His main areas of research interest include psycholinguistics and interlanguage pragmatics.

[DOI: 10.29252/ijal.20.1.81]