The purpose of this study was to investigate the effect of the input-based and output-based tasks with different and identical degrees of involvement loads on Iranian EFL learners’ incidental vocabulary learning. The participants were 120 pre-intermediate EFL learners from five English language institutes in Iran. In Phase 1, the participants received input-based and output-based tasks with identical involvement loads. In phase 2, they received both types of tasks but the involvement load of input-based tasks was higher than that of the output-based ones. Finally, in phase 3, the participants received output-based tasks with higher involvement loads. To measure the amount of vocabulary learning, immediate post-tests and delayed post-tests were administered upon completion of the tasks and a week after the post-tests, respectively. The results revealed that input- and output-based tasks with identical involvement loads had a positive significant effect on students’ vocabulary learning and retention at both the post-test and the delayed post-test. Also, input-based tasks with higher involvement loads had no significant effect on students’ vocabulary learning and retention at the post-test and delayed post-test. Finally, output-based tasks with higher involvement loads had no significant effect on students’ vocabulary learning and retention at the post-test and delayed post-test.

ABOUT THE AUTHORS

Shiva Kaivanpanah is an Associate Professor at the University of Tehran. She has published numerous articles in the area of language teaching. Her current research interests include vocabulary studies, teacher education, and issues in teaching academic writing to English language learners.

Seyyed Mohammad Alavi is a Full Professor at the University of Tehran. He has published many articles in the area of language testing and assessment. His current research interests include, teacher education, research methodologies, vocabulary studies and cognitive and meta cognitive strategies.

Afsaneh Ravandpour is a PhD candidate in TEFL at the University of Tehran (Kish International Campus, Iran). She is a lecturer at the University of Kish. Her research interests are language teaching, learning assessment, effective teaching, teachers’ professional development, vocabulary learning, and educational technology.

PUBLIC INTEREST STATEMENT

McCarthy (1988) states, “no matter how well the students learn grammar, without words to express meanings, communication cannot happen meaningfully” (p. 1). Therefore, several researchers have explored ways to increase vocabulary knowledge. Laufer & Hulstijn (2001) proposed the Involvement Load Hypothesis (ILH). The involvement construct is composed of three components: need, search and evaluation. “The need is the motivational dimension of involvement. Search is the attempt to find the meaning of an L2 word by a dictionary or authority. Evaluation is a comparison of a given word with other words. The present study investigated the effect of input-based and output-based tasks with different and identical degrees of involvement loads on Iranian EFL learners’ incidental vocabulary learning. The results revealed that the condition with identical involvement load in the output-based group had a positive significant effect on students’ vocabulary learning and retention.
tasks with higher involvement loads had a positive significant effect on students’ vocabulary learning in both the post-test and delayed post-test. The findings have pedagogical implications for teachers and materials developers in developing effective tasks with a sufficient amount of involvement load.

Subjects: Language & Linguistics; Language Teaching & Learning; Literature

Keywords: incidental vocabulary learning; Involvement Load Hypothesis (ILH); input-based tasks; output-based tasks

1. Introduction

Learning a second or a foreign language involves not only learning the rules governing its grammar but also its vocabulary. In fact, vocabulary is generally considered a crucial part of L2 learning by both learners and teachers (Bao, 2015). As Nation (2001) maintains, expanding L2 learners’ vocabulary is not an aim in itself; instead, the ultimate goal is to assist them to listen, read, speak and write more efficiently. Limited or insufficient vocabulary can disrupt students’ learning process; therefore, language teachers should use suitable tasks to improve vocabulary acquisition processes.

While it is generally believed that the majority of vocabulary is acquired indirectly through reading and listening (See Nagy, Herman, & Anderson, 1985), direct learning tasks, like word pairs and word-focused tasks, have also been shown as effective methods for quick vocabulary learning (Laufer, 2005; Nation, 2001). Of primary concern in L2 vocabulary research, acquisition is an exploration of tasks that provides the best opportunity for learners to learn new words.

A significant amount of L2 vocabulary learning takes place incidentally, i.e. as a by-product of reading (Davis, 1989; Fraser, 1999; Jenkins, Stein, & Wysocki, 1984; Nation, 2001; Rieder, 2003) when individuals process new information without any intention to commit it to their memory. In fact, incidental vocabulary learning is “learning of vocabulary as a byproduct of any activity not explicitly geared to lexical learning” (Hulstijn & Laufer, 2001, p. 10). Moreover, cognitive involvement load of the tasks has also been known as a significant factor in vocabulary acquisition. Therefore, controlling the cognitive load of instructional tasks as well as their difficulty level could pave the way for improving the comparison between the different tasks and their usefulness. In this study, vocabulary learning was based on a cognitive processing perspective of learning and examined through the implementation of input- and output-based tasks with identical and different cognitive loads. Moreover, based on the Involvement Load hypothesis (ILH), vocabulary learning was assessed with the basic argument that learning words generally depends on their involvement load in processing. In fact, the more the vocabulary learning task is demanding for the learner, the more probable the word would be acquired.

According to ILH, originally proposed by Laufer and Hulstijn (2001), vocabulary acquisition is conditional upon the degree of involvement in processing new words. Moreover, the more the need, search and evaluation are involved in a task, the more effective the vocabulary learning task would be. In this hypothesis, need is defined as the quest for the required linguistic feature for achieving and completing the desired task such as knowing a particular word to comprehend a passage. Search refers to the effort to find the required information. Evaluation is the comparison of the word, or information about a word, with the context of utilization to determine to check if it fits or is the best choice. Notwithstanding the defined aspects of ILH and their contribution to learning, reading comprehensiveness would not necessarily need the whole word processing features happening in long-term retention (Rott & Williams, 2003). Thus, if the aim is to make new words acquired and retained, many of textual enhancement techniques such as adjunct aids (Marefat & Ghahari, 2009; Robinson, 1994), increased word frequency, or glosses in reading vocabulary guides could be used.
Due to the fact that remembering a word is dependent on the amount of attention allocated to the word, dedicating greater attention to new vocabulary may possibly lead to a greater likelihood for its acquisition (Rott & Williams, 2003). Thus, cognitive load involvement is of high importance in vocabulary learning literature. Krashen (1989) argues that we acquire vocabulary and spelling through exposure to comprehensible input. For him, comprehension is at the heart of language acquisition, and production is only a sign of second language acquisition that has already taken place (Krashen, 1981, 1989). However, researchers agree that exposure to the input may not necessarily help learners reach high L2 proficiency levels. Beside input, output plays an important role in the process of second language acquisition. Acknowledging the crucial role of comprehensible input in second language acquisition, Swain (1985) argues that it is not sufficient for learners of EFL learners to fully develop L2 proficiency. How input and output influence the comprehension and production of L2 forms and structures have been the subject of much research in SLA literature. Several studies have investigated the relative influences of input- and output-based instructions on vocabulary acquisition (e.g., Allen, 2000; DeKeyser & Sokalski, 1996; Erlam, 2009; Nagata, 1998).

Among the different factors in vocabulary acquisition, task type can be considered as one of the most challenging aspects examined by Laufer and Hulstijn (2001) as well as some other researchers (e.g., Walsh, 2009; Xu, 2009), emphasizing the advantages of output tasks over input ones. Considering the little attention paid to the role of input and output in the Involvement Load Hypothesis, the present study addressed this issue by examining the effectiveness of input- and output-based tasks with identical and different involvement loads on learning and retention of words.

2. Review of the literature

Learning a foreign language involves the acquisition of thousands of words. Therefore, vocabulary acquisition is considered as an important area of language teaching and learning research. As McCarthy (1990) argues, “no matter how well the student learns grammar, no matter how successfully the sounds of L2 are mastered, without words to express a wide range of meanings, communication in an L2 just cannot happen in any meaningful way” (p. 1). Words can be acquired incidentally or intentionally. Previous research has addressed incidental and intentional learning inseparably (Hunt & Belgar, 2005). Intentional learning requires focused attention to linguistic items whereas incidental learning requires attention to meaning with the difference that it allows attention to be directed to form, i.e., message content (Ellis, 1999). Thus, any learning, whether intentional or incidental, can only happen with some degree of attention (Schmidt, 1994), and the nature of information processing primarily determines the retention and acquisition of vocabulary (Hulstijn, 2001). Accordingly, ILH considers vocabulary acquisition dependent on the degree of involvement in processing. The construct of involvement is composed of three main components: need, search and evaluation. Need is “the motivational, non-cognitive dimension of involvement” (Hulstijn & Laufer, 2001, p. 543) which could be externally imposed by the teacher, etc. or self-imposed by the learner. Search is “the attempt to find the match between the form and meaning of an unknown word” (Bao, 2015, p. 85). Evaluation “involves a decision about the meaning of a given word, a comparison of its meaning with those of other words or its proper use in the specific context” (Bao, 2015, p. 85). Involvement is operationalized by designing tasks with varying degrees of need, search and evaluation.

In 2001, Hulstijn and Laufer conducted two parallel experiments in which their advanced Dutch and Hebrew English language learners were classified into six intact groups. Retention of 10 unfamiliar words in incidental learning context was examined in three tasks types with various involvement loads. Their findings indicated strong support for the ILH. Likewise, Kim (2011) presented empirical evidence for ILH in a carefully designed study involving two experiments. She observed the effect of three tasks with different involvement loads in two different levels of proficiency. Reading, gap-fill and composition tasks were randomly given to the learners in each proficiency group. Two immediate and delayed posttests examined the short-term or long-term vocabulary retention. The composition group with an involvement index of three achieved significantly better results than the reading group with an involvement index of one and gap-fill group with an involvement index of two. However, the results for the gap-fill and reading groups in
the delayed posttest indicated the superiority of the gap-fill group over the reading group. Similarly, in Kim's (2011) experiment, two tasks with equal load involvement indexes were used to examine if they can result in similar retention of vocabularies. He contrasted writing composition task and writing sentence task with three involvement indexes and found equal vocabulary retention in both tasks in immediate and delayed posttests.

In another study, Toth (2006) examined the role of input and output in second language learning of Spanish morphosyntax and found that both groups equally improved on a grammar task, but the group receiving output-based tasks outperformed the other group in the controlled production task. In a similar vein, Maftoon and Sharifi (2012) found that output-oriented tasks could push learners to notice a wider web of form-meaning associations between real concepts and vocabulary to be learned. In 2013, Sarani, Mousapour Negari and Ghaviniat found that productive tasks were more effective than receptive tasks, and task involvement load was a crucial factor for vocabulary learning. Some recent studies (e.g., Alavinia & Rahimi, 2019; Bao, 2015; Tahmasbi & Farvardin, 2017) found that output-based tasks are very effective in enhancing learners’ vocabulary knowledge. Finally, Kaivanpanah and Miri (2018) investigated whether the difference in task-induced involvement could affect the actual realization of the evaluation stage in ILH and found that utilizing target words in the composition task, as compared to cloze tests, could induce a higher degree of evaluation.

Review of recent studies indicates that output-based tasks can help learners effectively establish stronger links between form and meaning of a word (Hu & Nassaji, 2016). Previous studies have mostly used tasks with similar levels of cognitive involvement to facilitate foreign language vocabulary learning. To address gaps in the literature, the present study was an attempt to consider the effect of task type (input- vs. output-based) and involvement load (identical or different) over time (post-test & delayed post-test). More specifically, the current study was motivated by the following research questions:

1. Is there any significant difference in the effectiveness of input-based and output-based tasks with identical involvement loads on learning and retention of words?
2. Is there any significant difference in the effectiveness of input-based and output-based tasks with different involvement loads on learning and retention of words?

3. Method

3.1. Participants
The participants were 180 Iranian EFL learners (age range = 17–26) from five private English language institutes in two Iranian cities and were selected via convenience sampling. A version of Oxford Proficiency Test (OPT) was administered to check the homogeneity of the learners’ English language proficiency. The learners whose score was one standard deviation below and above the mean (N = 120) comprised the main participants of the study.

3.2. Materials

3.2.1. Reading text
At the piloting stage, 15 randomly selected reading texts from Select Reading Series (Lee & Gundersen, 2002) were administered to sample of 100 students in order to check if there were any difficult words, determine the time required for completing the tasks, and any probable problems in implementing the tasks. Finally, one pre-intermediate text with approximately 325 words was chosen.

3.2.2. Tasks
Four tasks including two input-based and two output-based with different involvement loads were developed based on the selected reading text. Students were required to complete the tasks after reading each text.
3.2.2.1. Input-based tasks. These tasks focus more on comprehension and less or none on production (Krashen, 1985; Takimoto, 2007)

(1) Reading and comprehension questions; target words not glossed in texts but relevant to the tasks: This type of task-induced need, search to look up the words in a dictionary, and some evaluation. Therefore, the involvement index was 3 \( (1 + 1 + 1 = 3) \).

(2) Reading and comprehension questions; target words glossed in text but irrelevant to the tasks: In this task, unknown words were glossed and the comprehension questions could be answered without any reference to these words. It did not induce any need to focus on the glossed words irrelevant to the task, nor did it require any search for the meaning of unknown words since they were glossed. In addition, the task did not induce any evaluation. Therefore, the involvement index was 0 \( (0 + 0 + 0 = 0) \).

3.2.2.2. Output-based tasks. These tasks focus on both comprehension and production, but more emphasis is put on production (Van Patten & Sanz, 1995).

(1) Reading and comprehension questions and filling gaps; target words relevant to reading comprehension and listed with glosses at the end of the text: 10 new words were deleted from the text, leaving 10 gaps numbered 1–10. The 10 new words, along with five words which did not appear in the original text, were printed with their L2 explanations in a random order on a separate page. The participants were required to read the text, fill in the gaps with the missing words from a list of 15 words, and answer the comprehension questions. In terms of involvement load, the task-induced moderate need, no search, and moderate evaluation. Therefore, the involvement index was 2 \( (1 + 0 + 1 = 2) \).

(2) Writing a composition with glossed words in texts: This task involved writing a composition on a proposed topic. More specifically, the students were provided with new words and asked to write a composition. In terms of the involvement load, the task-induced moderate need, no search, and a strong evaluation as the new words were evaluated against suitable collocations in a learner-generated context. Therefore, the involvement index was 3 \( (1 + 0 + 2 = 3) \).

3.3. Instrumentation

3.3.1. Oxford placement test (OPT)
A version of OPT was employed to check the homogeneity of the participants in terms of English language proficiency. This test consists of 200 multiple-choice items (100 grammar and 100 listening comprehension questions) and is considered a reliable measure of English proficiency.

3.3.2. Immediate post-test
Immediately after the treatment and without any prior notice, a post-test was administered to measure the amount of vocabulary learned by participants in the experimental groups. The test included 30 vocabulary items instructed during the study. Following Hulstijn and Laufer’s (2001) scoring rubrics, any word “not translated” or “incorrectly translated” received a zero score; a “correct” answer was given a full score; and a “semantically approximate” answer received half a score.

3.3.3. Delayed post-test
One week after the immediate post-test, the delayed post-test was administered to measure students’ vocabulary retention.

3.4. Procedure
The present study was conducted in two phases:
3.4.1. Phase one: homogenizing the participants and piloting the materials and target words

In the first week, the OPT was administered to 180 language learners to ensure their homogeneity in terms of language proficiency. Based on OPT scores ranging between 120 and 134, learners who scored one standard deviation above and below the mean (N = 120) were selected as the main participants. Then, they were randomly assigned to six experimental groups. In the second week, 15 randomly selected texts on a wide variety of topics from Select Readings Series (pre-intermediate level) were given to a similar group of participants (N = 10) to identify unknown/new words. Each text had at least 20 unknown words; a word was classified as unknown if it was selected by at least 70% of the participants. Afterward, a text with 10 unknown words was chosen for the treatment and post-test sessions. Then, tasks similar to those used by Hulstijn and Laufer (2001) were developed for the present study and piloted on a sample of 10 participants who were similar to the participants in the main study in terms of language proficiency. The purpose of piloting was to determine the time required for completing the tasks and determining any probable problems in implementing the tasks. In terms of the required time, it was found that 45–60 min was sufficient for completing the tasks. Similar to Hulstijn and Laufer (2001), we considered time on task as an inherent property of the task. The pilot study revealed that some parts of the texts were ambiguous for the learners; therefore, necessary explanations were added to the tasks. In short, the pilot study aimed to examine and improve the tasks and their implementation procedure.

3.4.2. Phase two: main study

The main study was conducted after the piloting phase. The participants were not informed about the test time since test announcement is considered part of an intentional vocabulary learning process (Hulstijn, 2003). Tasks were completed during usual class periods and time on task was different for all tasks.

The following steps were taken in the main study:

In the fourth week, a two-week instructional treatment started. The participants in each group received different tasks. Immediately after the treatment, a post-test was administered to measure the amount of vocabulary learning. In the seventh week, i.e. 1 week after the immediate posttest, a delayed posttest test was administered to measure learners’ vocabulary knowledge retention. Therefore, at the first phase, for one group, the input-based tasks and for another group, the output-based tasks with identical involvement loads tasks were practiced. As noted earlier, input-based tasks included reading and comprehension questions which were not glossed and relevant to the task while output-based tasks included writing a composition with glossed words in texts. Students in output-based group were asked to write a composition on the topic of “the book of future”.

At the second phase, there were two groups, one receiving input- and the other output-based tasks with different involvement loads. In this phase, the involvement load of input-based tasks was higher than output-based tasks. The input-based task included reading and comprehension questions which were not glossed but relevant to the task. Here, the involvement load index was three. Nevertheless, the output-based tasks included reading comprehension plus “fill in” task. Students in this group were given the same text. For this group, the 10 new words were deleted from the text, leaving 10 gaps numbered 1–10. The 10 new words, along with five words that had not appeared in the original text, were given in random order as a list on a separate page, with their L2 explanations. The task required learners to read the text, fill in the 10 gaps with the missing words from the list of 15 words, and answer the comprehension questions. In terms of the involvement load, filling gap task-induced moderate need, no search and moderate evaluation because the context was provided. Its involvement index was 2.

At the third phase, the involvement loads of tasks for both groups were different. More specifically, the involvement load of the output-based task was high than the input-based task. The input-based task included reading comprehension questions glossed in the text but irrelevant to
tasks. Here, involvement load index was zero. In contrast, the output-based task included reading comprehension plus “fill in”. After completing tasks, an immediate post-test was administered without informing the students. Also, 1 week after the immediate post-test, the delayed post-test was administered to measure the amount of vocabulary retention of experimental groups. Finally, the results for groups with identical and different involvement loads in three different phases were compared and analyzed through independent samples t test.

4. Results

To check the normality of data distribution, Shapiro–Wilk test was employed (Table 1).

As it can be seen, the obtained p value in all groups is higher than .05. Therefore, it can safely be concluded that the data is normally distributed across all the variables.

The Difference in the effectiveness of input-based and output-based tasks with identical and different involvement loads on learning and retention of words

4.1. Post-tests of input-based and output-based groups with identical involvement loads

To examine whether vocabulary learning differed significantly between the two groups (input-based and output-based) after the treatment (identical involvement load), an independent-samples t-test was performed (Table 2).

As indicated in Table 2, there was a statistically significant difference between the groups receiving output-based tasks (M = 17.95, SD = 1.46) and input-based tasks (M = 16.55, SD = 1.57), (t(38) = −2.91, p = .006, d = .92) in the post-test.

In addition, in order to examine whether the retention of words differs significantly between the two groups (input-based and output-based) in the delayed post-test (identical involvement loads), an independent samples t-test was performed (Table 3).

4.2. Phase 2

4.2.1. Post-tests of input-based and output-based groups with different involvement loads (Input-based tasks with higher involvement load)

To answer the second research question aiming to examine whether the learning of words differs significantly between the groups receiving input-based and output-based tasks (input-based tasks with higher involvement load), an independent-samples t-test was performed to compare their performance after the treatment (Table 4).

4.2.2. Delayed post-tests of input-based and output-based groups with different involvement loads (Input-based tasks with higher involvement load)

An independent-samples t-test was performed to examine whether the retention of words differs significantly between the groups receiving input-based and output-based tasks (input-based tasks with more involvement load) at the delayed post-test. Table 5 shows the descriptive and inferential results.

| Tests                  | Group                                         | Df  | Sig. |
|------------------------|-----------------------------------------------|-----|------|
| (1) Post-test (phase 1)| Input-based and output based with identical involvement loads | 40  | .15  |
| (2) Post-test (phase 2)| Input-based and output based with different involvement loads | 40  | .07  |
| (3) Post-test (phase 3)| Input-based and output based with different involvement loads | 40  | .09  |
Table 2. The results of independent sample t-test for input-based and output-based groups in vocabulary post-test with identical involvement load

| Vocabulary post-test phase 1 | Means | SD | Levene’s Test for Equality of Variances | t-test for Equality of Means |
|----------------------------|-------|----|----------------------------------------|----------------------------|
|                            |       |    | F | Sig. | T   | df | Sig. (2-tailed) | Mean Difference | Std. Error Difference |
| Input-based                | 16.55 | 1.57 | .413 | .52 | -2.91 | 38 | .006 | -1.40 | .48 |
| Output-based               | 17.95 | 1.46 |     |     |       |    |       |       |     |
As it can be seen in Table 3, there was a statistically significant difference between the groups receiving output-based tasks ($M = 17.65$, $SD = 1.30$) and input-based tasks ($M = 15.20$, $SD = 1.36$), ($t$ (38) = −5.80, $p = .000$, $d = 1.84$).

Table 3. Results of independent samples t-test for input-based and output-based groups for tasks with identical involvement loads in delayed post-test

| Voc Delayed post-test with identical involvement load | Means | SD  | Levene's Test for Equality of Variances | t-test for Equality of Means |
|------------------------------------------------------|-------|-----|----------------------------------------|-----------------------------|
|                                                      |       |     |                                        | t        | df    | Sig. (2-tailed) | Mean Difference | Std. Error Difference |
| Input-based                                           | 15.20 | 1.36| .413                                  | −2.91  | 38    | .006                         | −1.40           | 48                   |
| Output-based                                          | 17.65 | 1.30| .52                                   |         |       |                              |                 |                      |
Table 4. Results of independent sample t-test for input-based and output-based groups in vocabulary post-test with different involvement load (input-based tasks with higher involvement load)

| Post-test phase 2 | Means | SD  | Levene's Test for Equality of Variances | t-test for Equality of Means |
|-------------------|-------|-----|----------------------------------------|-----------------------------|
|                   |       |     | F           | Sig.     | T       | Df | Sig. (2-tailed) | Mean Difference | Std. Error Difference |
| Input-based       | 16.95 | 1.35 | .126       | .72      | .89     | 38 | .37            | .40              | .44               |
| Output-based      | 16.55 | 1.46 |            |          |         |     |                |                  |                  |

As seen in Table 4, there is no statistically significant difference between the groups receiving output-based tasks ($M = 16.55, \text{SD} = 1.46$) and input-based tasks ($M = 16.95, \text{SD} = 1.35$), ($t(38) = .89, p = .37$) in the post-test.
Table 5. Results of independent sample t-test for input-based and output-based groups in vocabulary delayed post-test with different involvement load (input-based tasks with higher involvement load)

| Delayed Post-test phase 2 | Means | SD  | Levene’s Test for Equality of Variances | t-test for Equality of Means |
|---------------------------|-------|-----|----------------------------------------|-------------------------------|
|                           |       |     | F          | Sig. | T      | Df  | Sig. (2-tailed) | Mean Difference | Std. Error Difference |
| Input-based               | 16.10 | 1.32 | .15        | .69  | -.69   | 38  | .49             | -.30             | .43               |
| Output-based              | 15.80 | 1.41 |            |      |        |     |                |                 |                   |

According to Table 5, there is no statistically significant difference between the groups receiving output-based tasks (\(M = 15.80, SD = 1.41\)) and input-based tasks (\(M = 16.10, SD = 1.32\)), \(t(38) = -.69, p = .492\) in delayed post-test.
4.3. Phase 3

4.3.1. Post-tests of input-based and output-based groups with different involvement load (Output-based tasks with higher involvement load)
To examine whether vocabulary learning differs significantly between the groups receiving input-based and output-based tasks (output-based tasks with higher involvement load) after the treatment, an independent-samples t-test was performed (Table 6).

4.3.2. Delayed post-tests of input-based and output-based groups with different involvement load (Output-based tasks with higher involvement load)
To examine whether the retention of words differs significantly between the two groups (input-based and output-based) after the treatment (different involvement load with more output-based tasks), an independent-samples t-test was performed. The results are presented in Table 7.

5. Discussion
The present study examined the effect of the input-based and output-based tasks with different and identical involvement loads on Iranian EFL learners’ incidental vocabulary learning. The findings indicated that there was a significant difference in terms of vocabulary learning between the groups receiving input-based and output-based tasks with identical involvement loads in post-test and delayed post-test. In fact, students receiving output-based tasks outperformed those receiving input-based tasks in both post-test and delayed post-test. In addition, there was no significant difference between vocabulary learning of students receiving input-based tasks with higher involvement loads in both post-test and delayed post-test. Finally, there were significant differences between students receiving output-based tasks with higher involvement loads in both the post-test and delayed post-test. It is noteworthy that students receiving output-based tasks had a better performance in the post-test while students receiving input-based tasks had a better performance in the delayed post-test.

Although the task involvement load is an essential factor for vocabulary acquisition, the current study showed that output-based tasks—irrespective of their involvement load—have led to better vocabulary learning and retention in almost all phases. The results of the present study are in line with previous research (e.g., Al-Hadlaq, 2003; Bao, 2015; Khonamri & Hamzenia, 2013; Marmol & Sanchez-Lafuente, 2013; Sarani, Negari, & Ghaviniat, 2013; Soleimani & Rahmanian, 2015; Soleimania, Rahamian, & Sajedia, 2014; Xu, 2009), all confirming the significant positive effect of output-based tasks on L2 vocabulary learning and retention. More specifically, the findings of this study confirmed Laufer and Hulstijn’s (2001) argument that learning and retention of vocabulary is more effective when output-based tasks including fill-in task with a higher involvement index are employed.

It was only in the delayed post-test of the third phase when students receiving input-based tasks outperformed students receiving output-based tasks. This finding is in line with Laufer (2003) who found that sentence completion and dictionary consultation tasks were more effective than the sentence-writing task in improving target vocabulary retention. Similarly, Fourse’s (2006) study indicated that cloze-exercises were more effective than sentence-writing exercises for vocabulary learning and retention.

6. Conclusion
Involvement load and task type are two crucial factors in vocabulary learning and retention. In fact, vocabulary learning tasks with higher involvement load index are more cognitively demanding for learners and facilitate their learning. In addition, task type is an essential factor in vocabulary learning and retention; more specifically, output-based tasks are generally more facilitative and effective in vocabulary learning. Therefore, teachers should design a variety of output-based tasks with higher involvement loads to improve students’ vocabulary learning processes.
Table 6. Results of independent samples t-test for input-based and output-based groups in vocabulary post-test with different involvement load (output-based tasks with higher involvement load)

| Post-test phase 3 | Means | SD  | Levene's Test for Equality of Variances | t-test for Equality of Means |
|-------------------|-------|-----|----------------------------------------|-----------------------------|
|                   |       |     | F | Sig. | T  | Df | Sig. (2-tailed) | Mean Difference | Std. Error Difference |
| Input-based       | 16.05 | 1.19| 52 | .47  | -6.54 | 38 | .000 | -2.55 | 38 |
| Output-based      | 18.60 | 1.27|    |      |     |    |      |        |     |

According to Table 6, there is a statistically significant difference between the groups receiving output-based tasks (M = 18.60, SD = 1.27) and input-based tasks (M = 16.05, SD = 1.19), (t(38) = -6.54, p = .000, d = 2.07) in the post-test.
Table 7. Results of independent sample t-test for input-based and output-based groups in vocabulary delayed post-test with different involvement load (output-based tasks with higher involvement load)

| Post-test phase 3 | Means  | SD    | Levene’s Test for Equality of Variances | t-test for Equality of Means |
|-------------------|--------|-------|----------------------------------------|-----------------------------|
|                   | F      | Sig.  | T           | Df  | Sig. (2-tailed) | Mean Difference | Std. Error Difference |
| Input-based       | 15.05  | 1.66  | 1.45        | .23 | -6.23          | .000            | -3.00                 | .48          |
| Output-based      | 11.05  | 1.36  |             |     |                |                 |                      |             |

As presented in Table 7, there is a statistically significant difference between the groups receiving output-based tasks ($M = 11.05$, $SD = 1.36$) and input-based tasks ($M = 15.05$, $SD = 1.66$), ($t(38) = -6.23$, $p = .000$) in the delayed post-test.
The present study was limited in a number of ways. This study assessed the incidental vocabulary development of learners in an EFL context, where the language input outside the classroom is very limited. Future research could be conducted in ESL contexts where students are exposed to English outside the classroom context. Also, the current study employed an equally limited number of vocabulary items in output- and input-based tasks. It is suggested that future studies examine a wider range of vocabularies during a longer period of time in order to gain a clearer profile of word knowledge improvement.

Funding
The authors received no direct funding for this research.

Author details
Shiva Kaivanpanah
E-mail: shkaivan@ut.ac.ir
Seyed Mohammad Alavi
E-mail: smalavi@ut.ac.ir
Afsaneh Rovandpour
E-mail: afsanehrovandpour2018@gmail.com

1 University of Tehran, Tehran, Iran.
2 English Department, University of Tehran, Kish International Campus, Tehran, Iran.

Citation information
Cite this article as: The effect of input-based and output-based tasks with different and identical involvement loads on Iranian EFL learners’ incidental vocabulary learning. Shiva Kaivanpanah, Seyed Mohammad Alavi & Afsaneh Rovandpour, Cogent Psychology (2020), 7: 1731223.

References
Alavinia, P., & Rahimi, H. (2019). Task type effects and task involvement load on vocabulary learning of EFL learners. International Journal of Instruction, 12(1), 1501–1516. doi:10.29333/iji.2019.12196a
Al-Hadlaq, M. (2003). Retention of words learned incidentally by Saudi EFL learners through working on vocabulary learning tasks constructed to activate varying depth of processing (Unpublished Doctorate’s dissertation). Ball State University, Indiana.
Allen, L. (2000). Form-meaning connections and the French causative: An experiment in input processing. Studies in Second Language Acquisition, 22, 69–84. doi:10.1017/S0272263100001030
Bao, G. (2015). Task type effects on English as a foreign language learners’ acquisition of receptive and productive vocabulary knowledge. System, 53, 84–95. doi:10.1016/j.system.2015.07.006
Davis, J. N. (1989). Facilitating effects of marginal glosses on foreign language reading. The Modern Language Journal, 73(1), 41–48.
DeKeyser, R. M., & Sokalski, K. J. (1996). The differential role of comprehension and production practice. Language Learning, 46, 613–642. doi:10.1111/long.1996.46.issue-4
Ellis, R. (1999). Modified oral input and the acquisition of word meanings. Applied Linguistics, 16, 409–441. doi:10.1093/applin/16.4.409
Erlam, R. (2009). The elicited oral imitation test as a measure of implicit knowledge. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philip & H. Reinders, Implicit and explicit knowledge in second language learning, testing and teaching (pp. 65–93).
Folse, K. S. (2006). The effect of type of written exercise on L2 vocabulary retention. TESOL Journal, 40(1), 273–293. doi:10.2307/4264523
Fraser, C. A. (1999). Lexical processing strategy use and vocabulary learning through reading. Studies in Second Language Acquisition, 21, 225–241.
Hu, H., & Nassaji, H. (2016). Effective vocabulary learning tasks: Involvement load hypothesis versus technique feature analysis. System, 56, 2839. doi:10.1016/j.system.2015.11.001
Hulstijn, J. (2003). Incidental and intentional learning. In C.J. Doughty & M. H. Long (eds.), The handbook of second language acquisition (pp. 349–381). Blackwell Publishing.
Hulstijn, J. H. (2001). Intentional and incidental second language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity. In P. Robinson (Ed.), Cognition and second language instruction (pp. 258–286). Cambridge: Cambridge University Press.
Hunt, A., & Belgar, D. (2005). A framework for developing EFL reading vocabulary. Reading in a Foreign Language, 17, 23–59.
Jenkins, J. R., Stein, M. L., & Wysocki, K. (1984). Learning vocabulary through reading. American Educational Research Journal, 21(4), 767–787.
Kaivanpanah, S., & Miri, M. (2018). Inspecting task-induced involvement from the perspective of sociocultural theory. Journal of Teaching Language Skills, 37(1), 159–192.
Khonamri, F., & Hamzenia, Z. (2013). The role of task-induced involvement in vocabulary learning of Iranian language learners. English review. Journal of English Education, 1(2), 171–181.
Kim, Y. (2011). The role of task-induced involvement and learner proficiency in L2 vocabulary acquisition. Language Learning, 58(2), 285–325. doi:10.1111/j.1467-9922.2008.00442.x
Krahenbuhl, J. (2011). Second language acquisition and second language learning. Oxford: Pergamon Press.
Krahenbuhl, J. (1985). The input hypothesis: Issues and implications. New York: Longman.
Krahenbuhl, J. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. Modern Language Journal, 73, 440–464. doi:10.1111/j.1540-4781.1989.tb05325.x
Lauper, B. (2003). Vocabulary acquisition in a second language: Do learners really acquire most vocabulary by reading? Some empirical evidence. Canadian Modern Language Review, 59, 567–587. doi:10.3138/cmlr.59.4.567
Lauper, B. (2005). Focus on form in second language vocabulary learning. Euroslav Yearbook, 5, 223–250.
Lauper, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. Applied Linguistics, 22(1), 1–26.
Maftoon, P., & Sharifi, H. M. (2012). The relative effectiveness of input and output-oriented tasks with different involvement loads on the receptive and productive vocabulary knowledge of Iranian EFL learners. The Journal of Teaching Language Skills (JTLS), 4(2), 27–52.
Marefat, H., & Ghafari, S. (2009). Incorporating adjunct displays: A step toward facilitation of reading comprehension. Porta Linguarum, 11, 179–188.
Marmol, G. A., & Sanchez-Lafuente, A. L. (2013). The involvement load hypothesis: Its effect on vocabulary learning in primary education. RESLA, 26, 11–24.
McCarthy, M. (1990). Vocabulary. Oxford: Oxford University Press.

Nagata, N. (1998). The relative effectiveness of production and comprehension practice in second language acquisition. Computer Assisted Language Learning, 11 (2), 153–177. doi:10.1076/call.11.2.153.5686

Nagy, W., Herman, P., & Anderson, R. (1985). Learning words from context. Reading Research Quarterly, 20 (2), 233–253. doi:10.2307/747758

Nation. (2001). Learning vocabulary in another language. New York, NY: Cambridge University Press.

Rieder, A. (2005). A cognitive view of incidental vocabulary acquisition: From text meaning to word meaning. Vienna English Working Papers, 12, 53–71.

Robinson, D. H. (1994). Textbook selection: watch out for ‘inconsiderate’ texts. In (Eds.), Handbook of college teaching (pp. 415–422). Westport, CT: Greenwood.

Rott, S., & Williams, J. (2003). Making form-meaning connections while reading: A qualitative analysis of word processing. Reading in a Foreign Language, 15, 183–222.

Sarani, A., Negori, G. M., & Ghaviniat, M. (2013). The role of task type in L2 vocabulary acquisition: A case of involvement load hypothesis. Acta Scientiarum. Language and Culture, 35(4), 377–386. doi:10.4025/actascilangcult.v35i4.21135

Schmidt, R. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. AILA Review, 11, 11–26.

Soleimani, H., & Rahmani, M. (2015). Visiting involvement load hypothesis and vocabulary acquisition in similar task types. Theory and Practice in Language Studies, 5(9), 1883–1889. doi:10.17507/tpils.0509.16

Soleimania, H., Rahmanian, R., & Sajedia, S. (2014). A revisit to vocabulary acquisition in involvement load hypothesis. Procedia - Social and Behavioral Sciences, 192, 388–397. doi:10.1016/j.sbspro.2015.06.055

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Goss & C. Madden (Eds.), Input in second language acquisition (pp. 235–253). Rowley, MA: Newbury House.

Tahmasbi, M., & Farvardin, M. T. (2017). Probing the effects of task types on EFL learners' receptive and productive vocabulary knowledge. The case of involvement load hypothesis. SAGE Open, 7(3), 1–10. doi:10.1177/2158244017730596

Tokimoto, M. (2007). The effects of input-based tasks on the development of learners' pragmatic proficiency. Applied Linguistics, 28(1), 1–25.

Toth, P. D. (2006). Processing instruction and a role for output in second language acquisition. Language Learning, 56(2), 319385. doi:10.1111/j.0023-8333.2006.00349.x

Van Patten, B., & Sanz, C. (1995). From input to output: Processing instruction and communicative tasks. In F. R. Eckman, D. Highland, P. W. Lee, J. Mileham, & R. R. Weber (Eds.), Second language acquisition theory and pedagogy (pp. 169–185). Mahwah, NJ: Lawrence Erlbaum.

Walsh, M. I. (2009). The involvement load hypothesis applied to high school learners in Japan: Measuring the effects of evaluation (Unpublished MA thesis). Birmingham University, Edgbaston, UK.

Xu, J. (2009). An experimental study on the effects of different reading tasks on L2 vocabulary acquisition. English Language Teaching, 2, 69–76. doi:10.5539/elt.v2n3p69

© 2020 The Author(s). This open access article is distributed under a Creative Commons Attribution (CC-BY) 4.0 license.

You are free to:
Share — copy and redistribute the material in any medium or format.
Adapt — remix, transform, and build upon the material for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:
Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made.
No additional restrictions
You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.