Manipulation of the Involvement Load of L2 Reading Tasks: A Useful Heuristic for Enhanced L2 Vocabulary Development

Ehsan Namaziandost1, Murad Hassan Mohammed Sawalmeh2, Shouket Ahmad Tilwani3, Meisam Ziafar4, Arin Arianti5, Ronald M. Hernández6, Oleg Anatolevich Razzhivin7, Yolvi Ocaña-Fernández8, Doris Fuster-Guillén8, and Jessica Palacios Garay8

Abstract
Ensuring second language (L2) learners have an adequate breadth and depth of L2 vocabulary knowledge is a key pedagogical objective in L2 learning contexts. For this reason, establishing guiding principles that successfully enhance the efficacy of L2 vocabulary knowledge development is of strong importance. The current study investigated the value of applying principles from the Involvement Load Hypothesis (ILH) as part of a reading comprehension task among 40 intermediate English as a foreign language (EFL) students. Half of the group undertook a high involvement reading task, whereas the other half undertook a low involvement reading task. After the reading task, an unannounced Vocabulary Knowledge Scale test was administered to measure incidental vocabulary gains. Results showed the high involvement group remarkably outflanked the low involvement groups in terms of the target words learned from the reading task. A delayed post-test indicated that the retention of target word knowledge was more robust among the high involvement group, but that this difference did not maintain a level of statistical significance after 2 weeks. We conclude with suggestions about how EFL/ESL instructors can apply the principles of the ILH in efforts to systematically enhance learners’ L2 vocabulary knowledge.

Keywords
ILH, L2 reading tasks skill, vocabulary knowledge

Introduction
The importance of second language vocabulary is so self-axiomatic that it makes learning and teaching a must. To demonstrate the importance, Thornbury (2002, p. 13) claims that “without grammar very little can be conveyed, nothing can be conveyed without vocabulary.” All languages have been composed of words and they first are given birth to as words. Vocabularies never stop being coined in languages and consequently, the process of acquisition is a non-stop process. Even in their native language, language users are always grappling with the continuity of coinage, acquisition, and learning (Anova et al., 2015; Thornbury, 2002).

When it comes to learning a second language, the importance doubles. The language learner perceives himself or herself as being enveloped by a massive pile of unknown lexemes and their many features to learn. It is self-evident to both students and instructors that mastering a foreign language necessitates learning a plethora of new terms (Alqahtani, 2015). That’s why they are so apprehensive when faced with the tasks (Laufer & Hulstijn, 2001; Namaziandost, Rahimi Esfahani et al., 2019). As a result, it appears paradoxical that research into vocabulary acquisition is not nearly as extensive as research into other aspects of second language learning (Khabiri & Charmgar, 2012; Nasri et al., 2018; Nation, 2002).

There has recently been a surge in interest in language learning-related topics. Authored and edited publications devoted solely to vocabulary are evidence of the rebirth of interest in L2 vocabulary (Bogaards & Lafer, 2004; Fhonna, 2014; Hosseini et al., 2017). Much endeavor has been

1Islamic Azad University, Shahrekord Branch, Iran
2Dhofar University, Salalah, Oman
3Prince Sattam Bin Abdulaziz University, Alkhair, Saudi Arabia
4Islamic Azad University, Ahvaz Branch, Iran
5Universitas Veteran Bangun Nusantara Sukoharjo, Indonesia
6Universidad de San Martin de Porres, Lima, Perú
7Elabuga Institute of Kazan Federal University, Moskva, Russian Federation
8Universidad Nacional Mayor de San Marcos, Lima, Perú

Corresponding Author:
Ehsan Namaziandost, Department of English, Islamic Azad University, Shahrekord Branch, Shahrekord, Ahvaz 6176568964, Iran.
Email: e.namazi75@yahoo.com

Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (https://creativecommons.org/licenses/by/4.0/) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).
devoted in determining how L2 vocabulary can be learnt under various learning situations, as well as what factors impact the efficiency and patterns of L2 vocabulary learning (Azadi et al., 2018; Haratmeh, 2012; Jiang, 2002). The relation between vocabulary knowledge and reading proficiency has been an ongoing subject in the field of second language learning. To summarize, the problem has been concerned with the connection between vocabulary development and incidental reading. According to Chen and Truscott (2010), it is widely acknowledged in vocabulary acquisition researches that learning words happens incidentally during reading and that this type of learning is important. Research on both first and second language development supports the idea that words are mostly learnt naturally when language learners attempt to discern the meaning of new words encountered in reading and listening tasks (Alqahtani, 2015). Such learning has been labeled “incidental” as it “occurs while learners focus on anything other than word acquisition itself.” (Paribakht & Wesche, 1999, p. 16).

**Literature Review**

**Knowing a Word**

Words are not distinct subsets of languages, but rather components of intertwined and interconnected systems. As a result, numerous elements and dimensions of word knowledge are necessary for learners to utilize words effectively and accurately (Nation, 2001; Van Polen, 2014). Thus, one must understand what is meant by knowing.

Anderson and Freebody (1981) presented another paradigm that vocabulary learning researchers find beneficial. “This is a sort of distinction between breadth and depth of knowledge” (p. 37). The amount of words a person knows refer to his/her breadth of knowledge. Furthermore, depth of knowledge refers to a learner’s understanding of numerous elements of a single word. The concept of depth of vocabulary knowledge can allude to the relationships between words and includes knowledge of word association, collocation, and colligation (Anderson & Freebody, 1981; Namaziandost, Nasri et al., 2019).

According to Xu (2010), “the intricacy of word knowledge is never entirely comprehended by simple two-fold classifications such as receptive and productive, or breadth and depth” (p. 79). Nation (2001) has developed a more thorough and accurate word knowledge framework. He categorizes word knowledge into four classifications: form awareness, meaning awareness, and knowledge of use. Each category is further subdivided, with both expressive and receptive characteristics. Form knowledge encompasses both spoken and written patterns, as well as word components (Nasri et al., 2019; Nation, 2001). According to Xu (2010), “knowledge of meaning can also be divided into form and meaning, concepts, referents, and associations. Knowledge of use covers grammatical functions, collocations, and constraints on use” (p. 61).

**ILH and Reading Skill**

Depth of processing and elaborative learning are conceptualized by the ILH in terms of three primary task components: need, search, and evaluation (Laufer & Hulstijn, 2001). The strength of each of the three components is different. For instance, “need” is thought to be either moderate or strong. When a teacher imposes a need on a student, it is regarded moderate (e.g., The teacher wants the learner to find the meaning of a word). However, whether the students are intrinsically driven or impose it on themselves, the need is strong (e.g., the need to look up the meaning of a word in a dictionary when reading a text). If the definitions are supplied in the margins, there is no need for a search. Depending on whether it is receptive retrieval or productive retrieval, search might be moderate or strong (Nation & Webb, 2011). If the student needs to hunt for or recover the meaning of a word, search is moderate, while search is strong if the learner needs to seek the word form. If the student has to compare the precise meaning of a term with various meanings, the evaluation is moderate. If there is a requirement to determine if the meaning of a word fits a certain linguistic context, evaluation is strong. According to the ILH, the extent to which a vocabulary task aids L2 learners in acquiring new target words is determined by how much the task encourages each of the aforementioned involvement load (IL) components. It implies that the higher the participation load in a particular task, the greater the vocabulary learning and retention would be.

Laufer and Hulstijn (2001) presented the following combination of different tasks and how their ILs varied. One task requires the student to construct sentences using a bunch of different words, the definitions of which are provided by the teacher. They contended that this activity needs no search because the meanings are supplied. However, it generates a moderate need and a strong assessment since the learner must assess the appropriateness of the terms in context. They anticipated that the task had an involvement index of 3 [0 (search) þ 1 (need) þ 2 (evaluation)] in terms of overall IL. The second task requires the student to read a text and respond comprehension questions, with the definitions of the terms supplied in the margins. The assignment here does not need any assessment or searches, but it does require the student to look at the glosses, which is a moderate requirement. They claimed that this task had an overall involvement load index of 1 [0 (search) 1 (need) 0 (evaluation)]. Task one, as per the investigators, would be more successful than task two for vocabulary learning.

The main argument of ILH is that retention of unknown words is typically dependent on one’s level of involvement in the processing of those words. The concept of involvement may be experimentally examined by devising a variety of tasks with varying degrees of need, search, and assessment. Tasks with varying levels of involvement, for instance, can be assigned to different groups of participants. After
accomplishing the assignment, the outcomes may be examined and compared to see if there is a connection among task IL and word recall.

According to Paribakht and Wesche (1997), after analyzing researches on the association between reading and vocabulary learning, “these studies all point to the role for reading processes in vocabulary acquisition, but an unexpected one, and not always the most effective.” They explain that the uncertainty of what vocabularies would most likely be learned and the amount to which learning would occur makes the situation murkier. As a result, vocabulary cannot be left to fend for itself, and learners should not be permitted to pick up vocabularies at their own. To compensate for the foregoing pitfalls, systematic development in vocabulary acquisition should be the target” (Van Polen, 2014, p. 36). Although numerous researches have been conducted on incidental learning and vocabulary acquisition, there has been no effective way to present a theoretical foundation for these experiments, that is, how activities used for incidental learning may be modified to improve vocabulary acquisition (Soleimani et al., 2015). Keating (2008) observes that one of the primary obsessions of scholars and teachers is recognizing the tasks that provide opportunities for learners to learn words.

Driven by studies in the literature, Laufer and Hulstijn (2001) propose a framework called task-induced IL. Investigating the tasks employed and reviewing the literature so far, Laufer and Hulstijn claim that they have determined the elements of incidental tasks that promote the type of complex processing required for learning.

In order to trigger theoretical and empirical study in the domain of L2 vocabulary, Laufer and Hulstijn (2001) introduce a new construct termed Involvement. They feature the construct “as composed of three motivational and cognitive dimensions: need, search, and evaluation” (Laufer & Hulstijn, 2001, p. 75). They define the Need element as the “drive to comply with the task requirements, whereby task requirements can be either externally imposed or self-imposed.” (p. 14). In addition to the definition, they characterize Need as being either “moderate” or “strong.” By the former it is meant the need is imposed on the learner externally. In the case of the latter, need is imposed internally by the learner (Ziafar & Namaziandost, 2019).

Contrariwise, when the same learner decides to use a vocabulary that he needs to use in his performance, the Need is considered to be self-imposed, that is, strong. Thus, Laufer and Hulstijn (2001) maintain that the need is strong because this is something self-imposed by the learner him/herself. What was discussed so far, actually formed the motivational part of the construct of involvement. The cognitive components of the construct are the other two to be addressed. Search is the first one. Search is defined as “the attempt to find the meaning of an unknown L2 word or trying to find the L2 word for expressing a concept (e.g., trying to find the L2 translation of an L1 word) by consulting a dictionary or another authority (e.g., a teacher)” (Laufer & Hulstijn, 2001, p. 14). Evaluation is the third element of the construct of Involvement. Laufer and Hulstijn (2001) define it “a comparison of a given word with other words, a specific meaning of a word with its other meanings, or combining the word with other words to assess whether a word (i.e., a form-meaning pair) does or does not fit its context” (Laufer & Hulstijn, 2001, p. 14). Moreover, according to Van Polen (2014), “the concept of involvement can be operationalized by advising tasks with varying degrees of need, search, and evaluation and therefore can be submitted to the empirical investigation” (p. 34).

Prior Researches on ILH

Keating (2008) showed that “word learning and retention in a second language are contingent upon a task’s IL (i.e., the amount of need, search, and evaluation it imposes),” (p. 17) as suggested by Laufer and Hulstijn (2001). To achieve this goal, 79 beginning Spanish learners finished one of three vocabulary learning tasks that differed in the percentage of involvement (i.e., mental endeavor) they required: reading comprehension (no endeavor), reading comprehension and as well as target word supplementation (moderate endeavor), and sentence writing (strong endeavor).

Hu and Nassaji (2016) carried out an empirical research on ILH and TFA regarding their predictability for efficient L2 vocabulary learning tasks. They assigned 96 adult EFL learners to four groups and were asked to learn the meaning of 14 unfamiliar words. Each group practiced through one of the vocabulary tasks from the two different frameworks. Finally, they discovered that the TFA group outperformed the ILH group and that the former had a better explanatory power in predicting vocabulary learning gains.

Similarly, Reinhard and Sporer (2008) tried to determine the utility of dual-process frameworks in comprehending the trustworthiness attributing mechanism. Only strong task engagement and high cognitive capacity, according to the premises of dual-process frameworks, result in intense analysis of verbal and nonverbal information while generating trustworthiness judgments. People primarily utilize nonverbal information for trustworthiness identification when task involvement and/or cognitive ability are minimal. Their research found that individuals with high cognitive ability, as opposed to those with poor cognitive capacity, used linguistic data to develop trustworthiness.

Keyvanfar and Badrughi (2011) explored whether word learning and maintenance in a subsequent language were dependent upon an errand’s contribution load (i.e., the measure of need, search, and assessment it forces), as suggested by Laufer and Hulstijn (2001). According to the findings of their study, learners learned more from sentence formation that involved the contrast of fresh words with terms already known for the goal of output. Finally, they found that the assessment element may be the most important factor in
task-induced IL. In another study, Kim (2008) conducted two studies to demonstrate the role of ILH in L2 vocabulary learning. Experiment 1 was designed to look at how varying levels of task-induced engagement influenced L2 learners’ preliminary learning and retention of target words. Kim (2008) investigated whether two activities (writing composition and writing sentences) purporting to have the same level of task-induced involvement (involvement index = 3) would have equal impacts on the preliminary acquisition and retention of target words in his experiment 2. His study’s results indicated that the two exercises were equally successful in boosting both preliminary acquisition and retention of new words.

In another examination, Asadzadeh Maleki (2012) tried to see if word learning and maintenance in a subsequent language is dependent upon a task ILH, for example, the measure of need, search, and assessment. Based on findings obtained through administering the immediate and delayed vocabulary retention post-tests, she came to the conclusion that there was a substantial difference in retention effects between the three activities, verifying the validity of the ILH and confirming that tasks with a greater participation load lead to greater retention influences. Tahmasbi and Farvardin (2017) also investigated the effects of task types on EFL learners’ receptive and productive vocabulary knowledge. They reported that all output tasks outperformed the control task in terms of improving respondents’ receptive and productive vocabulary knowledge.

Furthermore, Karalik and Merc (2016) conducted a research to investigate the effect of task-induced IL on incidental vocabulary growth and retention. Following completion of the tasks, a post-test and subsequently a delayed post-test were given. The findings of this study indicated that activities with a greater level of IL resulted in greater vocabulary development and retention. On the delayed post-test, the only important differences were between the retelling by searching and fill-in groups, indicating the importance of the task-induced ILH.

Xu and Yan (2018) investigated the effects of task-induced engagement and different task categories on incidental vocabulary learning. Reading tasks with higher ILs lead to a more positive preliminary and delayed word acquisitions than activities with less ILs, according to the findings. The outcomes also revealed that there is no notable change in the influence of the interpretative task and the production task on immediate word learning, but the production task appears to be far stronger to the interpretative task on word recall. In another study, Gohar et al. (2018) attempted to compare the estimated predictive power of two techniques namely ILH and TFA in vocabulary learning. Three vocabulary tasks of sentence making, composition, and reading comprehension were assigned to 90 high proficiency EFL students. The ILH was shown not to be a strong predictor, and TFA was a key determinant of pretest to posttest score improvement but not in during-task activity.

Kang and Shin (2019) conducted a study to explore the effects of task categories and task IL on the vocabulary retention of Korean EFL undergraduate students. To study the interaction between engagement indices and task varieties, three kinds of productive word-focused activities (gap-filling using a dictionary, creating original phrases, and gap-filling using word transformation) were employed. The results showed substantial interaction effects of task types on retention tests, proficiency levels on retention tests, and task type proficiency levels on retentions. Furthermore, Alarjani (2020) conducted a research to evaluate the influence of ILH task-based learning on Saudi university students’ recall of meaning. The findings of this research revealed that the involvement-loaded task not only developed the participants’ knowledge of the target words but also assisted in meaning retention.

Few researches have been done in the Iranian context on the influence of task varieties and IL on EFL students’ vocabulary learning. Yaqubi et al. (2012), for example, randomly assigned 60 EFL students to three groups: First group accomplished an input-oriented assignment with a participation load of 3, Second group accomplished the very same type of task but with a participation load of 2, and Third group accomplished an output-oriented task with a participation load of 3. The findings contradicted the ILH’s forecast, indicating that Task 2 outperformed Task 1, which had a higher index. Furthermore, notwithstanding their index equivalence, students who undertook Task 3 performed considerably better than those who finished Task 1. Similarly, Soleimani and Rahmanian (2015 randomly assigned 33 Iranian EFL IL load of 2), and sentence composition (IL of 3). The sentence writing exercise performed much better than the other two activities, according to the findings. The findings corroborated the ILH hypotheses. Nonetheless, this research has numerous drawbacks, including no word type control, a small sample size, no evaluation of productive knowledge, and inadequate information for power analysis.

All in all, investigating the impact of applying principles from the ILH as part of a reading comprehension task seems to be quite telling when it comes to the prominence of the aforementioned determining factors in language learning efficiently. Further, the findings of the current study may prove revealing in language teachers’ and learners’ search for an approach that satisfies language learners’ needs to be more task-savvy and thus more effective in fulfilling language learning goals. Thus, the current study tried to lessen the limitations of the previous researches such as having a lack of IL. Moreover, this study focused on reading comprehension tasks that have not been already considered in prior studies. Finally, rare researches have been accomplished to examine vocabulary retention through task embedding and incidental learning. As a result, this study sought to prevail over the shortcomings in previous researches.
**Research Questions**

The following research questions were addressed in this study:

1. Does different levels of involvement load (i.e., high involvement vs. low involvement) have any significant effect on the participants’ vocabulary learning from the reading task?
2. Is there any significant difference between the participants’ vocabulary retention through task embedding and incidental learning?

**Methodology**

**Participants**

This study included 40 male and female students who were chosen from two intermediate intact BA classes majoring in translation studies at Payam Noor University of Ahvaz, Iran. Then they were divided into two equal experimental groups namely the high involvement load group (HILG) and the lack of involvement load group (LILG). The participants’ age ranged from 18 to 25. They were studying their seventh semester, out of eight. Until the sixth semester, they had passed about 80 units in total; 70 units concerned courses related to translation.

**Materials and Instruments**

The study drew on two reading tasks with different demands on test-takers. In the first reading task, the unknown words were glossed for the test takers but the follow-up comprehension questions of the task were designed in a way to be answered without reference to the target words (irrelevant to the task requirement). The way to do it was to direct the questions to the sections in the reading that the target words were not present or they were irrelevant to the answers looked for. According to Laufer and Hulstijn (2001), “this type of reading task triggers neither Need (being irrelevant to the task) nor Search for the meaning (because of the glossary provided) and no evaluation; in other words, the IL for the task is 0 (-+ - =0)” (p. 65). The minus symbol denotes the lack of the three elements of the task. As a result, the task’s load was at its lowest feasible degree, and as Laufer and Hulstijn (2001) assert, the task generates the least potential of inadvertent learning of target languages.

In the second task, again use was made of the same reading task but with some modifications. This time, the target words were omitted from the text. These words were recorded at the base of the content with their clarifications and the errand required the students to fill the content holes with the right words from the rundown (the undertaking can be made increasingly complex by adding two additional words to the ones that don’t fit the content by any stretch of the imagination). The fill-in assignment instigates a moderate Need (+), no Search (the words are clarified), and a moderate Evaluation (+) since every one of the words in the rundown must be assessed against one another and the setting of the gap. If we assign one point to each presentation of the three components in their moderate version and two points to their strong versions, so the load number for the first and second task was 0 and 2 respectively. Therefore, based on Laufer and Hulstijn (2001), it is predicted that the second task leads to more gains in vocabularies in comparison to the first one in that the latter induces more IL.

The two reading texts were controlled for their difficulty level. Attempts were made to have a mediocre difficulty level. Every reading comprehension test’s level of difficulty was tailored to the level of language competence of the participants who were scheduled to take the test at the end of the teaching period. The passage difficulty levels were computed using the Fog Index Formula and were found to be fairly consistent with the Michigan test. The level of difficulty of a text, as per Fulcher (2005), should be appropriate for the readers. Suitable learning texts are selected in a variety of ways. The underlying premise in all formulas is that the more polysyllabic words in a sentence, the more difficult it is, and the fewer sentences in a paragraph, the more difficult the paragraph. The underlying premise in all formulas is that the more polysyllabic words in a sentence, the more difficult it is, and the fewer sentences in a paragraph, the more difficult the paragraph. The Gunning Fog Index is one of these formulas. (Gunning, 1952) (http://gunning-fog-index.com/). These texts were shown to be equal by the Fog index of readability (Farhadi et al., 1994). Using the Fog index, the readability levels of the two passages were computed to be 18 and 21. The average readability was 19 and the standard deviation was 2.01. The Fog index of readability of the texts selected for this study was calculated to be 19 that gives it an appropriate level of difficulty because it was within the range of 19 ± 1.99.

**Target Vocabularies**

In total, nine vocabularies were investigated in the study. The frequency of the vocabulary was the criteria used to choose them, with the target words having a low frequency. This was done on purpose to ensure that respondents had not previously been subjected to the vocabulary. As a result, these vocabularies were chosen based on Nation’s (1994) taxonomy of vocabularies. In addition, to account for inherent difficulty (the number of syllables may impact vocabulary learning), all target vocabularies were chosen with the same number of syllables. The researchers intended to use vocabulary with two syllables in the research. Furthermore, to illustrate the variety of vocabularies, nouns, verbs, and adjectives were all given an equal proportion of 3 out of 9. The current study did not aim to find functional terms.

**Instrument: Vocabulary Knowledge Scale (VKS)**

The VKS was designed in the framework of a university-based study on the vocabulary growth of ESL students (Paribakht & Wesche, 1993). This test catches in a moderately effective
manner certain phases in the underlying advancement of center learning of given words. Paribakht and Wesche (1996) declare that the VKS ought to be seen as a down-to-earth instrument for use in investigations of the underlying acknowledgment and utilization of new words. The VKS instrument utilizes a scale joining self-report and execution things to inspire both self-saw and exhibited learning of explicit words in composed structure. The scale appraisals extend from complete newness, through acknowledgment of the word and some thought of its importance, to the capacity to utilize the word with syntactic and semantic exactness in a sentence. Our essential objective in building up the VKS has been to catch starting stages or levels in word discovering that are liable to exact self-report or productive exhibition, and that is exact enough to reflect gains during a moderately short instructional period. Based on the levels of knowledge proved by the performance of the learners, Paribakht and Wesche (1996, p. 98) suggested the scoring system for the levels of knowledge as follows:

| Self-report categories | possible scores | meaning of scores |
|------------------------|-----------------|-------------------|
| I.                     | 1               | *The word is not familiar at all. |
| II.                    | 2               | *The word is familiar but its meaning is not known. |
| III.                   | 3               | *A correct synonym or translation is given. |
| IV.                    | 4               | *The word is used with semantic appropriateness In a sentence. |
| V.                     | 5               | *The word is used with semantic appropriateness and grammatical accuracy in a sentence. |

**Figure 1.** Self-report categories and the corresponding scores.

The feasible scores for a word on this tool and their connection to the self-report classifications are given in Figure 1. As it is delineated, wrong reactions in self-report classifications III, IV, or V will prompt a score of 2. A score of 3 shows that a proper equivalent word or interpretation has been given for self-report classifications III or IV. A score of 4 is given if the word is utilized in a sentence exhibiting the student’s learning of its importance in that unique situation however with erroneous grammar (e.g., a target noun utilized as a verb: “This famous player announced his retire”), or an erroneously conjugated or determined structure is given (e.g., “losed” for “lost”). A score of 5 reflects both semantically and syntactically right utilization of the objective word, regardless of whether different pieces of the sentence contain errors.

**Data Collection Procedures**

The whole procedure of the study pivoted around two stages. In the first stage, participants of the first group were given the target words and the VKS to check that they did not know the vocabulary. Those familiar with vocabularies were excluded from the study. The first reading comprehension exam (E1) was then administered to participants (E1). VKS1 was utilized promptly after they were completed to assess if any learning had occurred in the target vocabulary. At the same time, the second experimental group (E2) was given the second reading comprehension test at a different location. Following the completion of the exam, participants were administered the VKS1 to assess any vocabulary learning in the target vocabularies as a consequence of the intervention. We had the treatment’s preliminary findings at this point. We sought to assess vocabulary retention following 2 weeks of immediate and incidental application of the VKS2 (a revamped version of VKS2). Students in the two groups (E1 and E2) did not take reading comprehension exams this time. They were only provided the list of vocabulary they had been subjected to during the reading comprehension tasks 2 weeks prior. They were required to respond the VKS2 questions.

**Data Analysis Procedures**

For answering the research questions, Independent Samples t-tests were run to find out the effectiveness of high and low ILs on EFL learners’ vocabulary learning (VKS1) learning through reading tasks and also to compare the performance of both groups (i.e., high IL and IL groups) on delayed vocab test (VKS2).

**Results**

The first question of this study aimed to check “Does different levels of IL (i.e., high involvement vs. low involvement) have any significant effect on the participants’ vocabulary learning learned from the reading task?” To give a logical answer to this question, both groups were given the relevant reading passages before the treatment and the VKS was
result is also evident in Figure 3. Between LILG's performance on VKS 1 and VKS 2, this difference is less than one in reality. Overall, based on this graph, it is possible to conclude that the mean scores of both groups achieved almost similar levels in the retention test, albeit the group with a high level of IL was somewhat higher.

### Discussion and Conclusion

The results of the current study are in agreement with the results of several literature studies. Several prior studies have concluded that higher levels of IL on the part of the learners have resulted in better learning. The present study suggested that a higher participation load leads to better learning of vocabulary in reading activities so that it is compatible with the inquiry line. The findings of Keating (2008), Reinhard and Sporer (2008), Haratmeh (2012), Keyvanfar and Badraghi (2011), and Lu and Huang (2009) were consistent with the findings of this study. They all concluded that IL can play a significant role in improving the acquisition of vocabulary. All these studies, together with the present one, also imply that precise utilization of this IL is fundamental and requires the utmost consideration for the target students, level of difficulty in the task, environment in the classroom, level of proficiency, etc.

More importantly, Keating (2008) showed that the load involvement of a task determines learning and retention of words in a second language (i.e., the amount of need, search, and evaluation it imposes), as Laufer and Hulstijn (2001) suggested. Similarly, Yaqubi et al. (2012) affirmed the advantages of utilizing high IL tasks by suggesting that teachers and language learners may use tasks with higher involvement indexes regardless of their form to boost their acquisition of vocabulary.

In addition, Reinhard and Sporer (2008) supported this finding by claiming that “participants with high cognitive ability utilized verbal information to attribute their credibility” (p. 11). Likewise, Haratmeh (2012) supported one type of task by expressing that shared task-selection control results in higher task involvement. In other words, they advocated the implementation of tasks resulting in higher learning outcomes combined with intense straightforwardly committed learning attempts.

The results obtained in this study are also consistent with those found by Keyvanfar and Badraghi (2011). They confirmed the examination of tasks for vocabulary merging and proposed that the component of the evaluation might play a fundamental role in the task-induced load of involvement.

To sum up, the present study provided strong support for the IL construct of Laufer and Hulstijn (2001). Generally speaking, enhancing the general IL levels has shown to be

---

**Table 1. Students' Performance on VKS 1 (After Reading).**

| Group    | N  | Mean | SD  | SE Mean |
|----------|----|------|-----|---------|
| HILG     | 20 | 4.72 | 1.31| .29     |
| LILG     | 20 | 2.05 | 1.41| .31     |

*Note. LILG = lack of involvement load group; HILG = high level of involvement load group.*
influential in promoting vocabulary learning by tasks (Nasri & Biria, 2016). Moreover, controlling for student-related inadequacies, for example, dictionary use propensities, writing abilities, and attention span which were available in Haratmeh (2012) and Van Polen (2014) was discovered to be helpful for the segments of search and assessment to produce results. The present study, in a nutshell, provided potent support for the predictive value of the task-induced load hypothesis of involvement, which many studies could not achieve due to learner-related factors. It can be proposed that taking into account the particular characteristics of the learners in a special context, the IL construct can be exerted to design various tasks that are conducive to incidence.

Concerning the present results, it can be inferred that in the field of vocabulary acquisition, designing tasks with high ILs can serve as a new strategy and solution to current vocabulary acquisition problems. EFL teachers, instructors, and vocabulary trainers can use the findings to help them design reading tasks with appropriate levels of difficulty with a suitable level of involvement in vocabulary. In doing so, they can test vocabulary load levels to be integrated into the reading activities, and then give them to the participants. In that way, they will educate learners with better knowledge of vocabulary.

In addition, the present results may also be used by EFL curriculum designers to integrate activities with high levels of interaction loads. These can be included with reading tasks to make reading an activity that needs more involvement and attraction. Furthermore, it would possible for them to exert high load vocabulary involvement to other language skills. For example, listening, writing, and speaking skills practitioners can give their students vocabulary tasks with a higher IL. In this way, it is possible to better and more effectively acquire vocabulary which is obligatory parts L2 material.

Since the current research includes only two levels of IL, high level and lack of IL, future researches can take account of a higher number of IL. In this way, a variety of tasks may be discovered with different levels of participation IL. Consequences would be effective to vocabulary practitioners, and they would have an extensive range of tasks to utilize for a particular group of learners. This could lead to outlandish methods of vocabulary instruction and could make learning English more absorbing. A final suggestion for further studies might be implementing various delayed post-tests with different time intervals. The current study used a delayed posttest after a single time interval. Future studies can take account of posttest with different time intervals to see which time interval would serve as the best for the retention of vocabularies with high IL. Therefore, it is recommended to take posttest with various time intervals into account to see which time interval would be ideal for retaining vocabulary with high loads of involvement.

This study was not out of limitations, like any other study. First of all, this study included only 40 participants who were selected from a single university and did not have a large sample size. Because of this, generalization must be considered with caution. Secondly, only two reading tasks with two levels of IL (i.e., lack of IL and high level of IL) were used in this study as instructing materials. Other levels of IL were not considered. Besides, the third constraint involves participants sampling. If the participants were randomly selected from various universities, more reliable results would have been achieved. Furthermore, these results would have been more generalizable if participants were selected from amongst the learners of various EFL contexts. If so, consequences can be generalized to a larger group of Iranian EFL learners of a particular level of language proficiency.

Table 2. Independent Samples Test on VKS 1.

|                      | Levene’s test for equality of variances | t-Test for equality of mean |
|----------------------|----------------------------------------|-----------------------------|
|                      | F | Sig. | t | df | Sig. (two-tailed) |
| Equal variances assumed | 0.45 | .50 | 6.19 | 38 | .000 |
| Equal variances not assumed | 6.19 | 37.80 | .000 |

Table 3. Students’ Performance on VKS 2 (Retention).

| Group     | N  | Mean | SD  | SE  | Mean |
|-----------|----|------|-----|-----|------|
| High Involvement Load Group (HILG) | 20 | 2.72 | 1.18 | 0.26 | 0.29 |
| Lack of Involvement Load Group (LILG) | 20 | 2.35 | 1.31 | 0.29 | 0.29 |
tasks have been confirmed for better vocabulary acquisition. Excellent outcomes were achieved when the participants in the reading exercises were more familiar with the vocabulary. Secondly, it became apparent that the same beneficial and important results of tasks filled with higher involvement were not noticeable in the delayed posttest. Nevertheless, although these outcomes were not notable, it was revealed that higher IL participants performed better on the delayed posttest than the lack of IL group.

Declaration of Conflicting Interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD
Ehsan Namaziandost https://orcid.org/0000-0002-8393-2537

References
Alarjani, F. A. (2020). Involvement load hypothesis and the retention of word meaning among Saudi EFL learners. Asiatic: IIUM Journal of English Language and Literature, 14(1), 160–173.
Alqaahani, O. (2015). The importance of vocabulary in language learning and how to be taught. International Journal of Teaching and Education, 3(3), 21–34. https://doi.org/10.52950/TE.2015.3.3.002
Alqahtani, M. (2015). The importance of vocabulary in language learning and how to be taught. International Journal of Teaching and Education, 3(3), 21–34. https://doi.org/10.20472/TE.2015.3.3.002
Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge (Technical report no. 136). Illinois University Publication.
Anova, C., Antoni, R., & Kasyulita, E. (2015). The correlation between students’ vocabulary mastery and speaking skill at fifth semester of English study program in Pasir Pengaraian University. Jurnal Ilmiah Mahasiswa FKIP Prodi Bahasa Inggris, 1(1), 1–9.
Asadzadeh Maleki, N. (2012). The effect of the involvement load hypothesis on improving Iranian EFL learners’ incidental vocabulary acquisition in listening comprehension classes. Australian Journal of Basic and Applied Sciences, 6(9), 119–128.
Azadi, G., Biria, R., & Nasri, M. (2018). Operationalising the concept of mediation in L2 teacher education. Journal of Language Teaching and Research, 9(1), 132–140.
Bogaards, P., & Laufer, B. (2004). Introduction. In P. Bogaards & B. Laufer (Eds.), Vocabulary in a second language: Selection, acquisition, and testing (pp. vii–xiv). Benjamins.
Chen, C., & Truscott, J. (2010). The effects of repetition and L1 lexicalization on incidental vocabulary acquisition. Applied Linguistics, 31(5), 693–713.
Fhonna, R. (2014). The correlation between mastering vocabulary and speaking ability (case study at SMA 10 Fajar Harapan Banda Aceh). Visipena, 5(1), 49–58.
Fulcher, G. (2005). Text difficulty and accessibility: Reading formulae and expert judgment. System, 25(4), 497–513.
Gohar, M. J., Rahmanian, M., & Soleimani, H. (2018). Technique feature analysis or involvement load hypothesis: Estimating their predictive power in vocabulary learning. Journal of Psycholinguistic Research, 47(4), 859–869.
Gunning, R. (1952). The technique of clear writing. McGraw-Hill.
Haratmeh, M. S. (2012). Involvement load and task type in task effectiveness: Two aspects of vocabulary knowledge. International Journal of Academic Research, 4(8), 86–95. https://doi.org/10.7813/2075-4124.2012/4-4/b.13
Hosseini, E. Z., Nasri, M., & Afghari, A. (2017). Looking beyond teachers’ classroom behavior: Novice and experienced EFL teachers’ practice of pedagogical knowledge to improve learners’ motivational strategies. Journal of Applied Linguistics and Language Research, 4(8), 183–200.
Hu, H. C. M., & Nassaji, H. (2016). Effective vocabulary learning tasks: Involvement load hypothesis versus technique feature analysis. System, 56(2), 28–39. https://doi.org/10.1016/j.system.2015.11.001
Jiang, N. (2002). Form–meaning mapping in vocabulary acquisition in a second language. *Studies in Second Language Acquisition, 24*(4), 617–637.

Kang, H. R., & Shin, Y. (2019). The involvement load hypothesis and vocabulary retention: The interplay of task types and involvement index on L2 vocabulary learning. *Foreign Languages Education, 26*(4), 89–111. http://db.koreascholar.com/article.aspx?code=386525

Karalik, T., & Merc, A. (2016). The effect of task-induced involvement load on incidental vocabulary acquisition. *Mustafa Kemal University Journal of Graduate School of Social Sciences, 13*(35), 77–92. https://dergipark.org.tr/tr/download/article-file/226468

Keating, G. D. (2008). Task effectiveness and word learning in a second language: The involvement load hypothesis on trial. *Language Teaching Research, 12*(3), 365–386.

Keyvanfar, A., & Badraghi, A. H. (2011). Revisiting task-induced involvement load and vocabulary enhancement: Insights from the EFL setting of Iran. *Man & the Word/Zmogus ir zodis, 13*(3), 1–17.

Khabiri, M., & Charmgar, S. (2012). The comparative effect of non-task and provide-a-model preparation activities on EFL learners’ vocabulary retention. *International Journal of Applied Linguistics & English Literature, 1*(1), 113–122.

Kim, Y. (2008). The role of task-induced involvement and learner proficiency in L2 vocabulary acquisition. *Language Learning, 58*, 285–325.

Lafer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics, 22*(1), 1–26.

Lu, J., & Huang, J. (2009). An empirical study of the involvement load hypothesis in incidental vocabulary acquisition in EFL listening. *Polyglossia, 16*, 1–11.

Namaziandost, E., Nasri, M., & Ziafar, M. (2019). Comparing the impacts of various inputs (I+1 & 1-1) on pre-intermediate EFL learners’ reading comprehension and reading motivation: The case of Ahvazi learners. *Asian-Pacific Journal of Second and Foreign Language Education, 4*(1), 1–20. https://doi.org/10.1186/s40862-019-0079-1

Namaziandost, E., Rahimi Esfahani, F., & Nasri, M. (2019). Texts with various levels of hardness, reading comprehension and reading motivation: 1+1 versus 1-1. *ELT Forum: Journal of English Language Teaching, 8*(1), 60–77.

Nasri, M., & Biria, R. (2016). Integrating multiple and focused strategies for improving reading comprehension and L2 lexical development of Iranian intermediate EFL learners. *International Journal of Applied Linguistics & English Literature, 6*(1), 311–321.

Nasri, M., Biria, R., & Karimi, M. (2018). Projecting gender identity in argumentative written discourse. *International Journal of Applied Linguistics & English Literature, 7*(3), 201–205.

Nasri, M., Namaziandost, E., & Akbari, S. (2019). Impact of pictorial cues on speaking fluency and accuracy among Iranian pre-intermediate EFL learners. *International Journal of English Language and Literature Studies, 8*(3), 99–109.

Nation, I. S. P. (2001). *Learning vocabulary in another language.* Cambridge University Press.

Nation, I. S. P. (2002). Learning vocabulary in lexical sets: Dangers and guidelines. *TESOL Journal, 9*(2), 6–10.

Nation, P. (1994). *New ways in teaching vocabulary.* TESOL.

Nation, P., & Webb, S. (2011). *Researching and analyzing vocabulary.* Heinle.

Paribakht, S., & Wesche, M. (1996). Enhancing vocabulary acquisition through reading: A hierarchy of text-related exercise types. *Canadian Modern Language Review, 52*, 155–178.

Paribakht, T. S., & Wesche, M. (1993). The relationship between reading comprehension and second language development in a comprehension-based ESL program. *TESL Canada Journal, 11*, 9–29.

Paribakht, T. S., & Wesche, M. (1997). Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition. *Second Language Vocabulary Acquisition: A Rationale for Pedagogy, 5*(4), 174–200.

Paribakht, T. S., & Wesche, M. (1999). Reading and “incidental” L2 vocabulary acquisition. *Studies in Second Language Acquisition, 21*(2), 195–224.

Reinhard, M. A., & Sporer, S. L. (2008). Verbal and nonverbal behaviour as a basis for credibility attribution: The impact of task involvement and cognitive capacity. *Journal of Experimental Social Psychology, 44*(3), 477–488.

Soleimani, H., & Rahamanian, M. (2015). Vocabulary acquisition and task effectiveness in involvement load hypothesis: A case in Iran. *International Journal of Applied Linguistics & English Literature, 4*, 198–205.

Soleimani, H., Rahamanian, M., & Sajedi, K. (2015). A revisit to vocabulary acquisition in involvement load hypothesis. *Procedia – Social and Behavioral Sciences, 192*, 388–397. https://doi.org/10.1016/j.sbspro.2015.06.055

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. https://doi.org/10.1177/2158244017730596

Thornbury, S. (2002). *How to teach vocabulary.* Pearson Education Limited.

Van Polen, Z. (2014). The effects of task-induced involvement load and word exposure frequency on L2 incidental vocabulary learning through reading (Doctoral dissertation). Universiteit Van Amsterdam.

Xu, Q., & Yan, L. (2018). The effect of different reading tasks on Chinese senior high school English learners’ incidental vocabulary acquisition. *Journal of Advances in Education Research, 3*(4), 216–223. https://doi.org/10.22606/jaer.2018.34003

Xu, X. (2010). An empirical study on the effect of task on L2 incidental vocabulary acquisition through reading. *Asian Social Science, 6*(7), 126–131. https://doi.org/10.5539/ass.v6n7p126

Yaqub, B., Rayati, R. A., & Allemzade Gorgi, N. (2012). The involvement load hypothesis and vocabulary learning: The effect of task types and involvement index on L2 vocabulary acquisition. *Journal of Teaching Language Skills, 29*(1), 145–163.

Ziafar, M., & Namaziandost, E. (2019). *Linguistics, SLA and lexicon as the unit of language.* *International Journal of Linguistics, Literature and Translation, 2*(5), 245–250.