# Involvement Load Hypothesis and the Retention of Word Meaning Among Saudi EFL Learners

Faten Ahmed Alarjani[1]
King Saud University, Saudi Arabia

## Abstract

The Involvement Load Hypothesis (Laufer and Hulstijn) claims that in incidental learning situations, the retention of forms (words) and meaning depend on the manipulation of the cognitive and motivational variables within tasks. This study attempted to investigate the effect of task-based learning of the Involvement Load Hypothesis on Saudi university students' retention of meaning. The study examined two tasks developed based on the hypothesis and their effect on Saudi university students' retention of meaning. The participants in the study were female university students enrolled at Princes Nourah University in Riyadh, who were learners of English as a foreign language and spoke Arabic as their first language. The participants were randomly assigned to two groups: experimental group and control group. The experimental group received the writing composition task, while the control group received only the glossary of terms to read. The participants were pre-tested before the implementation of the task and post-tested one week later. This test examined whether the Involvement Load Hypothesis had affected the learners' retention of meaning or not. The pretest scores of the two groups were compared to examine whether they are compatible. An independent sample *t*-test was used to calculate equivalency between them. Also, the data was analysed quantitatively by using a paired sample *t*-test and an independent *t*-test to support the qualitative analysis.

## Keywords

Learning a language and successfully using it communicatively requires learning a large amount of vocabulary which is a fact long acknowledged by researchers in second language acquisition and foreign language instructors. An extensive vocabulary helps learners "to outperform their competence" (Nunan 103), as this will allow them to handle unpredictable communicative situations (Laufer and Hulstijn 6). Theoretical advancement in the field of vocabulary L2 learning and acquisition was limited when compared with grammar teaching, despite its

---

[1] **Faten Ahmed Alarjani** is an English language instructor and a PhD candidate in Applied Linguistics in the Department of English Language and Literature, College of Arts, King Saud University. She is interested in quantitative approaches to research within the areas of Psycholinguistics, Language Acquisition and Corpus Linguistics. Email: fatenalarjani@gmail.com.

importance, until the Involvement Load Hypothesis was developed by Laufer and Hulstijn in 2001. The Involvement Load Hypothesis claims that the retention of forms (words) and meaning depend on the manipulation of the cognitive and motivational variables within tasks, especially in incidental learning situations (Laufer and Hulstijn). Keating defined the Involvement Load Hypothesis as:

> a task-induced involvement and a motivational-cognitive construct that consists of three task factors: need, search and evaluation. Need is the motivational, non-cognitive component of involvement and refers to whether knowledge of novel words is required to complete a task…. [while] Search and evaluation are the cognitive components of involvement because they entail information processing (i.e. noticing new words and allocating attention to them). (366)

Significant role of the knowledge of word meaning and retention of word meaning has motivated research in second language acquisition and second language learning. Won defines knowledge of word meaning as "every dimension of complex word knowledge related to comprehension" (11). This knowledge and its retention play a significant role, as selecting the best techniques of vocabulary learning relies on it (Yaqubi et al.). Laufer and Hulstijn claim that "[i]t is generally agreed that retention of new information [word meaning or vocabulary] depends on the amount and the quality of attention that individuals pay to various aspects of words" (541).

## Literature Review
The Involvement Load Hypothesis was proposed by Laufer and Hulstijn, who argue that the effort an individual invests in a task mentally, or its involvement load is an essential factor in learning. Laufer and Hulstijn's motivational-cognitive construct of the task-induced involvement is based on the framework of the depth of processing, originally proposed by Craik and Tulving. Soleimani and Rahmanian explained "depth of processing" as a hypothesis that "deals with the internal processing stages of learning a stimulus in mind… holds that this depth of processing can have the outcome of more durable and firmer traces for learning new items. They bring us two boxes: sensory memory holding information which has gone through threshold analyses; Short-term memory (STM) holding information which has gone through deeper analyses" (198).

The motivational-cognitive construct of involvement consists of three components: *need, search* and *evaluation*. Laufer and Hulstijn argue that the *need* component is the motivational, non-cognitive dimension of involvement. It has two degrees of prominence: moderate and strong. The *need* component is moderate when it is imposed by an external agent, such as the need to use a word in a sentence that the teacher has asked for. On the other hand, the *need* component is claimed to be strong when it is intrinsically motivated, that is, self-

imposed by the students, for example, the decision to look up a word in a bilingual dictionary when writing an essay. The second component, *search*, is the attempt to find out the meaning of an unknown L2 word by consulting a dictionary (Laufer and Hulstijn 543). The third component is *evaluation* which entails a comparison of a given word with other words, in order to assess whether a word does or does not fit its context (Laufer and Hulstijn 543). All three factors can be absent or present when processing a word in a natural or artificially designed task (Laufer and Hulstijn 543). The accumulation of these factors with their degrees of prominence constitutes involvement load. The basic contention of the Involvement Load Hypothesis is that "retention of unfamiliar words is conditional upon the degree of involvement in processing these words. It is conditional upon who has set the task, whether the new word has to be searched, and whether it has to be compared or combined with other words" (Laufer and Hulstijn 544).

### Operationalising Involvement Load

Laufer and Hulstijn operationalised involvement load (i.e., depth of processing) by proposing the "involvement index or load" where the factor is indexed by a number which indicates its strength of involvement. For example, the absence of a factor (*need, search* or *evaluation*) is given a 0 index, the strong presence of a factor is given a 2 index and a moderate presence of a factor is given a 1 index. They gave an example of two tasks with two different involvement indexes. The first task required the participants to write original sentences using words the teacher has provided them with. According to Laufer and Hulstijn, there is a moderate *need* component here since the task is imposed by the teacher but no *search* component since the teacher has provided the meaning; howbeit, there is a strong *evaluation* component in the task since the participants have to use the new words. Therefore, the task has an involvement index of 3 (1+0+2). The second example required that the participants read a text with a glossary of the new words and answer the accompanying comprehension questions. Laufer and Hulstijn argued that the task would have a moderate *need* component, but no *search* or *evaluation* components. The task's involvement index was 1 (1+0+0). They added that the first task induces a greater involvement than the second task. Laufer and Hulstijn argued that the concept of involvement can be empirically investigated by devising incidental-learning tasks with varying degrees of *need, search* and *evaluation*. This is taken up in the current study.

Laufer and Hulstijn went on to conduct a study to investigate whether retention of vocabulary acquired incidentally is dependent on the amount of task-induced involvement. Also, the study examined the short and long-term retention of ten unfamiliar words in three learning tasks (reading comprehension, comprehension plus filling in target words and composition-writing with target words) with varying degrees of "involvement loads" of various combinations of

*need, search* and *evaluation*. The participants were divided into six groups of advanced university learners of English as a foreign language in the Netherlands (three groups) and Israel (three groups). The total number of participants in the study was 97 in the Netherlands and 128 in Israel. The first group received the reading comprehension with marginal glosses task (involvement load [IL] of 1). They were provided with a text and a set of ten multiple-choice comprehension questions. The second group received the reading comprehension plus a fill-in task (IL of 2). They were given the same text and the same questions as those in the first group; however, the ten words were deleted from the text, leaving ten gaps. The ten target words, plus the five words that did not appear in the original text, were written in random order as a list on a separate page, with their L1 translations and L2 explanations. The third group received the task of writing a composition and incorporating the target words (IL of 3). The participants in the third group were asked to write a letter to the editor of a British newspaper. All participants were only post-tested. They were asked to give the L1 equivalents or English explanations for these words. Also, they were asked to indicate whether they had known the words before the task. One week later in the Netherlands and two weeks later in Israel, the participants received the same test again. The study found that the writing composition groups scored higher (significant task effect [$F(2, 84) = 11.50$; $p < .001$; $\eta2 = .22$]) than the reading plus fill-in groups, which, in turn, was higher than that of the reading groups. The participants outperforming on the composition task support Swain's Output Hypothesis (Swain), given that the composition task required the participants to stretch their linguistic resources. The Involvement Load Hypothesis proposed "that higher involvement in a word induced by the task will result in better retention, regardless of whether it is an input or an output task" (Laufer and Hulstijn 546).

Laufer and Hulstijn found that the Involvement Load Hypothesis increased vocabulary learning and acquisition. However, their results cannot be generalised, since there were some issues in the research. First, they did not pretest the participants; they were only post-tested. Therefore, the amount of improvement and the effect of prior knowledge are not accurately measured. The study might have benefited from a pretest to ensure that the development was the result of the intervention rather than prior knowledge. Also, there was insufficient time between the immediate post-test and the delayed post-test (only one week), which was not enough to examine the effect of long-term retention. Second, the participants were all from an advanced-level course in English. This is problematic for two reasons; first, the study might have presented more empirical results if they had a more diversified group with varying levels of English; second, the results might suffer from the ceiling effect because of the participants' high language proficiency. Finally, the methodology also had some issues such as the participants were asked after the intervention to indicate whether they had known the words before the task. This was not enough to ensure they did not have prior

knowledge of the words, and that might have influenced the post-test results. The current study considered all this by having participants with varying levels of language proficiency (from low to mid). A pretest was administered to the participants to ensure the validity of the results and that the outcome of the study is the result of the intervention not the participants' prior knowledge.

### *Empirical Investigations of the Involvement Load Hypothesis*
Keating conducted a study where he used three tasks with different involvement loads to assess the predictive nature of the Involvement Load Hypothesis, to investigate whether the hypothesis can be extended to low-proficiency learners. The participants ($n$=79) who were beginning learners of Spanish were divided into three groups with one task each. The first group received the task that consisted of a reading passage with marginal glosses (IL was 1), and the second group received the reading comprehension plus fill-in task (IL was 2) while the third group was given the writing task where they were required to write original sentences using the targeted words (IL was 3). The participants' passive and active knowledge of the target words was immediately post-tested after the intervention (i.e., the tasks). A delayed post-test was administered two weeks later to see effects on retention. Two tests were used to measure the participants' passive and active knowledge: the test for passive word knowledge was a Spanish-English word translation task, while the test for active word knowledge was an English-Spanish sentence translation task. The results strongly supported the hypothesis that the involvement load hypothesis can increase the learners' retention of the vocabulary and that the results can be generalised to low-proficiency learners. The third group (the writing sentence task) performed better on the immediate post-test than the other two groups however, the difference was not statistically significant. Furthermore, the benefit associated with more involving tasks faded over time indicated by the participants' performance on the delayed post-test.

Soleimani and Rahmanian conducted a study to investigate whether tasks with higher involvements generated better results in the retention of vocabulary. The study examined the effect of different involvements on low proficiency EFL learners in Iran. The participants were selected after they underwent a Nelson Proficiency Test to ensure that they are of low proficiency. Afterwards, the low proficient EFL learners ($n$=33) were randomly assigned to three experimental groups: the fill-in-the-blank group (IL of 1), the reading comprehension group (IL of 2) and the sentence writing group (IL of 3). The participants were given the posttest after two days and delayed post-tested after two weeks to examine the effect of the tasks on their retention of vocabulary. In the posttest, the participants were asked to provide the English or the Iranian equivalent of the ten targeted words they learned from the tasks during the intervention stage. The mean scores of the three groups were compared using the one-way ANOVA test. The study found that the sentence-writing task group performed significantly

better than the other two tasks F (2, 30) = 16.72; $p$ = .000 < .05. The sentence making task resulted in significantly better retention. The results of the study were in line with Laufer and Hulstijn's findings and support the Involvement Load Hypothesis assumptions. However, the study had some limitations such as small sample size and no measurement of language production.

Tahmasbi and Farvardin's study investigated the effects of task types on English EFL learners' receptive and productive vocabulary knowledge. The implemented tasks were designed based on the Involvement Load Hypothesis that learning of vocabulary is reliant on the amount of task-induced involvement. The participants ($n$=130) in the study were EFL learners who were randomly assigned to one of six groups where each group received a different task to learn the thirty target words. The tasks implemented in this study produced the same or different involvement loads with respect to the presence and strength of each component: the paragraph writing task (IL of 4), the sentence writing task (IL of 3), the combining task (IL of 2), the fill-in-the-blank task (IL of 3), the translation task (IL of 2) and the control group that did not receive any task (IL of 0). The participants' receptive and productive knowledge of the target words were measured by a vocabulary knowledge test that was adopted from Keating's four-item Vocabulary Knowledge Scale. A post-test was immediately administered on the participants followed by a delayed post-test a month later. The results of the study revealed that all the five experimental groups performed better than the control group on the test measuring the participants' receptive and productive vocabulary knowledge. However, the experimental group that received the paragraph writing task performed significantly better compared to the other experimental groups in both the immediate post-test and the delayed post-test. Additionally, the sentence writing group outperformed the combining and control groups on both post-tests.

**The Study**

Learning and using words requires the initial acquisition of the words and relatively long-term retention of their meanings. Saudi university students have limited incidental exposure to word meanings and their usage because English is not used outside the classroom, thus, the majority of them have a limited lexicon. The purpose of the study is to examine the effectiveness of the involvement load hypothesis on the retention of the meaning of selected ten English words by adult EFL learners in Saudi Arabia in an incidental learning setting.

The research question for the study is: What is the effect of the involvement load of the "writing short sentences task" on the retention of word meaning of ten English words of Saudi female learners in Saudi Arabia?

The following hypothesis was posited on the basis of the relevant literature reviewed:

H1:  The involvement load of the "writing short sentences task" would result in considerably higher scores in the posttest measuring word meaning retention, in line with previous findings on the effects of involvement loaded tasks among Saudi students.

## Methodology

According to Laufer and Hulstijn's study, the retention of word meanings was the highest in the writing composition task, lower in reading plus fill-in task and lowest in the reading task. Therefore, this study used only the composition-writing with target words task. This study had two groups of participants:  the experimental group and the control group. The experimental group received the writing composition task and a glossary with definitions, while the control group was  given only the glossary to read. The control group was asked to read, while the experimental group was asked to form sentences using the given words in addition to reading the glossary and their definitions. The participants were pre-tested and post-tested to check whether the involvement load of each task affected the learners' retention of word meanings.  Only ten items were chosen for treatment and testing in the study. They included *rigmarole, wrath, grist, , privy to, morally derelict, curb* and *inflammatory*. The items were expected to be unfamiliar to the participants. The mean score for both groups on the pretest was 0.00 which illustrated that the participants did not have any prior knowledge of the vocabulary items.

### Participants

The participants in the study were university female learners of English (*n*=30) enrolled in Princesses Nourah University in Riyadh.  They were randomly assigned to two groups: experimental group and control group, with 15 participants per group. The experimental group received the writing task while the control group received only the glossary of terms to read.

The pretest scores of the two groups were compared. An independent sample *t*-test was used to calculate equivalency between the two groups to examine whether the difference was statistically significant. Table 1 illustrates the results of the independent *t*-tests before the experiment. The *p* value was less than 0.05, which meant that there was no statistically significant difference between the two groups in the knowledge of the vocabulary items being tested before the experiment.

### Research Tools

This study adapted the research design from Laufer and Hulstijn and included an intervention (i.e., the task) and a testing stage where the participants were pretested and post-tested.

*The Task*

The composition task where participants had to write original sentences using the target words was selected to investigate its involvement load effect on the retention of word meaning. The task was administered on the experimental group only. The participants in the experimental group were asked to read the same list of 10 words (i.e., a gloss of the words and their definitions) as the participants in the control group. The instructions for the task were as follows: "Write one sentence for each of the words listed below, you can use the vocabulary reading list provided by the teacher to help you know the meaning of the words." The task was untimed and ended when the participants completed the task. In terms of involvement load, the task induced a moderate *need* because it was imposed by the task and there was no *search*. The task had a strong *evaluation* because the words had to be used in an original context where the participants use the new words with previously known words to create original sentences. Its involvement load index was 3.

*The Tests*

The study also adopted the tests from Laufer and Hulstijn. On both the pretest and post-test, the participants were asked to provide Arabic equivalents or English explanations for each word. The task that was given to the participants was designed according to the Involvement Load Hypothesis. The reliability of the pretest and post-test was assessed using internal consistency of responses to every item in each of the tests. Cronbach's alpha coefficient was calculated for the posttest that had 10 items and reliability of 0.6.

The test used in this study to measure the participants' retention of word meaning had both construct validity and face validity. Construct validity is defined as "the extent to which we can interpret a given test score as an indicator of the ability[ies] or construct[s] we want to measure" (Larson-Hall 169). The test required the participants to recall the meaning of the words they had learned one week before during the intervention stage. If the participants were able to recall the meaning either in Arabic or English, then this was seen as evidence that the test measured what it claimed to measure, which is retention. Face validity is when the "measure appears to be assessing the intended construct under study" (Larson-Hall 182) and this was demonstrated by the participants' pretest scores where the majority in both groups scored zero (the control group $M=0.0$, and the experimental group $M=.13$), since they did not have prior knowledge of the word meanings. Thus, the test measured what it was supposed to measure. The researcher warned the participants not to exchange whatever they had learned from the task with each other in order to be much safer about the values of each component of the tasks. This in effect could decrease the threat to internal validity.

The tests were scored by the researcher and a second rater to ensure the

reliability of the scoring. The whole test was out of ten marks with each item assigned one point. A word that was left empty, wrongly translated or provided with a wrong equivalent received a score of zero. A correct response received the full point.

### Procedure

The study was conducted in a regular university class. In the intervention, the experimental group received a reading task and a composition task where the participants were given a list of 10 words with definitions and to write a single original sentence for each word. The control group received the list of 10 words and their definitions to read only (see appendix B). There was no time limit implemented for both groups during the intervention. The targeted words were unfamiliar words and expressions to ensure that the learners would not have prior knowledge of them. The pretest was implemented before the intervention, while the post-test was implemented one week later. Table 3 illustrates the different stages of the study.

Data were analysed quantitatively by using a paired sample $t$-test and an independent $t$-test to support the qualitative explanation that followed. All the data were analysed using the SPSS statistical package. To answer the research question – "What is the effect of the involvement load of the writing short sentences task on the retention of word meaning of ten English words of Saudi female learners in Saudi Arabia?" – a comparison of mean scores of the experimental group participants on the pretest was compared with the mean scores of the post-test using the paired sample $t$-test. Additionally, the scores of both groups were compared on the post-test using the independent sample $t$-test.

### Results
### Effects of Involvement Loaded Task on the Participants' Retention of the Vocabulary

The hypothesis stated that the writing task with the involvement load of 3 would result in considerably higher scores on the post-test in line with previous findings on the effects of involvement loaded task (Keating; Laufer and Hulstijn; Tahmasbi and Farvardin; Soleimani and Rahmanian). Results of the vocabulary test supported the hypothesis.

*The pretest and posttest*
Before the intervention, the experimental group scored higher on the pretest ($M$ = .13 out of 10) than the control group ($M$ = 00 out of 10). However, the difference between the experimental group and the control group is not statistically significant ($p$ = .326 > .05). Thus, it is assumed that both groups have equal knowledge of the terms. After the intervention (i.e., the involvement load task), the experimental group ($M$ = 5.60 out of 10) outperformed the control

group ($M$ = 4.13 out of 10). Therefore, the experimental group retained word meaning of the words after the intervention better than the control group. However, the difference between the groups' performance on the post-test was not statistically significant ($p$ = .095 > .05). This suggested that when participants were exposed to the involvement load task, their knowledge of the word meanings increased and was retained; however, the results cannot be generalised to other groups doing the same tasks.

The power analysis and the effect size for both groups were calculated post-hoc. The lack of significant difference between the experimental group and the control group can be attributed to the low power level of 38 % ($k$ =2, $n$= 15, *power* = .385). Having a low power level (i.e., not having enough participants) could have resulted in a Type II error (the possibility of Type II error is 62%), where a researcher concludes that there is no effect of the treatment while there is an effect. However, this is avoided here by concluding that there is a possibility of an effect of the treatment (involvement load task) over the experimental group.

A paired *t*-test that compares between the experimental group means on the pretest and posttest was carried out to examine whether there is a statistical difference as a result of the treatment. The results of the paired *t*-test show that there is a difference between the experimental group performance on the pretest ($M$=.13 out of 10) and the post-test ($M$= 5.60 out of 10). The difference between the pretest and post-test scores ($p$ = 0.00 <.05) is statistically significant within the experimental group. This suggested that, when exposed to the loaded task during the intervention stage, the experimental group gained knowledge of the word meanings and retained that knowledge after a week, as suggested by the post-test results.

Thus, to answer the research question, the involvement loaded task (i.e., writing short sentences using the words introduced by the glosses provided by the teacher) did have a positive and significant effect on knowledge and retention of word meanings. The effect was evident in the results of the experimental group post-test. The experimental group outperformed the control group on the post-tests and sustained the knowledge gained for a week after the intervention. Additionally, the experimental group performed significantly better in the post-test than in the pretest which provides additional evidence to support that the task positively affected the experimental group.

## Discussion
### The Effect of the Involvement Loaded Task on Knowledge of Word Meanings

The participants in the study were female English translation students enrolled in a Princesses Norah University in Riyadh. Students of translation often need a rich and diverse knowledge of English words and their meanings to be able to translate written texts or interpret speeches. Using traditional methods, such as

memorising lists of words and their meanings, to help students develop their lexical knowledge is tedious and usually does not result in long-term retention of knowledge. After the participants in the experimental group did the writing task, there was a significant increase (M=5.60, p =,000 < .05) in their knowledge of the word meanings. The increase in the participants' knowledge was apparent in their scores in the posttest (M= 5.60) when compared with their scores in the pretest (M= .13). However, the results of the tests only measured the development of the participants' active knowledge since the participants were able to recall the word meanings. This assumption is in line with Keating and Tahmasbi, and Farvardin who claim that the involvement loaded task develops the participants' active knowledge of the words. Nonetheless, the posttest results cannot show whether the learners can use the words in new contexts, as it only reflects their ability to recall the word meanings and express them in writing. This is only a part of active knowledge. The posttest did not measure all aspects of the participants' active knowledge; thus, the results of the study cannot be generalised. Future studies will benefit from investigating the effects of involvement load tasks on the learners active and passive knowledge.

### The Effect of the Involvement Loaded Task on Retention of Word Meaning

Results from the tests that measure word meaning retention showed that participants in the experimental and the control groups had no knowledge of the words before the intervention. The purpose of choosing unfamiliar vocabulary was to ensure that the gained knowledge of word meanings was the result of the treatment and not the result of prior knowledge. There was a slight difference between the means of the two groups in the pretest. However, it was not statistically significant. Thus, it was assumed that both groups had similar prior knowledge of the words.

The improvement of the experimental group in recalling the word meanings in the posttest demonstrated the positive effect of the involvement loaded task on the participants' retention. The experimental group did not gain much in comparison to the control group which was evident in the mean scores of the experimental group (M=5.60 out of 10) and the control group (M= 4.13 out of 10) in the posttest. The difference between the two groups was not statistically significant. However, the importance of the results cannot be ruled out. The reason behind the absence of a significant difference between the groups can be attributed to the low power of the effect size. The calculated power for both groups was around 38% which is a lot below the expected power margin of 60% to 80%. Therefore, it is recommended that future studies have a higher number of participants to avoid the low power issue.

The task had a long-term effect on the experimental group, with retention of the word meanings one week after the intervention. The experimental group's

results in the pretest (M =.13) when compared with their posttest results (M = 5.60) demonstrate a significant development in retention (p=0.00 < .05).

The findings of this study support Hulstijn and Laufer's hypothesis of involvement load where they suggested that composition writing tasks yield better outcomes when learning words than other tasks with lower involvement index. The experimental group in this study also improved in the immediate post-test after the writing composition task. Participants in Hulstijn and Laufer study were not tested before implementing the tasks. Therefore, there was no way to make sure that their scores were the results of prior knowledge or the effect of the implemented task during the intervention stage. The participants in the current study, however, were pretested to ensure that the improvement in the experimental group was truly the result of the treatment.

Findings from the posttest could be compared to the results of previous studies, such as by Hulstijn and Laufer, Keating, Tahmasbi and Farvardin, and Soleimani and Rahmanian. The current study used one task only which is the composition task, with the involvement load index of 4. Participants in the experimental group in the current study scored higher in the posttest when compared with the pretest, meaning they scored higher on a measure of word meaning retention. The finding was similar to the results of previous studies where participants outperformed on measures of vocabulary retention (i.e., posttest) compared to pretest. As a result, these studies (Hulstijn and Laufer, Keating, Tahmasbi and Farvardin, Soleimani and Rahmanian) suggested that the higher the involvement index a task has, the better the retention will be for the learners. This is supported by the results of the current study.

All the discussed studies, except for Keating, asked the participants to write an original composition. However, in the current study, participants did not write a composition but unrelated sentences. Keating argued that production of a connected discourse might involve more elaborate and deeper processing of the word meanings than producing disconnected sentences. He added that any difference between writing related sentences in an essay form and writing unrelated sentences on their own, such as the ones done in this study, might not be accounted for by the Involvement Load Hypothesis because the tasks are similar regarding the amount of *need, search* and *evaluation* they prompt. Both writing related sentences in an essay and writing unrelated sentence tasks have the same involvement index of 3. Future studies should compare the two kinds of composition tasks and investigate their effects on learners' retention and whether the two tasks should be assigned different involvement indexes.

**Conclusion**

The significant role of acquiring meaning and vocabulary in EFL and ESL prompted this study since Saudi learners suffer from lack of exposure to the target language (i.e., English). The main objective for conducting this study was to test

the claim that involvement loaded task had a facilitative role in developing learners' retention of word meanings and increasing their lexical knowledge. This study employed a task based on the Involvement Load Hypothesis (Laufer and Hulstijn) because of its previous successes with learners in Netherland and Israel. Furthermore, this study was an attempt to provide empirical evidence for the hypothesis in second language learning. Results also showed that loaded tasks with a higher index improved the participants' retention and knowledge of word meanings.

The results of this study suggested that the involvement loaded task not only developed the participants' knowledge of the target words but also helped in meaning retention. Further verification of this claim would benefit from more empirical evidence with participants with varying proficiency levels. Larger samples would lend more support to the trends found in the research thus far. Future studies are advised to have a larger sample size to avoid the low power issue that might have affected this study. Furthermore, the impact of the involvement loaded task on the participants' retention of word meaning could have been verified on a second and even a third delayed post-test implemented 4 and 6 weeks after the intervention to examine delayed retention over different periods.

To conclude, this study investigated the effectiveness of the involvement loaded task on the retention of word meaning. The quantitative data suggested that the involvement loaded task affected the participants' retention of word meaning. The findings of this study is a valuable contribution to the SLA and TESL fields, concerning the role of the involvement loaded task in increasing the learners' retention and knowledge of word meaning.

## Acknowledgement

## Works Cited

Craik, Fergus IM and Endel Tulving. "Depth of Processing and the Retention of Words in Episodic Memory." *Journal of Experimental Psychology: General* 104.3 (1975): 268-85

Hulstijn, Jan H., and Batia Laufer. "Some empirical evidence for the involvement load hypothesis in vocabulary acquisition." Language learning 51.3 (2001): 539-558.

Keating, Gregory D. "Task Effectiveness and Word Learning in a Second Language: The Involvement Load Hypothesis on Trial." *Language Teaching Research* 12.3 (2008): 365-86.

Larson-Hall, Jenifer. A guide to doing statistics in second language research using SPSS and R. Routledge, 2015.

Laufer, Batia and Jan Hulstijn. "Incidental Vocabulary Acquisition in a Second Language: The Construct of Task-Induced Involvement." *Applied Linguistics* 22.1 (2001): 1-26.

Nunan, David. *Second Language Teaching and Learning,* 1999.

Soleimani, Hassan and Mahboubeh Rahmanian. "Vocabulary Acquisition and Task Effectiveness in Involvement Load Hypothesis: A Case in Iran." *International Journal of Applied Linguistics and English Literature* 4.5 (2015): 198-205.

Swain, Merrill. "Communicative Competence: Some Roles of Comprehensible Input and Comprehensible Output in Its Development." *Input in Second Language Acquisition* 15 (1985): 165-79.

Tahmasbi, Maryam and Mohammad Taghi Farvardin. "Probing the Effects of Task Types on EFL Learners' Receptive and Productive Vocabulary Knowledge: The Case of Involvement Load Hypothesis." *SAGE Open* 7.3 (2017): 1-10.

Won, Mijin. "The Effects of Vocabulary Instruction on English language Learners." Diss. Texas Tech University, 2008.

Yaqubi, Bagher, et al. "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 (2012): 145-63.