



THE UNIVERSITY OF  
**WAIKATO**  
*Te Whare Wānanga o Waikato*

*Faculty of Education*

Te Kura Toi Tangata

**Waikato Journal of Education**

Te Hautaka Mātauranga o Waikato



Volume 20, Issue 2: 2015

## Waikato Journal of Education Te Hautaka Mātauranga o Waikato

Special Section Editor: Carl Mika & Vanessa Paki  
Current General Editors: Clive Pope and Noeline Wright  
Editorial Board: Bronwen Cowie, Deborah Fraser, Richard Hill, Clive Pope, Margie Hōhepa, Sally Peters, Noeline Wright.  
International Board Members: Tony Brown (England), Alec Couros (Canada), Agnes Hulme (England), Cathy Reischl (USA), Iram Siraj (England), Christine Sleeter (USA), John Smyth (Australia), Janet Soler (England).

The *Waikato Journal of Education* is a peer refereed journal, published twice a year. This journal takes an eclectic approach to the broad field of education. It embraces creative, qualitative and quantitative methods and topics. The editorial board is currently exploring options for online publication formats to further increase authorial options.

The Wilf Malcolm Institute of Educational Research (WMIER), which is part of the Faculty of Education, The University of Waikato, publishes the journal.

There are two major submission deadline dates: December 1 (for publication the following year in May); June 1 (for publication in the same year in November). Please submit your article or abstract on the website <http://wje.org.nz/index.php/WJE>.

Submissions for special sections of the journal are usually by invitation. Offers for topics for these special sections, along with offers to edit special sections are also welcome.

Contact details: The Administrator Wilf Malcolm Institute of Educational Research, Faculty of Education, The University of Waikato, Private Bag 3105, Hamilton, 3240, New Zealand. Email: [wmier@waikato.ac.nz](mailto:wmier@waikato.ac.nz)

Copyright:



This work is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-nc-sa/4.0/).

Publisher: Faculty of Education, The University of Waikato  
Cover design: Donn Ratana  
ISSN: 2382-0373

# Waikato Journal of Education

## Te Hautaka Mātauranga o Waikato

Volume 20, Issue 2, 2015

### Special Section: Te Kōhao o Te Rangahau—The indigenous lens on research

Special section editorial: Te kōhao o te rangahau—The Indigenous Lens on Research <i>Carl Mika &amp; Vanessa Paki</i>	3
Indigenous collaboration towards hybrid musical production builds a better community: Taioreo—Māui, Izanagi and Izanami <i>Te Manaarooha Pirihira Rollo</i>	5
Ngā kaiako mō āpōpō: Waiho mā te iwi e kōrero <i>Nā Karaitiana Tamatea, Margie Hōhepa, Ngārewa Hāwera, me Sharyn Heaton</i>	17
Ngā tapuwae o mua mō muri: Footprints of the past to motivate today's diverse learners <i>Angus Hikairo Macfarlane</i>	27
Whatu: A Māori approach to research <i>Lesley Rameka</i>	39
Exploring whakapapa (genealogy) as a cultural concept to mapping transition journeys, understanding what is happening and discovering new insights <i>Vanessa Paki &amp; Sally Peters</i>	49
The thing's revelation: Some thoughts on Māori philosophical research <i>Carl Te Hira Mika</i>	61

### General Section

Relational and culturally responsive supervision of doctoral students working in Māori contexts: Inspirations from the Kingitanga <i>Ted Glynn &amp; Mere Berryman</i>	69
The effects of task type and task involvement load on vocabulary learning <i>Mandana Hazrat</i>	79
School leaders growing leadership from within: A framework for the development of school leaders <i>Anthony Fisher &amp; Tracey Carlyon</i>	93



## The effects of task type and task involvement load on vocabulary learning

**Mandana Hazrat**

The Islamic Republic of Iran

### Abstract

*Vocabulary learning has been a challenging task for Iranian EFL learners. Forgetting and inability to use the learnt words productively are their common problems. In order to assist Iranian English teachers and material developers to satisfy the learners' needs, this study put the Involvement Load Hypothesis and the Depth of Processing Hypothesis into practice. The former emphasises the impact of a task's determining factor for vocabulary learning including a combination of need, search, and evaluation of the new words, which is called task involvement load, whereas the latter focuses on the effect of the task type on vocabulary learning. To this end, 39 adult Iranian EFL university students in three groups learnt 10 unfamiliar words via three types of tasks with the same involvement load including reading plus multiple-choice glosses, sentence writing and oral sentence production. Learning and remembering the new words were measured using the immediate and delayed productive and receptive vocabulary tests. The results suggested that, being equal in terms of the involvement load, the writing task resulted in more long-term productive vocabulary learning compared with the reading and speaking tasks. The speaking task was the least effective one regarding its impact on vocabulary learning.*

### Keywords

The involvement load hypothesis, the depth of processing hypothesis, active vocabulary learning, passive vocabulary learning.

### Introduction

Learning English vocabulary is a challenging task for Iranian English as a foreign language (EFL) learners, partly because English is not heard outside classrooms. In spite of such difficulties, there is a burgeoning interest in learning English for purposes such as furthering education in other countries or being qualified for job opportunities.

However, Iran's education system appears not to live up to the needs and expectations of EFL learners. Iranian public schools usually begin EFL classes from the age of 12. The prevalent teaching method is grammar translation. It is a common complaint among Iranian EFL learners that their English vocabulary knowledge is a passive (receptive) type that serves reading purposes without meeting their communicative needs. In other words, they are not satisfied gaining knowledge of words' meanings without being able to use their knowledge for speaking and writing. This is exacerbated by not having opportunities to practise English outside the classrooms. As a result, long-term active (productive) vocabulary learning is a challenge in Iranian contexts. To satisfy learners'



demands, English teachers trial ways of assisting learners to add to their active vocabulary knowledge, but it is a burdensome task in private language institutes where they try to provide optimal contexts for vocabulary learning during the class time using vocabulary-learning tasks. There is thus a need to assess the effectiveness of such tasks and their resultant learning and to design tasks that can contribute to vocabulary learning.

An English teacher undertook to assist her colleagues to evaluate the effectiveness of the vocabulary-learning tasks based on their students' learning needs. To achieve this goal, two hypotheses provided the bases for this study including Hulstijn and Laufer's (2001) *Involvement Load Hypothesis* and Craik and Lockhart's (1972) *Depth of Processing Hypothesis*.

As Laufer and Hulstijn (2001) proposed, based on the Involvement Load Hypothesis, the determining factors for vocabulary learning in a task include need, search and evaluation. The combination of these three factors is referred to as involvement. Involvement is a construct that can account for learners' success in vocabulary learning. The first component is learners' need to achieve a *learning goal*. For instance, if a learner is reading a text and an unknown word is necessary for comprehending the text, the learner will need to understand the word. The second component, *search*, is making an effort to find the meaning of an unknown L2 words. Lastly, *evaluation* involves a comparison of a given L2 word with other words, a decision about a specific meaning of an L2 word in a specific context, combining a word with other words in order to examine whether the word does or does not fit its context, or choosing the most suitable translation of an L1 word in order to do an L2 writing task (Laufer & Hulstijn, 2001).

The Involvement Load Hypothesis predicts that the higher the involvement load of a task, the higher the opportunity for vocabulary learning. The highest level of involvement is achieved when learners perceive the need to learn, search for the meaning, and make a decision about how a new word can combine with other words in a sentence. In other words, by simply receiving the meaning of unknown words or recognising the differences between the words by choosing a correct meaning in fill-in-the-blank tasks, learners seldom reach the level of involvement required for vocabulary learning (Rott, 2007).

In the process of vocabulary learning through natural or designed tasks, one or more components of involvement may be absent (Laufer & Hulstijn, 2001). In addition, the components can be presented in their moderate or strong versions. Need is moderate if it is imposed externally (e.g., by the teacher) and is strong if it is self-imposed. Regarding the evaluation component, it is perceived as moderate if a decision has to be made between the given words or between several senses of a given word. It is referred to as strong if a decision is made about combining a new word with other words in order to produce an original sentence (Hulstijn & Laufer, 2001). Consequently, a good task induces a high degree of involvement that can predict vocabulary learning. The degree of involvement can be shown by the involvement index, which is judged numerically: 0 (for absence), 1 (moderate), 2 (strong) (Laufer & Hulstijn, 2001). For instance, in a reading comprehension task where unknown words are glossed and they are necessary for answering the comprehension questions, the involvement index is 1 because this task induces moderate need (need is imposed by the task), no search (words' meanings are given) and no evaluation. If the glosses are eliminated, the search component will appear and the involvement index will be 2.

On the other hand, Craik and Lockhart (1972) in their Depth of Processing Hypothesis argued that the chance of storing new information in memory depends on the shallowness and depth of its initial processing. While processing the meaning of a new word happens at a deep level, processing phonological and orthographic forms occurs at a shallow level. Craik and Tulving (1975) further suggested that presence or absence of semantic encoding is not sufficient for retention. Richness of encoding also plays a part. Hulstijn and Laufer (2001) also stated that the likelihood of retaining a word depends on attention to a word's properties including pronunciation, orthography, meaning, grammatical information and semantic relations to other words. In this regard, it is postulated that reading, writing and speaking involve different types of processing. According to Field (2006), lower levels of reading involve decoding orthographic features and linking them to a concept, while lower levels of speaking and writing involve encoding a message by retrieving appropriate spoken and written forms respectively. At higher levels of processing, reading involves applying background

knowledge, interpreting a writer's intention, and constructing global meaning whereas written or spoken production requires planning and organising.

Consequently, in addition to the involvement load, task type requires a specific depth of processing that may affect vocabulary learning and retention. The aim of the study examined here was to compare the effects of task type and task involvement load for short- and long-term vocabulary learning. Moreover, if the type of processing matters, which task type involves the processing that is more influential for active and passive vocabulary learning?

To this end, the researcher decided to conduct the study with a group of participants who were learning English vocabulary. They were first-year Iranian university students studying English Literature at Tehran Payame-Noor University. To cope with the demands of their course, they were required to expand their vocabulary knowledge and use the learnt words productively.

## Literature review

In the area of vocabulary learning and retention, the groundwork laid by Craik and Lockhart's (1972) Depth of Processing Hypothesis suggested that learning and remembering a word hinges on the amount of attention given to it and the amount of manipulation involved at the early learning stages. Researchers investigated predictions of the Involvement Load Hypothesis by using tasks with various load degrees of need, search, and evaluation. For instance, Hulstijn and Laufer (2001) explored short-term and long-term retention of unknown words in three tasks with various involvement loads including reading comprehension, comprehension plus fill-in task, and composition writing. In line with the Involvement Load Hypothesis predictions, retention was highest in composition writing, lower in reading and fill-in task, and lowest in reading task. A similar study was carried out by Keating (2008), who set three vocabulary learning tasks with different degrees of involvement including reading comprehension (no effort), reading comprehension plus fill-in task (moderate effort), and sentence writing using target words (strong effort). The result was congruent with the predictions of the Involvement Load Hypothesis and Hulstijn and Laufer's (2001) findings, and reflected in Jing and Jianbin's (2009) study.

Rott (2007) also examined the Involvement Load Hypothesis, this time regarding levels of proficiency, adding knowledge about the impact of visual enhancement. Her study involved three tasks: the first one glossed target words four times, the second task glossed them first, retrieved them in the first language, and bolded them twice, and the third one glossed them first and bolded them three times. The results supported the hypothesis by showing that the first and the second tasks had similar impacts and both of them resulted in more productive word gain than the last task. However, repeated visual enhancement was found to have no impact on vocabulary learning.

However, not all the studies corroborated the Involvement Load Hypothesis predictions, such as Kim (2011). Two experiments were carried out to examine the Involvement Load Hypothesis predictions across levels of proficiency. Findings from the first experiment contradicted the above-mentioned predictions. This included three tasks hypothesized to have different levels of involvement including reading, gap filling, and composition writing. The second experiment involved two tasks hypothesized to offer the same levels of involvement including writing composition and writing sentences. The first experiment demonstrated greater vocabulary gain in the composition task, which required a higher level of involvement. However, no significant difference was found between the impacts of the reading and gap filling tasks regarding vocabulary learning. The second experiment indicated the same impact on vocabulary learning and retention. Moreover, the results were consistent across levels of proficiency.

Martínez-Fernández (2008) measured the impact of three tasks on vocabulary development. The first task involved need component, but not search and evaluation ones. The participants read the text with translation single glosses of target words. In the second task, target words were deleted from the text for participants to fill in from a provided list of L2 words with their translations. In the last task, a multiple-choice gloss, all three components of involvement were needed. The participants read the text with translation multiple-choice glosses with four options including the correct meaning, two additional meanings that would make sense in the context, and a do not know option. The result did

not confirm the Involvement Load Hypothesis predictions. The fill-in task led to higher vocabulary development than the multiple-choice gloss task. The fill-in and single gloss tasks did not show a significant difference regarding their effectiveness for vocabulary learning.

These studies mostly investigated the Involvement Load Hypothesis predictions using tasks with increasing involvement loads without addressing the impact of the task type on vocabulary learning, but this gap was investigated by Pichette, De Serres and Lafontaine (2011). They compared the effectiveness of reading and writing sentences. Participants of their study learned L2 words while reading and writing sentences. Vocabulary learning was assessed using immediate and delayed recall tests. The result of the immediate recall test suggested that the writing task was superior to the reading task, but, based on the delayed test, they concluded that this superiority disappeared over time.

Vidal (2011) also compared the effects of the listening and reading tasks on vocabulary acquisition across levels of proficiency. Participants either read the texts or watched the lectures for learning unknown words. The result indicated the superiority of the reading task over the listening task for vocabulary acquisition. However, the difference between the impacts of the tasks decreased as the participants' proficiency increased. The reading task's superiority was also evident in the delayed vocabulary test.

### 3. Method

#### Participants

Forty-seven first-year female Iranian university students, aged 18–24 and studying English Literature at Tehran Payame-Noor University, volunteered to take part in this study. Because the participants had to be assigned into three groups, homogeneity of the groups was checked, based on the participants' existing academic IELTS test scores. This excluded five participants and therefore data collection began with 42 participants. Later, three more participants were excluded due to not attending the sessions or having exposure to the target words outside the context of the study. Therefore, data were collected from 12 participants in a speaking group, 14 in a reading group, and 13 in a writing group, 39 participants in all.

#### Material

##### *Treatment passages*

Two passages from the reading sections of previous academic IELTS tests were used. Topics of the passages were *Party Labels in Mid-Eighteenth Century England* and *Medical Profession*. Each passage contained five target words to learn. For practical reasons, passages were modified: shortened, with target words bolded. Bolding the target words assisted the participants to find them in the passages without interfering with the aims of the study.

##### *Target words*

Ten low-frequency words were chosen for investigation. All the words were abstract nouns: *Spate, Tack, Plight, Veneer, Onus, Chagrin, Treason, Nemesis, Exile* and *Tenet*. Because word length may be considered an effective factor in vocabulary learning, there was an attempt to include more uniform words in terms of their length. Each target word appeared only once in the texts. The sentences in which the target words appeared included no other low-frequency word.

##### *Tasks*

Three tasks with the same involvement index but of different types were administered to provide opportunities for vocabulary learning. The involvement index refers to a task's determining factors for

vocabulary learning based on need, search, and evaluation components (see Table 1). In this table, *one plus* (+) means moderate and *double plus* (++) strong version of the components as defined in the introduction section of this paper. Minus (-) indicates the absence of the components. The first task was based on Martínez-Fernández's (2008) study. The reading passages had bolded target words and participants needed to circle the best option. This task induced moderate *need* (need was imposed by the tasks), *search* (words' meanings were not provided), and *evaluation* (participants had to evaluate the meaning that fit best in the context). Consequently, the reading task's involvement index was 3.

The second task was writing. Participants were given a list of the target words accompanied by their L1 translations and asked to write original sentences using them. This task involved moderate need (need was externally imposed), no search (the meanings were given), and strong evaluation (participants had to produce original sentences). As a result, the same as the previous task, the involvement index was 3.

The third task was speaking. As in the writing task, participants were given the list of the target words and their L1 translations. They were required to speak sentences containing the target words. This took place as a private session between each student and the researcher. This task had the same involvement index as the two previous ones. The speaking group sessions took place on one day.

Although the time participants took to complete these tasks varied, Hulstijn and Laufer (2001) argue that time on task is an inherent property of the tasks. Therefore, it did not introduce a separate variable to be controlled.

**Table 3.** The tasks' involvement indices

	Need	Search	Evaluation	Involvement Index
Reading Task	+	+	+	3
Writing Task	+	-	++	3
Speaking Task	+	-	++	3

## Procedures and measures

To ensure that all the target words were unfamiliar to the participants, a vocabulary checklist test was administered to all participants in the first session of the study. It included 10 target words and 20 distractors. Participants were required to provide meanings for the words either in English or in their native language (Persian) and cross out words they did not know.

Six days later, in the second session of the experiment, participants of each group received their vocabulary learning tasks and the immediate vocabulary tests. The reading group was first and their session lasted about 40 minutes. The writing group session took nearly 30 minutes for the whole group, and finally the speaking group session lasted about 20 minutes per participant. Immediately after each task, students' task papers were collected and short-term vocabulary learning was assessed using two measures based on Laufer, Elder, Hill, and Congdon's (2004) study: an immediate active recall test and a passive recall test.

After administering the immediate vocabulary tests, participants were not allowed to jot down the words they had practised or to take the task or test papers with them to prevent the learners from having uncontrolled exposures to the target words outside the research context. Previous research confirmed the impact of frequency of exposure on vocabulary learning (Eckerth & Tavakoli, 2012; Sun, 2014). Therefore, practising the target words outside the research context could introduce a new variable that might affect the research result.

For measuring long-term vocabulary learning, delayed tests sessions were held three weeks later. The same measures were administered to the three groups of participants in the same way as the previous administrations. Scores of the delayed active and passive vocabulary tests reflected the degree to which long-term productive and receptive vocabulary learning happened using the learning tasks. Before the tests, participants were asked if they had encountered the target words during the time

between the two test administrations or not. If the answer was yes, participants' scores were excluded from the study. As mentioned before, exposure outside the context of the study might lead to unequal learning opportunities for participants and would affect their test scores. Unreliable test scores could not provide the basis for drawing conclusions.

Finally, papers of the reading and writing groups and recorded oral productions of the speaking group were scored by two raters. Each correct translation received one point and each wrong translation or a word that had not been translated received zero. For consistency, the raters discussed any disagreements and reached a compromise.

#### 4. Results and discussion

To achieve a result, we compared the three groups' test scores using the Statistical Package for the Social Sciences (SPSS) to run the Kruskal-Wallis test, which makes between-group comparisons.

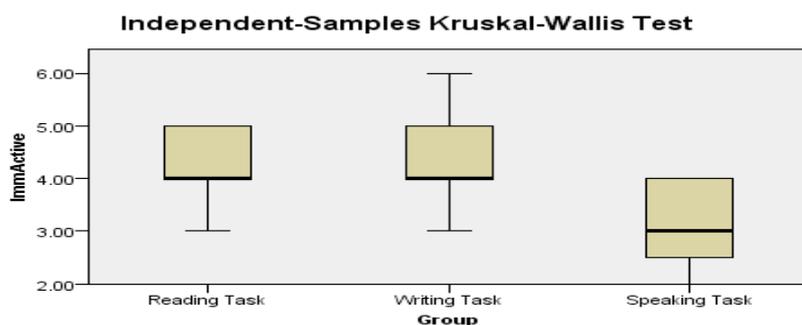
Regarding the immediate active vocabulary test, Table 2 highlighted that the reading (Med = 4) and writing (Med = 4) groups had the same median scores while the median score for the speaking group was 3.

**Table 4. Descriptive Statistics, Median Scores by Groups**

Group	Immediate. Active
Reading Task	4.00
Writing Task	4.00
Speaking Task	3.00

The results of the Kruskal-Wallis test ( $H(2) = 11.49, p < .05$ ) (Figure 1) indicated that there were significant differences between the three groups' performance on the immediate active vocabulary test.

**Figure 2. Kruskal-Wallis Test; Immediate Active Vocabulary by Groups**



<b>Total N</b>	39
<b>Test Statistic</b>	11.491
<b>Degrees of Freedom</b>	2
<b>Asymptotic Sig. (2-sided test)</b>	.003

1. The test statistic is adjusted for ties.

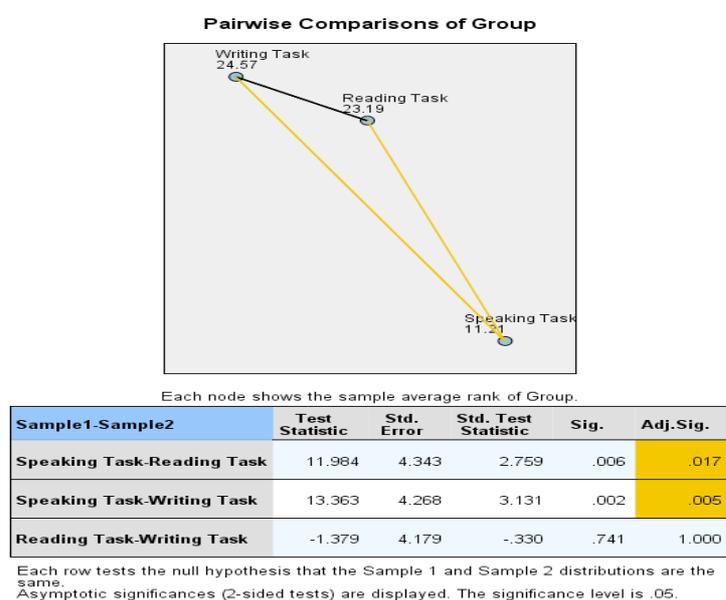
The results of the pairwise comparison tests (Figure 2) indicated that:

- The reading task group (Med = 4) significantly outperformed the speaking task group (Med = 3) on the immediate active vocabulary test ( $Z = 2.75, p < .05$ ).

- The writing task group (Med = 4) significantly outperformed the speaking task group (Med = 3) on the immediate active vocabulary test ( $Z = 3.13, p < .05$ ).
- There was not any significant difference between the median scores of the writing and reading tasks groups on the immediate active vocabulary test ( $Z = -.33, p < .05$ ).

The results signified that although the three tasks involved learners through need, search and evaluation components to the same degree and were expected to bring about the same learning, the writing and reading tasks were more effective for short-term active vocabulary learning than the speaking task. In addition, the writing and reading tasks had similar impacts on short-term active vocabulary learning. Because one of Iranian EFL learners’ difficulties regarding vocabulary learning is developing productive knowledge of words to be able to use them in oral and written production, the result may show that reading and writing tasks can help them reach their aim. English teachers and autonomous learners can take advantage of the reading and writing tasks for productive vocabulary learning. However, this result is only limited to short-term learning.

**Figure 3. Pairwise Comparison, Immediate Active Vocabulary by Groups**



Note: The statistics displayed on the nodes in Figure 2 are the mean ranks for the three groups based on which all computations are made.

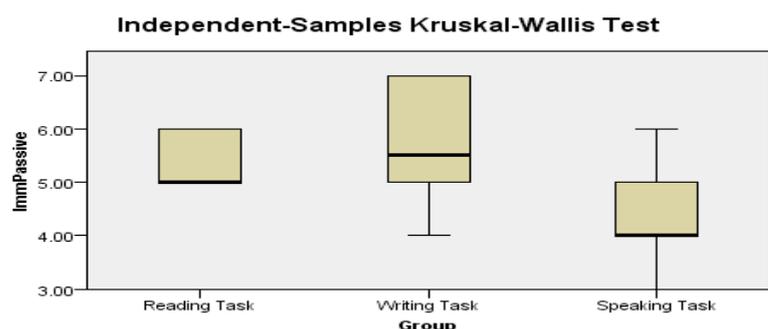
With respect to the immediate passive vocabulary test, Table 3 depicted that the reading (Med = 5) and writing (Med = 5.50) groups had similar median scores while the median score for the speaking group was 4.

**Table 5. Descriptive Statistics; Median Scores by Groups**

Group	Immediate. Passive
Reading Task	5.00
Writing Task	5.50
Speaking Task	4.00

The results of the Kruskal-Wallis test ( $H(2) = 11.57, p < .05$ ) (Figure 3) indicated that there were significant differences between the three groups’ performance on the immediate passive vocabulary test.

**Figure 4. Kruskal-Wallis Test; Immediate Passive Vocabulary by Groups**



<b>Total N</b>	39
<b>Test Statistic</b>	11.574
<b>Degrees of Freedom</b>	2
<b>Asymptotic Sig. (2-sided test)</b>	.003

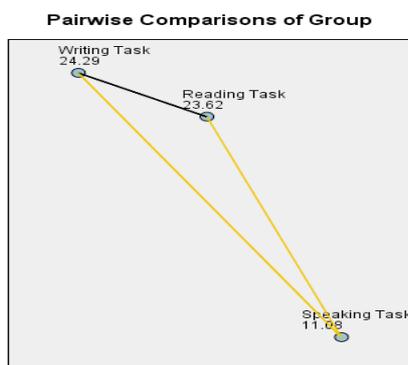
1. The test statistic is adjusted for ties.

The results of the pairwise comparison tests (Figure 4) indicated that:

- The reading task group (Med = 5) significantly outperformed the speaking task group (Med = 4) on the immediate passive vocabulary test ( $Z = 2.86, p < .05$ ).
- The writing task group (Med = 5.50) significantly outperformed the speaking task group (Med = 4) on the immediate passive vocabulary test ( $Z = 3.07, p < .05$ ).
- There was not any significant difference between the median scores of the writing and reading tasks groups on the immediate passive vocabulary test ( $Z = -.15, p > .05$ ).

The result indicated that the impacts the three tasks had on short-term active vocabulary learning was also true for short-term passive vocabulary learning. It means that the reading and writing tasks not only address Iranian EFL learners’ problem regarding active vocabulary learning, but also contribute to passive vocabulary learning, which is of importance for receptive skills such as reading. Although the knowledge gained is not active, at least a short-term passive learning took place that can be reinforced and activated later.

**Figure 5. Pairwise Comparison; Immediate Passive Vocabulary by Groups**



Each node shows the sample average rank of Group.

Sample1-Sample2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj.Sig.
Speaking Task-Reading Task	12.532	4.373	2.866	.004	.012
Speaking Task-Writing Task	13.202	4.297	3.072	.002	.006
Reading Task-Writing Task	-.670	4.207	-.159	.873	1.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same. Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

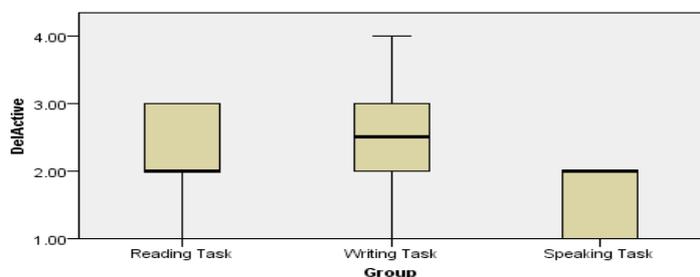
Considering the delayed active vocabulary test, Table 4 highlighted that the reading (Med = 2) and speaking (Med = 2) groups had the same median scores while the median score for the writing group was 2.5. The results of the Kruskal-Wallis test ( $H(2) = -2.37, p < .05$ ) (Figure 5) indicated that there were significant differences between the three groups' performance on the delayed active vocabulary test.

**Table 6. Descriptive Statistics; Median Scores by Groups**

Group	Delayed Activity
Reading Task	2.00
Writing Task	2.50
Speaking Task	2.00

**Figure 6. Kruskal-Wallis Test; Delayed Active Vocabulary by Groups**

Independent-Samples Jonckheere-Terpstra Test for Ordered Alternatives



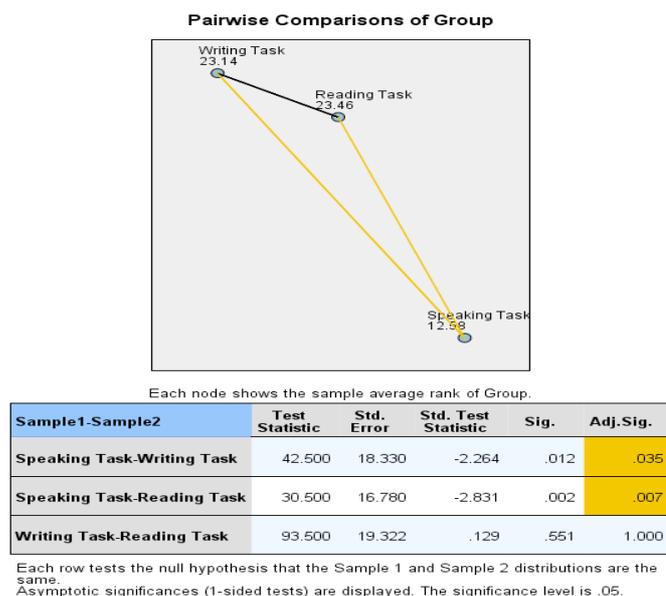
<b>Total N</b>	39
<b>Test Statistic</b>	166.500
<b>Standard Error</b>	36.409
<b>Standardized Test Statistic</b>	-2.376
<b>Asymptotic Sig. (2-sided test)</b>	.018

The results of the pairwise comparison tests (Figure 6) indicated that:

- The writing task group (Med = 2.5) significantly outperformed the speaking task group (Med = 2) on the delayed active vocabulary test ( $Z = -2.26, p < .05$ ).
- The writing task group (Med = 2.5) significantly outperformed the reading task group (Med = 2) on the delayed active vocabulary test ( $Z = -2.83, p < .05$ ).
- There was not any significant difference between the median scores of the speaking and reading tasks groups on the delayed active vocabulary test ( $Z = .125, p > .05$ ).

This demonstrated more long-term productive vocabulary learning as a result of the writing task in comparison with the speaking and reading tasks. However, the reading and speaking tasks led to the same amount of long-term productive vocabulary learning. Although Table 4.3 highlighted that using the reading and writing tasks can assist learners in reaching the aim of active vocabulary learning, learning was of a short-term nature. Long-term learning of words for productive purposes is perceived as the most valuable outcome by Iranian learners because this type of learning may contribute to writing and speaking fluently. The results indicate that writing tasks may be the best choice for keeping words active for a longer period (at least 3 weeks after learning).

**Figure 7. Pairwise Comparison; Delayed Active Vocabulary by Groups**

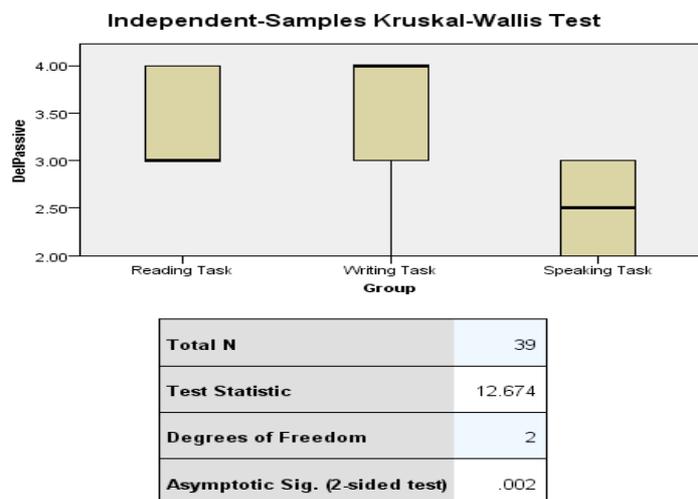


Regarding the delayed passive vocabulary test, as displayed in Table 5, the writing (Med = 4) and reading (Med = 3) groups had higher median scores on the delayed passive vocabulary test than the speaking group (Med= 2.5). Meanwhile, the results of the Kruskal-Wallis test ( $H(2) = 12.67, p < .05$ ) (Figure 7) indicated that there were significant differences between the three groups' performance on the delayed passive vocabulary test.

**Table 7. Descriptive Statistics; Median Scores by Groups**

Group	Delayed Passive
Reading Task	3.00
Writing Task	4.00
Speaking Task	2.50

**Figure 8. Kruskal-Wallis Test; Delayed Passive Vocabulary by Groups**



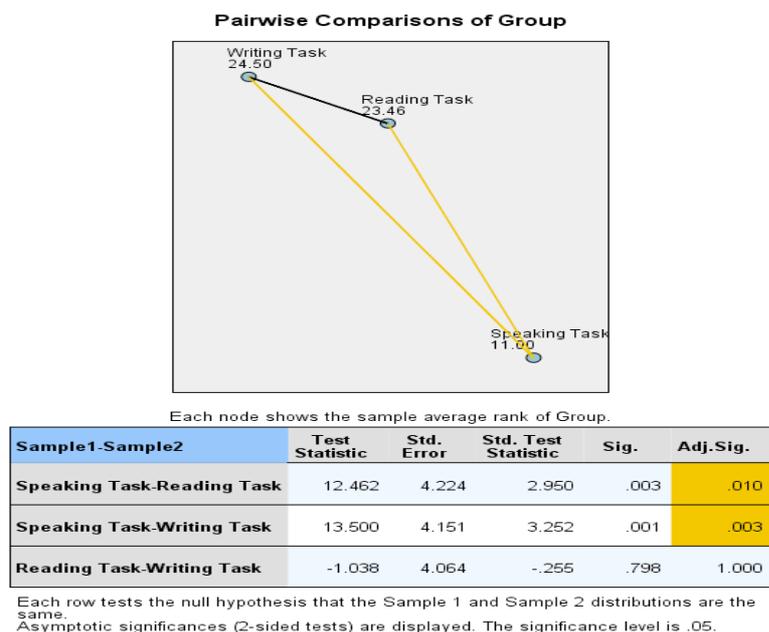
1. The test statistic is adjusted for ties.

The results of the pairwise comparison tests (Figure 8) indicated that:

- The reading task group (Med = 3) significantly outperformed the speaking task group (Med = 2.5) on the delayed passive vocabulary test ( $Z = 2.95, p < .05$ ).
- The writing task group (Med = 4) significantly outperformed the speaking task group (Med = 2.5) on the delayed passive vocabulary test ( $Z = 3.25, p < .05$ ).
- There was not any significant difference between the median scores of the writing and reading tasks groups on the delayed passive vocabulary test ( $Z = -.025, p > .05$ ).

As the result revealed, the reading and writing tasks not only led to the same amount of long-term receptive vocabulary learning (at least 3 weeks after learning), but they were superior to the speaking task in this regard. Based on the result, for learners interested in developing reading ability by expanding their passive vocabulary knowledge, writing and reading tasks can be of most assistance.

**Figure 9. Pairwise Comparison; Delayed Passive Vocabulary by Groups**



## Conclusion

This study was conducted to investigate the impact of different types of tasks with the same involvement load on short- and long-term active and passive vocabulary retention in order to compare the effectiveness of task type and task involvement load for vocabulary learning. According to the Involvement Load Hypothesis, tasks with the same involvement load (a combination of need, search and evaluation factors) should result in the same amount of vocabulary retention. However, the results of our study only partially supported this theory. Vocabulary learning was not only affected by the involvement load of the task, but also affected by the task type. The reading and writing tasks with the same involvement load resulted in similar short-term active and passive and long-term passive vocabulary learning. However, the speaking task with the same involvement load was inferior to the reading and writing tasks. In terms of involving the learners in the learning process, the reading and writing tasks resulted in more vocabulary learning than the speaking task (except for long-term active vocabulary learning). The result may be attributed to the type of processing required for oral output. Oral production of the target words in original sentences that demands simultaneous attention to semantic elaboration and verbalisation may exhaust the processing resources required for vocabulary learning. Nevertheless, this is a cautious conclusion because, according to Shumin (2002), oral production can be affected by many factors such as participants' emotions, their self-esteem and confidence, motivation for test taking, and anxiety. The impact of anxiety on speaking performance is corroborated by Woodrow (2006). If test-taking anxiety is added to speaking anxiety, oral

performance may be affected. In our study, because the participants' oral productions were recorded and measured during a twenty-minute session, the participants' emotions at the time of test taking and anxiety may have affected their performance. Although we attempted to control these variables by providing a relaxing atmosphere at the time of test taking, total elimination of anxiety was impossible, which may be a limitation of the study.

The result of the delayed active vocabulary test indicated long-term active vocabulary learning was an exception in our study. While it partially supported the Involvement Load Hypothesis by revealing that reading and speaking tasks with the same involvement indices had the same impact on vocabulary learning, it contradicted the other three tests. The writing task with the same involvement index was superior to both the reading and speaking tasks, resulting in more long-term active vocabulary learning than the reading and speaking tasks. This result is in accordance with Coomber, Ramstad and Sheets' (1986) findings. They argued that using words in original written sentences entails a type of semantic elaboration that facilitates new word learning. However, our result contradicts Barcroft's (2004) findings that sentence writing has a strong negative impact on productive vocabulary learning.

The findings can also be interpreted based on the type of processing each task entailed. Speaking a new word in an original sentence involves processing the word's meaning, relations to other words, grammatical category and pronunciation. The speaking task was less successful than other task types in vocabulary learning (except in long-term active vocabulary learning). However, the writing task, which involved attention to the word's meaning, relations to other words, grammatical category, and orthography, was more successful in learning vocabulary than the speaking task. Because the writing and speaking tasks had the same involvement indices, perhaps this result indicates that vocabulary learning depends on the type of processing a task involves rather than its involvement load. Even the reading task, which involved decoding rather than encoding a message, was superior to the speaking task (except in long-term active vocabulary learning). Perhaps oral output was the reason for lack of success. The reading and writing tasks had similar effectiveness for vocabulary learning (except for long-term active vocabulary learning). This finding may indicate that as far as the reading and writing tasks are concerned, vocabulary learning depends on the involvement load of a task rather than the type of processing.

Tasks that require learners to have oral word production in original sentences may not bring about the same amount of learning we expect based on the Involvement Load Hypothesis predictions. Since the Involvement Load Hypothesis predicts that tasks with the same involvement load (a combination of need, search, and evaluation components) result in the same amount of learning, this result conflicts the prediction by bringing the impact of task type into focus. Even the predictions may not be true for all types of vocabulary learning resulting from the learning tasks. As the results suggest, for short-term active and passive vocabulary learning and long-term passive vocabulary learning, reading and writing tasks can be more effective than speaking tasks irrespective of the involvement load of the tasks. Consequently, it seems that task involvement load does not guarantee learners' vocabulary learning and mere inclusion of need, search, and evaluation components in tasks does not lead to the expected result for all task types. When employing the Involvement Load Hypothesis in vocabulary teaching and learning, task type should also be considered.

This has some implications for vocabulary teaching. First, task involvement indices can be adjusted to meet the teaching and learning purposes. In other words, vocabulary-learning tasks should include need, search, and evaluation components to involve learners in the process of learning. When designing the learning tasks, teachers can consider these components. Learners should feel the need to learn a word even if it is imposed by a task. At this stage, including the target words in a reading passage can both impose the need to understand the word in order to comprehend the text and provide the type of processing (decoding orthographic features and linking them to a concept) that suits vocabulary learning. In addition, learners need to search for the words' meanings and to evaluate them in the context that is achieved by making original sentences.

Second, in addition to the task involvement load, task type should also be taken into account. The speaking task's lack of success in this study in comparison with the reading and writing tasks may indicate that oral learning tasks can be accompanied by other task types such as reading and writing to compensate for learners' reduced processing resource resulting from oral production. It may be more influential to encourage learners to make both oral and written original sentences with the target words

or to provide opportunities for learners to both encounter the target word in the reading passages and make original sentences. The superiority of the writing task to the reading and speaking tasks in long-term active vocabulary retention may indicate that assigning writing tasks can reinforce long-term productive vocabulary learning. Since gaining productive vocabulary knowledge that is retained for later use is the aim of EFL learners, this result may encourage them to value writing tasks.

There are also implications for vocabulary learning. Being aware of the effects of task involvement load and task type, autonomous learners can take advantage of the task type and the combination of need, search, and evaluation that are best for their vocabulary learning purposes. For instance, in order to develop long-term receptive vocabulary for reading purposes, learners can choose a reading passage with unknown words that affect understanding and ability to answer comprehension questions.

Finally, for less proficient learners, tasks such as unscrambling words or gap filling can be fruitful because to successfully complete these tasks, learners need the target words' meanings. Without having to make original sentences, they can evaluate the target words in a context of a short given sentence or make a sentence with the given words. As a result, without being proficient enough to do a reading comprehension or a writing task, unscrambling words or gap filling tasks can involve learners in the vocabulary learning process by including need, search, and evaluation components.

## Reference

- Barcroft, J. (2004). Effects of sentence writing in second language lexical acquisition. *Second Language Research*, 20(4), 303–334. doi:10.1191/0267658304sr233oa
- Coomber, J. E., Ramstad, D. A., & Sheets, D. R. (1986). Elaboration in vocabulary learning: A comparison of three rehearsal methods. *Research in the Teaching of English*, 20(3), 281–293.
- Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11(6), 671–684. Retrieved from [http://dx.doi.org/10.1016/S0022-5371\(72\)80001-X](http://dx.doi.org/10.1016/S0022-5371(72)80001-X)
- Craik, F. I., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, 104(3), 268–294.
- Eckerth, J., & Tavakoli, P. (2012). The effect of word exposure frequency and elaboration of word processing on incidental L2 vocabulary acquisition through reading. *Language Teaching Research*, 16(2), 227–252. doi:10.1177/1362168811431377
- Field, J. (2006). *Psycholinguistics*. New York, NY: Routledge.
- Hulstijn, J. H., & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning*, 51(3), 539–558. doi:10.1111/0023-8333.00164
- Jing, L., & Jianbin, H. (2009). An empirical study of the involvement load hypothesis in incidental vocabulary acquisition in EFL listening. *Polyglossia*, 16, 1–11.
- 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. doi:10.1177/1362168808089922
- 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.1467-9922.2011.00644.x
- Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, 22(1), 1–26. doi:10.1093/applin/22.1.1
- Laufer, B., Elder, C., Hill, K., & Congdon, P. (2004). Size and strength: Do we need both to measure vocabulary knowledge? *Language Testing*, 21(2), 202–227. doi:10.1191/0265532204lt277oa
- Martínez-Fernández, A. (2008). Revisiting the involvement load hypothesis: Awareness, type of task and type of item. In M. Bowles, R. Foote, S. Perrián, & R. Bhatt (Eds.), *Selected proceedings of the 2007 second language research forum* (pp. 210–228). Somerville, MA: Cascadilla Proceedings Project.
- Pichette, F., De Serres, L., & Lafontaine, M. (2011). Sentence reading and writing for second language vocabulary acquisition. *Applied Linguistics*, 33(1), 66–82. doi:10.1093/applin/amr037

- Rott, S. (2007). The effect of frequency of input enhancements on word learning and text comprehension. *Language Learning*, 57(2), 165–199. doi:10.1111/j.1467-9922.2007.00406.x
- Shumin, K. (2002). Factors to consider: Developing adult EFL students' speaking abilities. In J. C. Richards & W. A. Renandya (Eds.), *Methodology in language teaching: An anthology of current practice* (pp. 204–211). Cambridge, England: Cambridge University Press. doi.org/10.1017/CBO9780511667190.028
- Sun, H. (2014). The effects of exposure frequency and contextual richness in reading on Chinese EFL learners' vocabulary acquisition. *Chinese Journal of Applied Linguistics*, 37(1), 86–106. doi:10.1515/cjal-2014-0006
- Vidal, K. (2011). A comparison of the effects of reading and listening on incidental vocabulary acquisition. *Language Learning*, 61(1), 219–258. doi:10.1111/j.1467-9922.2010.00593.x
- Woodrow, L. (2006). Anxiety and speaking English as a second language. *RELC journal*, 37(3), 308–328. doi:10.1177/0033688206071315