

L2 Idioms Recall: On the Effectiveness of Task Focus and Involvement Load¹Maryam Mousavi²Abbas Ali Zarei Mavini*³Saeideh Ahangari

Research Paper

IJEAP- 2106-1728

DOR: [20.1001.1.24763187.2021.10.2.5.1](https://doi.org/10.1001.1.24763187.2021.10.2.5.1)

Received: 2021-05-08

Accepted: 2021-07-01

Published: 2021-07-02

Abstract

Finding effective ways to improve L2 idioms recall has been a long-standing concern of many practitioners. This study compared the effects of meaning-focused and form-focused tasks with varying involvement load indices on EFL learners' recall of L2 idioms. One-hundred eighty intermediate level EFL learners participated in this study and were randomly assigned to one of the six experimental groups. In three form-focused groups, the participants received multiple-choice, sentence-completion and sentence-making tasks with involvement indices of 2, 3, and 4, respectively. In three meaning-focused groups, they were asked to perform summary writing, writing with glossary, and writing without glossary with involvement degrees of 2, 3, and 4, respectively. A 30-item test in fill-in-the-blank format was administered to all groups to assess their recall of idioms after the treatment. To analyze the data, a two-way ANOVA and a series of independent-samples t-tests were used. The results showed that the tasks with higher degrees of involvement were more efficient on the recall of idioms. The results also revealed that although form-focused tasks were significantly more effective than meaning-focused tasks at lower loads of involvement, at higher involvement load, the difference between form-focused and meaning-focused tasks was not significant. The findings of the present study can have theoretical as well as pedagogical implications for learners, teachers, textbook designers, and curriculum developers.

Keywords: Form-focused tasks, Idioms, Involvement Load, Meaning-focused tasks

1. Introduction

Vocabulary is an essential element in learning a second language. Having a vast knowledge of vocabulary makes language learners competent in both understanding and producing a language (Kıvrak & Uygun Gökmen, 2019). Richards and Renandya (2002) state that vocabulary is an open-ended component of language, which is presented in a clear, well-organized package. Idioms are regarded as a considerable component of vocabulary. Wray and Perkins (2000) opine that to be able to read a variety of books, magazines, and stories, having a considerable knowledge of idiomatic expressions is essential. Hence, learners need to be capable of understanding and using them. Nonetheless, learning idioms and improving the ability to use them appropriately is not simply attainable. Even for native speakers, learning idioms may pose a bit of challenge (Hinkel, 2017). Among several suggestions made by many researchers and teachers to facilitate idioms learning is employing tasks. However, the use of different task types has caused its own share of controversy.

One issue is related to the focus of task. Whereas the proponents of Noticing Hypothesis support employing form-focused tasks, based on which attention is a necessary condition for language learning (Schmidt, 1990), others believe more in Krashen's (1985) Input Hypothesis and advocate the use of meaning-focused tasks in which learners pay more attention to meaning. Furthermore, based on the Involvement Load Hypothesis (ILH) proposed by Laufer and Hulstijn

¹ PhD candidate in TEFL, mousavimaryam69@yahoo.com; Department of English, Islamic Azad University, Tabriz Branch, Iran.

² Associate professor (Corresponding Author), a.zarei@hum.ikiu.ac.ir; Department of English, Imam Khomeini International University, Qazvin, Iran.

³ Assistant Professor of TEFL, ahangari@iaut.ac.ir; Department of English, Tabriz Branch, Islamic Azad University, Tabriz, Iran.

(2001), it is claimed that it is the involvement level of a task, not its focus, which determines its effectiveness. The ILH predicts that when tasks have identical load, whether form-focused or meaning-focused, they can be equally beneficial; otherwise, tasks with higher involvement indices will be more effective. Several researchers (e.g., Naserpour et al., 2020; Soleimani & Rahmanian, 2015; Zarei & Moftakhari Rezaei, 2016) have investigated the effect of involvement load. However, they have used different task types and task orientations, focused on different dependent variables (like vocabulary, collocations, etc, and have often reported conflicting results. The present study tried to resolve this issue by addressing the following questions:

Research Question One: Are meaning-focused and form-focused tasks with different involvement loads differentially effective on L2 idioms recall?

Research Question Two: Are meaning-focused and form-focused tasks with identical involvement loads differentially effective on L2 idioms recall?

2. Literature Review

2.1. Idioms

Using idiomatic language is an economical approach to communicating meaning (Liontas, 2003). Idioms are a kind of multiple and fixed units of words that have both literal and figurative meanings, and the figurative meaning of an idiom cannot be easily understood from the literal meaning of the individual parts (Haward, 1998; Irujo, 1986). Thawabteh (2011) states that idioms are considered to be one of the vital aspects of all languages and not being able to understand and use them correctly causes both linguistic and cultural problems for learners. And these problems may impede efficient communication. McCarthy and O'Dell (2017) refer to seven forms of idioms as follows: 1) Whole clause or sentence '*to cut a long story short*', 2) Prepositional phrase '*in a blink of an eye*', 3) Verb + object/component '*kill two birds with one stone*', 4) Compound '*a bone of contention*', 5) Simile '*as busy as a bee*', 6) Binomial '*rough and ready*', 7) Trinomial '*cool, calm and collected*'. As Moreno (2011) states, because idioms have high frequency, teachers must pay attention to teaching them, and it has always been a major concern of language teachers and researchers to design suitable tasks for the instruction of idioms.

2.2. Tasks

Prabhu (1987) defines task as an activity in which learners are required to arrive at an outcome using some given information and processing of thought. Different classifications of tasks exist in the literature. Regarding focus, there are form-focused and meaning-focused tasks (Spada & Lightbown, 2008). Form-focused tasks are those in which learners attend to specific forms of language by utilizing different strategies (Ellis, 2003). Based on Noticing Hypothesis, Schmidt (1990) claims that attention is a necessary condition to acquire a language. Moreover, Spada and Lightbown (2008) refer to isolated form-focused activities suggesting that if an activity is separated from communicative context, it will be more effective. On the other hand, meaning-focused tasks are those in which learners pay attention mainly to meaning (Ellis, 2015), and no strategies are used to draw learners' attention to particular linguistic features (Golshan, 2015). Krashen (1985) argues that merely exposure to input, without consciousness, is sufficient for learning. Meaning-focused tasks provide such a condition for language learning.

A vast body of research has investigated the effect of meaning and form focused tasks on vocabulary learning (Celik, 2019; Khonarmi & Roostaei, 2014; Zarei & Moftakhari Rezaei, 2016). For instance, Zarei and Moftakhari Rezaei (2016) examined the effectiveness of type of tasks (meaning- and form-focused) and task orientation (input- and output-oriented) on vocabulary learning. Regarding task type, meaning-focused tasks turned out to be more beneficial than form-focused tasks. Moreover, in accordance with Swain's (2000) Output Hypothesis, output-oriented tasks were more efficient when they were meaning-focused, whereas input-oriented tasks were effective when they were form-focused. In addition, Szudaeski (2012) concluded that a combination of form-focused and meaning-focused tasks provided better opportunities for learning collocations than meaning-focused tasks alone.

Nonetheless, a number of studies highlight the effectiveness of form-focused tasks compared to meaning-focused ones (Ahour & Ghorbani Shemshadsara, 2015; Lan & Wu, 2013; Celik, 2019). For instance, Elgort et al. (2016) compared the effects of meaning-focused and form-focused strategies on contextual vocabulary learning of Chinese and Dutch learners. The results showed better learning outcomes for form-focused elaboration. In another study, Celik (2019) compared the effects of meaning-focused and form-focused instruction on vocabulary. The results showed that the group with form-focused instruction outperformed the meaning-focused group. In addition, several studies (Fakhrzadeh & Yazdanjoo, 2020; Naseri & Khodabandeh, 2019) have shown the effectiveness of input enhancement strategy as a kind of form-focused technique. Although based on the above-mentioned studies, it was revealed that task focus affects vocabulary learning differently, Laufer and Hulstijn (2001) claim that the only determining factor regarding the effectiveness of a task is the degree that it involves learners, regardless of task type or focus.

2.3. *Involvement Load Hypothesis*

The Involvement Load Hypothesis (ILH) was proposed by Laufer and Hulstijn (2001) to answer the question of why some tasks are more effective than others in L2 vocabulary acquisition. Need, search and evaluation are three constructs of task-induced involvement load. The need factor is a non-cognitive, motivational dimension of involvement. Laufer and Hulstijn (2001) define it as "a drive to comply with the task requirements" (p. 14). When need is motivated externally, it is moderate (e.g., a teacher asks a learner to perform a task) and when it is motivated intrinsically, it is strong (e.g., a learner aspires to learn something new). Search is learners' attempt to determine the meaning of unknown words by using a dictionary or consulting other authorities (e.g., an expert). Evaluation is the learners' endeavor to make a comparison of words to see whether a word is appropriate in a given context. Evaluation is regarded moderate when learners try to choose between different words. Evaluation is strong when the learners need to make decisions on the combination of a new word with additional words to generate a new context. These constructs vary in their degree of involvement needed in doing a task: when there is no involvement (-) the score is 0, moderate involvement (+) receives the score 1, and 2 is assigned to the strong involvement (++). The need and evaluation components can be either moderate or strong while the search construct can be present or absent.

A number of recent studies (Ahmadi Fatalaki, 2014; Alcaraz-Mármol & Almela, 2013; Amini & Maftoon, 2017; Jahangard, 2013; Jing & Jianbin, 2009; Keyvanfar & Badraghi, 2011; Karalik & Merç, 2016; Naserpour et al., 2020; Pourakbari & Biria, 2015; Soleimani & Rahmadian, 2015; Zou, 2016) have submitted ILH to empirical investigation. One of the early studies on ILH was the one in which Hulstijn and Laufer (2001) presented three tasks simultaneously to students. They included reading texts with glosses, reading with gap-fill and composition writing. Strong empirical support was provided for the ILH, and the composition writing task with the highest involvement index was the most efficient. In another study, Teng (2015) examined the effects of three translation tasks on vocabulary learning. Task 1 included translation (index = 1), task 2 composed of translation and fill-in (index = 2), and task 3 involved translation with sentence writing (index = 3). The findings revealed that both vocabulary learning and retention were highest in task 3, followed by task 2, and finally task 1. Similarly, the results of the study by Alavinia and Rahimi (2019) were in accordance with ILH prediction.

Several studies (Hassanzadeh, 2016; Sarani et al., 2013; Tajeddin & Daraee, 2013) have examined the effects of involvement load as well as orientation and type of tasks on vocabulary learning. For instance, Kaivanpanah, et al. (2020) examined the effects of output-based and input-based tasks with identical and different involvement degrees on L2 vocabulary learning. The results revealed that the input- and output-based tasks with the same involvement load were effective on vocabulary learning. However, for the group who received higher involvement load tasks, output-based tasks were more effective. The findings did not fully support the ILH. In another study, Bao (2015) investigated the effects of task type and involvement load on EFL learners' vocabulary learning. Among four experimental groups, three groups received various tasks with identical involvement degree of 2, while one of the experimental groups received a task with a higher load of

3. The results were not in favor of the ILH and showed that the learners in all output task groups performed noticeably better than the control group. The findings showed that involvement load is not the only factor determining tasks effectiveness. Moreover, the tasks with the same level of involvement affected vocabulary learning differently. Additionally, the findings of Un-udom's (2018) study were not in accordance with the ILH predictions.

Considering the above review, it is evident that several studies have investigated the effects of both task type and involvement degree on vocabulary gain (e.g., Kırak, C., & Uygun Gökmen, 2019; Un-udom, 2018; Zou, 2016). However, they have used different task types and focused on general vocabulary learning. Little research, if any, has examined the effects of task focus and involvement load on idioms learning. To bridge this gap, this study aims at investigating the effects of meaning-focused and form-focused tasks with different levels of involvement on idioms recall.

3. Methodology

3.1. Participants

The participants of this study were 211 intermediate level learners (according to their educational profile) studying in two language schools of Safir Gostar and Shokouh in Abhar, Iran. They were male and female students studying Top Notch books. Based on availability, they were selected via convenience sampling. Before the study, the reading and writing sections of PET (Preliminary English Test) were utilized to check the participants' language ability. After administering the test, 31 learners were excluded from statistical analysis because they scored more than one standard deviation below or above the mean. Hence, 180 students (90 females and 90 males) remained. They aged from 16 to 32. Some were university students with different fields of study and some were high school students. Their mother tongue was Persian or Turkish.

3.2. Instruments and Materials

3.2.1. PET

To homogenize the participants, the writing and reading parts of the PET were utilized for practical reasons. The reading part consisted of five sections which included 35 multiple-choice items. The writing part included three sections; part one consisted of paraphrasing, while the other two sections consisted of productive writing tasks. Five texts were included in part one of the reading; the participants had to answer following three multiple-choice items. Part two contained the descriptions of some people and TV programs. The participants were told to decide which program is suitable for each person. In part three, the students received true-false exercises based on a reading text. Part four included a reading text followed by 5 multiple-choice items. Part five contained a cloze passage. In part one of the writing section, the students had to paraphrase five sentences. In part two, they had to write about their weekend. In part three, the students had to choose from among the two topics and write on one of them, employing 100 words. 90 minutes were allotted for this test. Although PET has been used mostly in different EFL/ESL contexts, to re-estimate the reliability of the sub-test, the KR-21 formula was utilized. It turned out to be .78.

3.2.2. Pre-test

Prior to the treatment sessions, a teacher-made pre-test was given to the participants to assess whether or not they had any prior knowledge of the selected idioms for the treatment. It contained 135 sentences, each of which included one bold-faced and underlined target idiom. The learners were expected to provide their Persian definitions. Eighty-five minutes were allocated to this pre-test. To check the reliability of this test, the KR-21 formula was used; the result was .80. The target idiomatic expressions were selected from *English Idioms in Use* (McCarthy & O'Dell, 2017), designed for intermediate level students. The main purpose of this pretest was to make sure that the target idioms were not already familiar to the learners, and to remove the familiar ones from the posttest.

3.2.3. Teaching materials

The researchers designed six idiom tasks for the purpose of this study. Three form-focused tasks induced the involvement levels of 2, 3, and 4, and three meaning-focused tasks induced the involvement indices of 2, 3, and 4. It is worth noting that the involvement indices of these tasks were specified on the basis of the ILH (Laufer & Hulstijn, 2001). The three form-focused groups received the following tasks.

Multiple-choice task: The participants in this group received some complete, isolated sentences. Input enhancement as a form-focused instruction technique was used and the target idioms included in the sentences were bold-faced. Each sentence contained one glossed idiom at the end of the sentence. The students were expected to select the right definition or synonym of the idioms from among four options. Then, the correct answers were written on the board by the instructor. This task had moderate need, search was absent, and evaluation was moderate. Thus, the involvement index was $2(1 + 0 + 1)$.

Sentence completion tasks: The students were exposed to the same sentences, but the idioms were eliminated from the sentences, and the learners were asked to find the correct idioms in the sentences and fill the blanks. The Persian equivalents of the target idioms and a part of their English equivalents were provided in bold face as cues to prevent the students from supplying idioms that might fill the gaps correctly without being the intended idioms. The need and evaluation of this task was moderate, and search was present. Therefore, the numerical involvement index of this task was $3(1 + 1 + 1)$.

Sentence making tasks: In this group, the participants were exposed to the same sentences in which the selected idioms were bold-faced. However, in this task, the sentences were not glossed. The participants had to read the sentences and then search for idioms using a dictionary. Then, they had to create new sentences using the target idioms. Next, the teacher corrected the sentences, and wrote the best ones on the board. This task had no search, moderate need and strong evaluation. Hence, its involvement index was $4(1 + 1 + 2)$.

The three meaning-focused groups received the tasks below.

Summary writing tasks: In this group, first the learners were told to read a glossed passage with the average length of 150 words, and then they had to summarize the text and incorporate all the target idioms introduced in the passage in their summary. The L1 translation of each idiom was written in parentheses beside each idiom. The instructor provided feedback for them and corrected all idiomatic errors. This task included moderate need and evaluation but no search. Therefore, its involvement load was $2(1 + 0 + 1)$.

Writing task with glossary: The participants were required to write a paragraph and utilize all the selected idioms, after reading an identical text. Within the text, the Persian definition of each idiom was provided next to it. They were told that the average length of the paragraph should be 150 words. The teacher then corrected idiomatic errors. The involvement components of this task were no search, moderate need and strong evaluation. So, the involvement level was $3(1 + 0 + 2)$.

Writing task without glossary: Similar to writing with glossary task, the students in this group had to write a paragraph, after reading the text. The average length of the paragraph was expected to be 150 words. However, in this task, the passage was not glossed, and the participants had to look up the meaning of the idioms in a dictionary. This task had moderate need, search was present, and evaluation was strong. Therefore, its involvement load index was $4(1 + 1 + 2)$.

3.2.4. Post-test

A post-test including 30 fill-in-the-blanks items was administered to check the students' recall of idioms after the treatment. The Persian definition of the idioms was provided in parentheses, and the learners were told to write their English equivalents. Fifty minutes were allotted for the post-test. Since the researchers constructed some of the tasks and the post-test, their contents were attentively checked by a panel of experts to make sure of their content validity. Additionally, the

KR-21 formula was used to estimate the reliability of the post-test. The reliability index was estimated to be .81.

3.3. Procedure

Initially, 211 intermediate level participants were chosen through convenience sampling based on availability from Safir Gostar and Skokouh institutes in Abhar, Iran. To homogenize the participants, a sample of PET was administered. For practicality reasons, the speaking and listening sections were not used. Those participants who scored between $\pm 1SD$ from the mean formed the main participants. One-hundred and eighty learners were left. Prior to the treatment, the teacher-made idiom pre-test was administered to ensure that the students had no knowledge of the new idioms. The idioms known to more than five percent of the students (19 idioms) were not used in the post-test.

The experimental intervention lasted 17 sessions, of which two sessions were devoted to the administration of the PET and the pre-test, 14 sessions to the intervention, and the last session to the post-test. The learners had been studying Top Notch text books (Saslow & Ascher, 2011) in the institutes. In both institutes, the class time for each session was 90 minutes. At the beginning of each session, around 45 minutes were allotted for the experiment. The students were randomly assigned to one of the experimental groups.

In the form-focused groups, the learners received the following tasks: Multiple-choice task (Task A); Sentence completion task (Task B); and Sentence making task (Task C) with the involvement indices of 2, 3, and 4. Isolated form-focused tasks were utilized and the idioms were included in isolated sentences to draw the participants' attention to them (Spada & Lightbown, 2008). Simply put, they were not contextualized. Additionally, to draw the students' attention to the intended idioms, they were bolded.

Each session, in the three meaning-focused groups, the participants were provided with a passage of an average length of 150 words which included an average of 10 intended idioms. The students had to perform one of the tasks in each group, after reading the same text: summary writing (Task D), writing with glossary (Task E); and writing without glossary (Task F) with involvement degrees of 2, 3, and 4, respectively. The selected idioms were contextualized, but not bolded, in these groups.

During the treatment, if any difficulty arose in the comprehension of the idioms or words while reading the texts or the sentences, the instructors offered help. To finish their tasks, all participants in the six groups had the same amount of time (45 minutes). In task A, the learners had to read the glossed sentences in which new idioms were bold-faced. They had to select the best definition or synonym of the idioms from among four options that were provided in multiple-choice format. In task B, the students received the same sentences; however, the idioms were deleted and only one part of each was provided as a cue. Moreover, the Persian definitions of the idioms were written in the sentences to assist the participants to find the correct idioms. In task C, the selected idioms were highlighted in bold face and included in non-glossed sentences. The students were required to use dictionaries and search for new idioms and then create new sentences using them. In task D, the learners were provided with a glossed text including the intended idioms which were not bolded. The participants had to read the text and summarize it. In task E, the students received the same glossed texts. After reading it, they were asked to write a new paragraph using new idioms. Task F resembled task E, but the passage was not glossed, and the participants were first told to look up the meaning of the target idioms and then generate a new paragraph.

3.4. Data Analysis

A two-way ANOVA procedure and three independent samples t-tests were used to answer the questions.

4. Results

Before doing the main statistical analyses, to ascertain learners' homogeneity, the results of the PET were analyzed. No outlier or extreme score was observed, based on the boxplot of the scores. Furthermore, the results of Kolmogorov-Smirnov (statistic $_{(180)} = .065$, Sig. = .39, $P > .05$) and Shapiro-Wilk tests (statistic $_{(180)} = .988$, Sig. = .137, $P > .05$) revealed that the distribution of the scores was normal.

The first research question was intended to investigate the differences in the effectiveness of form-focused and meaning-focused tasks with varying involvement indices on the recall of L2 idioms. For this purpose, a two-way ANOVA was utilized. First, its assumptions were tested. The results of Kolmogorov-Smirnov test ($KS_{(180)} = .31$, $p > .05$) and Levene's test ($F_{(5, 174)} = .143$, $p > .05$) showed that the assumptions of normality and homogeneity of variances were both met. Then, descriptive statistics on the recall test of L2 idioms were summarized (Table 1).

Table 1: Descriptive Statistics for Idioms Recall

Task-focus	Involvement load	Mean	Std. Deviation	N
Form-focused	IL=2	23.433	1.654	30
	IL=3	24.733	1.595	30
	IL=4	24.833	1.984	30
	Total	24.333	1.848	90
Meaning-focused	IL=2	21.433	1.755	30
	IL=3	22.500	2.417	30
	IL=4	24.433	1.654	30
	Total	22.788	2.315	90
Total	IL=2	22.433	1.968	60
	IL=3	23.616	2.322	60
	IL=4	24.633	1.822	60
	Total	23.561	2.227	180

Based on Table 1, it appears that form-focused tasks are more effective than meaning-focused ones. However, the mean scores of the learners in the form-focused and meaning-focused groups with various loads of involvement are different. The form-focused task groups with involvement indices of 4 and 3 have the highest mean scores, followed closely by the meaning-focused task group with the involvement index of 4, followed by form-focused task group with involvement load of 2. The lowest means belong to meaning-focused tasks with involvement levels of 2 and 3, respectively. The differences among the groups are illustrated in Figure 1.

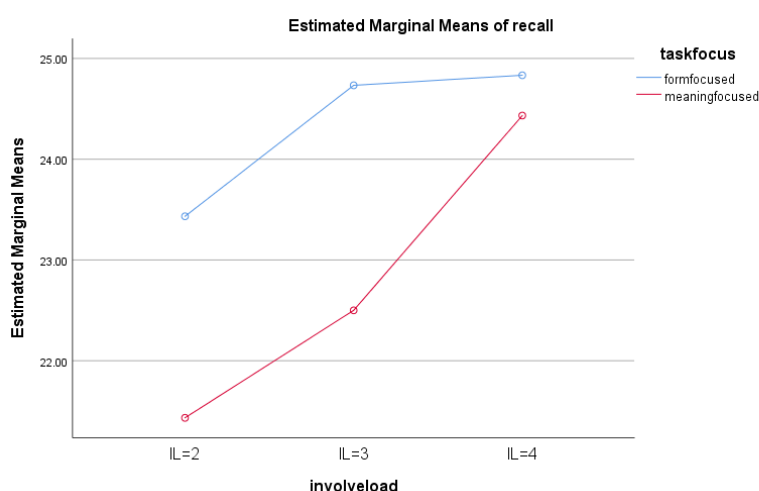


Figure 1. Means Plot of Idioms Recall Post-test

To investigate the possible interaction effect and the main effects of the independent variables, the results of the tests of between-subjects effects were checked (Table 2).

Table 2: Tests of Between-subjects Effects for Idioms Recall

	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	282.694 ^a	5	56.53	16.24	.000	.31
Intercept	99922.672	1	99922.67	28708.03	.000	.99
Task-focus	107.339	1	107.33	30.83	.000	.15
Involvement load	145.478	2	72.73	20.89	.000	.19
Task-focus * involvement load	29.878	2	14.93	4.29	.015	.04
Error	605.633	174	3.48			
Total	100811.00	180				
Corrected Total	888.328	179				

Table 2 shows a significant interaction between the effects of task focus and involvement load on idioms recall, ($F_{(2, 174)} = 4.29, p < .05$). In other words, the difference between meaning and form - focused tasks in terms of the effect of involvement level on idioms recall is statistically significant. The index of the partial eta squared indicates that the interaction accounts for .04% of the total variability among the groups. Furthermore, there is a meaningful difference between meaning-focused and form-focused tasks ($F_{(1,174)} = 30.83, p < .0005$) with the form-focused groups generally performing better than the meaning-focused groups on the idiom recall post-test. Besides, there are meaningful differences among the three indices of involvement ($F_{(2,174)} = 20.89, p < .0005$). Additionally, task focus and involvement load had the partial eta squared values of .15 and .19, respectively, showing large effect size, according to Cohen (1988). The Tukey HSD test was run to locate the differences among the three levels of involvement load.

Table 3: Tukey Test Results for the Effects of Involvement Loads on Idioms Recall

(I) involvement load	(J) involvement load	Mean Difference (I-J)	Sig.
IL=2	IL=3	-1.1833*	.001
	IL=4	-2.2000*	.000
IL=3	IL=4	-1.0167*	.003

Table 3 shows that there are statistically significant differences between all three involvement loads. These results suggest that the higher the involvement load of a task, the more effective it will be on idioms recall. However, because there was a significant interaction between task focus and involvement load, we cannot claim with certainty whether the difference between form-focused or meaning-focused tasks is the same across involvement load indices. That is why another research question was formulated to address this issue.

The second research question was intended to compare the effectiveness of meaning-focused and form-focused tasks on the participants' recall of idioms while the involvement load was controlled. To this end, the scores of the learners on the idioms recall test in form- and meaning-focused groups were compared at three different involvement load indices. Table 4 presents the summary of the descriptive statistics.

Table 4: Descriptive Statistics of Recall Post-test

Involvement load	Task-focus	N	Mean	Std. Deviation
IL=2	recall	30	23.43	1.65
	Form-focused			
IL=3	recall	30	21.43	1.75
	Meaning-focused			
IL=4	recall	30	24.73	1.59
	Form-focused			
IL=4	recall	30	22.50	2.41
	Meaning-focused			
IL=4	recall	30	24.83	1.98
	Form-focused			
IL=4	recall	30	24.43	1.65
	Meaning-focused			

Table 4 indicates that the mean scores are different among the groups that received tasks with different focuses and identical involvement degrees. Based on Table 4, it can also be seen that at involvement level of 2, the form-focused group has a higher mean score than the meaning-focused one. Similarly, at involvement load 3, the higher mean score belongs to the form-focused group. At involvement level of 4, the mean score of the form-focused group was higher than the meaning-focused group, although the difference was not significant.

Three independent-samples t-tests were run to see whether or not the observed difference between the form and meaning-focused groups at each level of involvement load is significant. The results of the t-tests are shown in Table 5. Based on Table 5, Levene's test for equality of variances showed that, at involvement load levels 2 and 4, the variances of the form- and meaning-focused groups were equal on the idioms recall post-test. Hence, the first rows were checked. At involvement load 3, since the assumption was not met, the lower row had to be checked.

Table 5: The t-test Results of Recall Post-test for Form and Meaning-focused Tasks

Involvement load		Levene's Test for Equality of Variances		t-test for Equality of Means			
		F	F Sig.	t	df	Sig.	Mean Dif.
IL=2	Equal variances assumed	.011	.917	4.54	58	.000	2.00
	Equal variances not assumed			4.54	57.79	.000	2.00
IL=3	Equal variances assumed	4.00	.050	4.22	58	.000	2.23
	Equal variances not assumed			4.22	50.24	.000	2.23
IL=4	Equal variances assumed	.069	.794	.84	58	.400	.400
	Equal variances not assumed			.84	56.18	.400	.400

In Table 5, the obtained t-value and significance level indicate a meaningful difference between the two groups with the involvement load of two ($t_{(1,58)} = 4.54$, $p < .0005$). Considering the two groups with similar involvement level of three, the result ($t_{(1,58)} = 4.22$, $p < .0005$) also indicates a meaningful difference between the meaning-focused and form-focused tasks. However, there is no considerable difference between the two groups with the involvement index of four ($t_{(1,58)} = .84$, $p > .05$). Therefore, it can be concluded that meaning-focused and form-focused tasks are differentially effective at lower involvement indices. Furthermore, the index of the strength of association for involvement loads of two, three and four were ($\eta^2 = .26$), ($\eta^2 = .23$), and ($\eta^2 = .01$), respectively, implying that 26%, 23%, and .01% of the total variability observed between the groups can be accounted for by task focus.

5. Discussion

One of the findings of the present study was that tasks inducing higher involvement index led to better recall of L2 idioms. This result is in line with the findings of several studies (Ahmadi Fatalaki, 2014; Alavinia & Rahimi, 2019; Naserpour et al., 2020; Zou, 2017), who showed that

more involving tasks resulted in better post-test scores. In a similar vein, Jing and Jianbin (2009) reported that with the increase of the involvement load index, vocabulary learning improved. However, unlike this study, they examined the impact of task-induced involvement load on lexical acquisition in listening comprehension.

Similarly, the finding of Jahangard (2013) is in agreement with that of this study. He reported that dictionary checking plus writing task yielded better results compared to the other two tasks. Soleimani and Rahmanian (2015) concluded that in similar types of tasks, i.e., input-oriented ones, the gap-filling task with a higher load of involvement led to better retention of words than the reading comprehension task with a lower index. Unlike our study in which the target idioms were bolded only in the form-focused tasks, in their study, all target words were bolded in both tasks. The results of Teng (2015) also confirm the finding of this study. However, he compared the effectiveness of task involvement load on vocabulary achievement through translation tasks, and, unlike the present study in which participants were at intermediate proficiency level, those in Teng's study were at elementary level.

Conversely, Un-udom (2018) found that higher degrees of involvement did not result in vocabulary gain. He reported that the learners in the sentence making task, which was form-focused (IL=3) performed better than those in the sentence creation task, which was not form-focused (IL=4). This result can be justified according to Schmidt's (1990) Noticing Hypothesis, based on which conscious attention is a necessary factor that facilitates language learning. Moreover, in the study of Bao (2015) no positive effect was found in favour of the ILH in vocabulary acquisition. The participants who received the definition task (Index=2) outperformed those who performed the writing tasks (Index=3). This contrast may be due to the evaluation component of the tasks. The writing task induced strong evaluation which was cognitively more demanding than the definition task which induced moderate evaluation. The students might have not been cognitively ready for more demanding tasks. Furthermore, similar to the finding of our study, the results of that study revealed that various tasks with the identical involvement index of 2 had different effects on vocabulary achievement. Additionally, in contrast with the findings of this study, Alcaraz-Mármol and Almela (2013) reported that the writing accompanied by dictionary use task (IL= 4) was less effective than the writing plus marginal glosses task (IL= 3). The reason for this contrast can be the fact that the participants were at elementary level and may not have been cognitively prepared to receive higher loaded tasks, that is, the ones with a load index of 4. Another reason may be the small sample size of the study including only 28 students.

Another finding of the present study was that the involvement load of a task is not the only factor determining efficiency; task focus is another. A significant interaction between task focus and task involvement load implies that at lower levels of involvement, form-focused tasks are more conducive to idioms recall. However, regarding higher involvement load, no significant difference was found between the performance of the learners in form-focused and meaning-focused groups. With regard to tasks with lower involvement indices of 2 and 3, the participants in the form-focused task groups showed significantly better performance than those in the meaning-focused ones. Regarding form-focused and meaning-focused tasks with the identical involvement load of 4, although the form-focused task group performed slightly better than the meaning-focused group, no significant difference was found.

The finding that at lower loads of 2 and 3, form-focused groups performed better than meaning-focused ones substantiate the Noticing Hypothesis that is considered as theoretical underpinning of FFI, based on which learners must consciously attend to forms as well as their meanings in the input to change it into intake (Schmidt, 1990). Another possible reason for this finding may be that in meaning-focused tasks, the target idioms were non-salient and the participants did not pay enough attention to them at lower loads of involvement. These findings can also be justified in light of Ellis's (2015) claim that form-focused instructions and tasks are based on information processing models, according to which students face difficulty while attending to form and meaning at the same time. While paying attention to meaning, learners need top-down strategies, whereas attending to form requires bottom-up processing. Form-focused tasks create a

condition for learners to distance themselves from merely focusing on meaning and provide opportunities for them to attend to particular forms of language as well.

However, unlike the tasks with involvement indices of 2 and 3, there was no significant difference between the performances of the learners who received different tasks with the same involvement load of 4. The reason may be that at higher levels of involvement, due to the demanding nature of tasks, learners may have to pay more attention to the target forms, regardless of task type. Another possible reason to explain this difference may be that, in the present study, a strong component of evaluation was used in both of these tasks. As Teng (2015) claims, while teaching vocabulary, special attention must be paid to the strong acts of evaluation. This way, students have opportunities first to consider the differences between words, and then use them in new contexts productively. Additionally, since both tasks with involvement load 4 (sentence making and writing without glossary) were output-oriented; this finding can be explained by Out Hypothesis, based on which Swain (2000) claims that when learners produce a language, they acquire it more effectively. Besides, the positive effect of highly loaded meaning-focused tasks in the study can be explained on with Input Hypothesis, based on which merely comprehensible input is essential for students to acquire a language (Krashen, 1985), i.e., consciousness is not necessary.

Similarly, the findings of Hassanzadeh (2016) corroborate those of this study. He concluded that not only the involvement load of a task but also modality is effective in determining task efficacy. Furthermore, Kaivanpanah et al. (2020) reported that, like involvement load, task type was important, and output tasks turned out to be more effective than input tasks. Further support for this finding comes from Sarani et al. (2013), who compared the effectiveness of involvement load and task type. However, they used input and output tasks with load indices of 1, 2, and 3, whereas in our study, the tasks were either form-focused or meaning-focused with load indices of 2, 3, and 4. Similar to our findings, Tajeddin and Daraee (2013) showed that tasks type was an effective factor in vocabulary retention.

In another study, Zou (2017) showed that in spite of inducing the same involvement load, the composition writing task was more effective than the sentence writing one. This finding is in contrast with that of our study in which the sentence writing task was better than the paragraph writing task, both inducing the same involvement index. The reason for this can be the evaluation component of the tasks. Regarding Zou's (2017) study, the evaluation component of sentence and composition writing was strong and very strong, respectively. However, in our study, the evaluation component was strong in both tasks. Furthermore, in contrast with our study, Pourakbari and Biria (2015) found no significant difference between the performance of the multiple-choice group (input-oriented task) and the sentence writing group (output-oriented task) with the identical involvement load of three. One way to explain this difference can be that in their study task orientation was studied, whereas in our study, task focus was examined.

Furthermore, the results of this study revealed that generally the use of form-focused tasks would lead to better idioms recall. A number of studies endorse the effectiveness of form-focused tasks in comparison with meaning-focused ones (Ahour & Ghorbani Shemshadsara, 2015; Celik, 2019; Lan & Wu, 2013). For instance, the finding of this study confirms that of Elgort et al. (2016) in that, like this study, they used isolated form-focused tasks and utilized sentential contexts for form-focused task group. However, unlike this study in which input enhancement strategy was used, they used copying and writing novel words as form-focused strategy. On the other hand, the results of this study about the superiority of form-focused tasks over meaning-focused ones are incongruent with that of Khonarmi and Roostaei (2014), who found no meaningful difference between the performance of participants regarding meaning and form-focused tasks in recalling collocations. One possible reason for such difference may be the small sample size of their study including 41 learners.

The better performance of form-focused tasks over meaning-focused ones in our study may be accounted for by the tenets of FFI, which emphasizes using different strategies to draw learners' attention to language forms. Another reason to justify this finding may be using input enhancement

as a kind of form-focused strategy. Several studies provide evidence backing up this finding (Fakhrzadeh & Yazdanjoo, 2020; Naseri & Khodabandeh, 2019). This may be further theoretically justified on the basis of isolated form-focused tasks which underlie the idea that acquiring problematic words and forms would occur in a context separate from communicative one in which the focus of learners is mainly on meaning (Spada & Lightbown, 2008). Additionally, this finding may be explained based on cognitive overload. With regard to the limited capacity assumption, this finding can be associated with cognitive overload that learners may face while considering the whole meaning of a text in comparison with sentential context in which less cognitive overload is expected on the part of learners.

6. Conclusion and Implications

Based on the findings of this study, it can be concluded that because the use of appropriately loaded tasks results in better idioms learning, they should be more included in classroom contexts, particularly those which induce higher levels of involvement. Additionally, according to ILH, the combination of different learning tasks results in the increase of the involvement index of tasks which, subsequently, involve learners in the learning process. Therefore, a fair conclusion to be drawn from this finding is that if appropriate meaning-focused and form-focused tasks are employed, idioms achievement would increase.

Regarding the superiority of form-focused tasks over meaning-focused ones in idioms recall, it may be concluded that Schmidt's (1990) Noticing Hypothesis holds true in idioms learning. From these findings, it can also be concluded that instructional programs need to be designed in such a way that idioms are presented through appropriate types of instruction rather than those in which tasks are haphazardly combined and used. Another conclusion to be drawn from this finding is that although in EFL contexts teachers employ different methods to draw students' attention to forms; they should focus much more on form-focused tasks, particularly isolated form-focused ones in addition to input enhancement strategies.

At higher levels of involvement, form-focused tasks showed better performance than meaning-focused tasks, however, the difference was not significant, and they both were almost equally effective. It is worth mentioning that both form and meaning-focused tasks with higher degrees of involvement were output-oriented, and this issue might have affected the performance of the participants. Thus, in line with Swain's (2000) Output Hypothesis, this finding may lead one to the conclusion that output-oriented tasks inducing higher loads of involvement affect recalling idioms in a similar way regardless of the kind of task. Furthermore, since these tasks induced strong evaluation, it may be concluded that this component (evaluation) has a significant effect on idioms recall at higher involvement indices. Thus, a balanced integration of form and meaning-focused, output-oriented tasks with higher involvement loads might be considered as a suitable condition for strengthening idioms recall.

These findings can have implications for multiple stakeholders. Material developers and teachers can benefit from the findings of this study and design tasks with other degrees of involvement load to improve idioms recall. Besides, learners can benefit from isolated form-focused tasks to focus more on target forms. Furthermore, researchers can conduct new studies in order to compare the effects of other task types and involvement load on the learning of idioms.

Meanwhile, it needs to be acknowledged that this study was carried out under certain limitations and delimitations. For one thing, the present researchers knew that a longer treatment period could help them see the differences in the effectiveness of the treatments more accurately. However, it was not possible for them to extend the experimental period. In addition, for manageability reasons, only task focus was considered, and task orientation (input versus output-oriented) was not regarded as a variable. More importantly, the aggregate involvement load of a task was considered, not the comparative effect of the index of a task on each of its components; namely need, search and evaluation. These points imply that a degree of caution is needed in generalizing the findings. At the same time, interested researchers can consider these points in their future studies.

References

- Ahmadi Fatalaki, J. (2014). Involvement Load Hypothesis: Word meaning retention across oral and written task types. *International Letters of Social and Humanistic Sciences*, 37(1), 29-45.
- Alcaraz-Mármol, G., & Almela, A. (2013). The involvement load hypothesis: Its effect on vocabulary learning in primary education. *Revista Espanola de Linguística Aplicada (RESLA)*, 26(1), 11-24.
- Ahour, T., & Ghorbani Shenshadsara, Z. (2015). The impact of focused and unfocused tasks on university students' grammatical achievement. *Theory and Practice in Language studies*, 5(1), 124-133.
- Amini, A., & Maftoon, P. (2017). The impact of skill integration on task involvement load. *The Journal of English Language Pedagogy and Practice*, 10(21), 29-48.
- Alavinia, P. & Rahimi, H. (2019). Task types effects and task involvement load on vocabulary learning of EFL learners. *International Journal of Instruction*, 12(1), 1501-1516.
- Bao, G. (2015). Task type effects on English as a foreign language learners' acquisition of receptive and productive vocabulary knowledge. *System*, 53, 84-95.
- Cohen, J.W. (1988). *Statistical power analysis for the behavioral science* (2nd ed.). Lawrence Erlbaum Associates.
- Celik, B. (2019). A comparison of form-focused and meaning-focused instruction types: A study on Ishik university students in Erbil, Iraq. *International Journal of English Linguistics*, 9(1), 201-228.
- Ellis, R. (2003). Designing a task-based syllabus. *RELC Journal*, 34(1), 64-81.
- Ellis, R. (2015). *Understanding second language acquisition*. Oxford University Press.
- Elgort, I., Candry, S., Boutorwick, Th. J., Eyckmans, J., & Brysbaert, M. (2016). Contextual word learning with form-focused and meaning-focused elaboration. *Applied Linguistics*, 25, 1-23. doi:10.1093/applin/amw029
- Fakhrzadeh, S., & Yazdanjoo, S. (2020). The impact of using input enhancement techniques in the use of frequent collocations via reading on restatement in writing of Iranian intermediate EFL learners. *MEXTESOL Journal*, 44(3), 1-22.
- Golshan, M. (2015). Task-based language teaching and focus on form. *Journal of Advances in English Language Teaching*, 3(4), 75-81.
- Haward, J. (1998). *Idioms in American life*. Prentice-Hal, Inc.
- Hassanzadeh, M. (2016). Experimenting the influence of input modality on involvement load hypothesis. *International Journal of Social, Behavioural, Educational, Economic, Business and Industrial Engineering*, 10(6), 2043-2048.
- Hinkel, E. (2017). Teaching idiomatic expressions and phrases: Insights and techniques. *Iranian Journal of Language Teaching Research*, 5(3), 45-59.
- Hulstijn, J. H., & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning*, 51(3), pp.539-558. <https://doi.org/10.1111/0023-8333.00164>
- Irujo, S. (1986). Don't put your leg in your mouth: Transfer in the acquisition of idioms in a second language. *TESOL Quarterly*, 20(2), 287-304.

- Jahangard, A. (2013). Task-induced involvement in L2 vocabulary learning: A case for listening comprehension. *Journal of English Language Teaching and Learning*, 12, 43-62.
- Jing, L., & Jianbin, H. (2009). An empirical study of the involvement load hypothesis in incidental vocabulary acquisition in EFL listening. *Polyglossia*, 16, 1-11.
- Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. Longman.
- Kaivanpanah, S., Alavi, M., & Ravandpour, A. (2020). The effect of input-based and output-based tasks with different and identical involvement loads on Iranian EFL learners' incidental vocabulary learning. *Cogent Psychology*, 7, 1-16. <https://doi.org/10.1080/23311908.2020.1731223>
- 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.
- Karalik, T., & MerÇ., A. (2016). The effect of task-induced involvement load on incidental vocabulary acquisition. *Journal of Graduate School of Social Sciences*, 13(35), 77-92.
- Khonamri, F., & Roostae, S. (2014). Does extensive reading combined with form-focused or meaning-focused activities affect lexical collocational knowledge of Iranian learners? *Theory and Practice in Language Studies*, 4(5), 1038-1044.
- Kıvrak, C., & Uygun Gökmen, D. (2019). The effects of task-induced involvement load and input modality on incidental vocabulary learning. *Journal of Language and Linguistic Studies*, 15(4), 1355-1375.
- Liontas, J. I. (2003). Killing two birds with one stone: Understanding Spanish VP idioms in and out of context. *Hispania*, 86(2), 289-301.
- Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, 22(1), 1-26.
- Lan, Y., & Wu, M. (2013). Application of form-focused instruction in English pronunciation: examples from Mandarin learners. *Creative Education*, 4(9), 29-34.
- Moreno, E. M. G. (2011). The role of etymology in the teaching of idioms related to colors in an L2. *Porta Linguarum*, 16, 19-32.
- McCarthy, M., & O'Dell, F. (2017). *English idioms in use*. Cambridge University Press.
- Naseri, E., & khodabandeh, F. (2019). Comparing the impact of audio-visual input enhancement on collocation learning in traditional and mobile learning contexts. *Applied Research on English Language*, 8(3), 383-422.
- Naserpour, A., Zarei, A. A., & Esfandiari. R. (2020). The effects of task orientation and involvement load on learning collocations. *Journal of Teaching Language Skills (JTLS)*, 39(1), 71-114. doi: 10.22099/jtls.2020.37032.2809
- Prabhu, N. S. (1987). *Second Language Pedagogy*. Oxford University Press.
- Pourakbari, A. A., & Biria, R. (2015). Efficacy of task-induced involvement in incidental lexical development of Iranian senior EFL students. *English Language Teaching*, 8(5), 122-131.
- Richards, J. C., & Renandya, W. A. (2002). *Methodology in language teaching: An anthology of current practice*. Cambridge University press.
- Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11, 129-158.
- Swain, M. (2000). The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 97-114). Oxford University Press.

- Szudaeski, P. (2012). Effects of meaning- and form-focused instruction on the acquisition of verb-noun collocations in L2 English. *Journal of Second Language Teaching and Research*, 1(2), 3–37.
- Spada, N., & Lightbown, P. M. (2008). Form-focused instruction: Isolated or integrated? *TESOL Quarterly*, 42(2), 181-207. <https://doi.org/10.1002/j.1545-7249.2008.tb00115.x>
- Sarani, A., Mousapour Negari, G., & Ghaviniat, M. (2013). The role of task type in L2 vocabulary acquisition: a case of Involvement Load Hypothesis. *Acta Scientiarum. Language and Culture, Maringá*, 35(4), 377-386.
- Soleimani, H., & Rahmanian, M. (2015). Visiting involvement load hypothesis and vocabulary acquisition in similar task types. *Theory and Practice in Language Studies*, 5(9), 1883–1889. doi:10.17507/tpls.0509.16
- Thawabteh, M. (2011). Linguistic, cultural, and technical problems in English-Arabic subtitling. *Journal of Translation and Interpretation*, 5(1), 24-44.
- Teng, M. F. (2015). Involvement load in translation tasks and EFL vocabulary learning. *The New English Teacher*, 9(1), 83-101.
- Tajeddin, Z., & Daraee., D. (2013). Vocabulary acquisition through written input: Effects of form-focused, message-oriented, and comprehension tasks. *Teaching English as a Second or Foreign Language*, 16(4), 1–19.
- Un-udom, N. (2018). The effects of task-induced involvement on EFL vocabulary. *Chophayom Journal*, 29(3), 181-196.
- Wray, A., & Perkins, M. R. (2000). The functions of formulaic language: An integrated model. *Language and Communication*, 20(1), 1-28.
- Zou, D. (2016). Vocabulary acquisition through cloze exercises, sentence-writing and composition-writing: Extending the evaluation component of the involvement load hypothesis. *Language Teaching Research*, 21, 54-75.
- Zarei, A. A., & Moftakhari Rezaei, G. (2016). The effect of task type and task orientation on L2 vocabulary learning. *Issues in Language Teaching*, 5(2), 255-278.