Iranian Journal of Applied Linguistics (*IJAL*)

**Vol. 23, No. 2, September 2020, 1-32**

# On the Effectiveness of Integrated Skills Approach in Language Teaching: A Meta-analysis

**Fatemeh Ahmadnattaj, Seyyed Ali Ostovar-Namaghi***

*English Language Department, Shahrood University of Technology, Shahrood, Iran*

## Abstract

This meta-analysis was conducted to synthesize the effect of 22 primary studies which have been conducted to test the effect of the integrated skills approach (ISA) on language skills and components. Three questions guide this analysis: What is the overall effect of ISA on language skills and sub-skills? To what extent moderator variables such as learners' level of education and proficiency modify the effect of the ISA? What is the magnitude of publication bias in this analysis? The overall effect size was found to be 1.18, which represents a large effect size based on Cohen, Manion and Morrison' (2007) scale. The results of moderator analysis revealed that the ISA has the largest effect for advanced learners and at a tertiary level of education. The symmetrical funnel plot together with fail-safe N test shows that publication bias does not have any significant effect on the effect size reported in this study. The findings of this meta-analysis have clear implications for practitioners, policymakers, and curriculum developers.

**Keywords:** Integrated skills approach, English proficiency, Research synthesis, Meta-analysis, Effect size.

**Article Information:**

**Received:** 10 April 2020 **Revised:** 3 August 2020 **Accepted:** 18 August 2020

**Corresponding author:** Associate professor of TEFL, English Language Department, Shahrood University of Technology, Iran. **Email:** ostovarnamaghi@shahroodut.ac.ir

## 1. Introduction

Up to 1970s, the discrete skills approach, along with its prime exemplar the Audio-Lingual method, was the dominant mode of language instruction. This approach was based on the premise that breaking the language into skills and components and focusing on one skill or component at a time catalyzes the process of language learning (Jing, 2006). Moreover, teachers who followed this approach believed that presenting one skill at a time leads to the accurate use of language by language learners (Klimova, 2014). This approach uses language as an end in itself and as such leaves no room for authentic communication (Dubin & Olshtain, 1986); hence, discrete skill approach is commonly referred "language-based approach" (Oxford, 2001).

It was in late 1970s that scholars in applied linguistics leveled severe criticisms against the discrete skills approach by arguing that skills are not divorced from each other (Corder, 1971, 1978; Kaplan, 1970; Stern, 1992). Widowson was the first linguist who attached the discrete skills approach by arguing "we can talk of skills in respect to usage, but if we talk about language use, we need a different concept, and perhaps a different term" (Widdowson, 1998, p. 325). More cogently, Oxford (2001) argued, If the strands are not woven together effectively and well, the instructional loom is likely to produce something small, weak, ragged, and pale--not recognizable as a tapestry at all. Other critics argued that: the discrete skills approach fails to prepare learners for social and academic language use (Widdowson, 1998); you may need to break a language into skills to learn it but you need to integrate skills to communicate (Hinkel, 2010); and that language skills are inseparable because the receptive skills and the productive skills are two sides of the same coin (Harmer, 2007).

These criticisms together with changes in the way language and learning are conceptualized led to the development of what is known as integrated skills approach (ISA). This approach gained popularity because it was a shift in language teaching away from boring mechanical practice towards authentic language use. Many scholars have advocated the use of ISA because they believe this approach:

- enables learners to have a more realistic access to authentic language learning, whereas a segregated approach does not offer a meaningful understanding of language or a motivating style to learning a foreign language (Myers & Hilliard, 1997).
- is enthusiastically accepted by students and most of them had a positive attitude toward this approach (Mitrofanova & Chemezov, 2011);
- helps learners carry over their skills and declarative knowledge from one skill to another which facilitates and simplifies the improvement of the other skills (Strang, 1972).

The popularity of the ISA not only developed its theoretical basis, but encouraged many researchers to test its efficacy under experimental conditions. However, despite a hefty body of research on the effect of ISA on the learners' proficiency (Durukan, 2011; Tajan et al., 2015; Alqouran & Samadi, 2016; Oxford, 1994; Mokhamar, 2016; Aldosari, 2016; Cheong et al., 2017), there is no consensus on the effect size of the ISA, which explains the inconclusiveness of the results. On the other hand, it is crucial for curriculum developers, EFL teachers, and policymakers to be assured of the effectiveness of this approach of teaching to make an informed decision based on this research. Therefore, a meta-analysis study is needed to synthesize the results of the previous related studies to provide a precise estimate of the population effect.

Research aims at helping policy makers make more informed decision. Despite the fact that ISA is more in line with theories and principles of second

language acquisition, the empirical findings reported are of little use in decision making because they are inconclusive and at times contradictory. To fill in this gap, a meta-analysis is needed to calculate the effect size related to each study, statistically synthesize them, and based on the overall effect size determine the efficacy of ISA. Moreover, some factors may moderate the effectiveness of ISA. Therefore, this study was conducted not only to determine the overall effect size of ISA but also to contextualize the use of this educational intervention through moderator analysis and by doing so help policy makers in making more informed decisions. This study will contribute to the theory and practice of ISA by answering questions:

(1) What is the overall effect of the ISA on EFL learners' proficiency?

(2) To what extent moderator variables modify the effect of ISA?

(3) What is the magnitude of publication bias in this study?

## 2. Literature Review

Integrated skills approach (ISA) and various offshoots of communicative language teaching are usually associated with teaching the four main skills (Hinkel, 2010). Widdowson (1978) defined ISA as teaching language in a way that brings linguistic skills and communicative abilities into close association with each other. This approach is consistent with communicative language teaching in that both of them focus on meaningful and authentic language use so that the oral and written language improves hand in hand (Su, 2007). Bothe of these approaches react against the discrete skills approach which breaks the language into its component parts and then in line with behavioristic psychology involves students in mechanical practice of discrete elements until they show automatic mastery of the target form.

While in real communication interlocutors subconsciously integrate language skills, in language teaching is not that straightforward. To help practitioners systematically integrate language skills, Nunan (1989) outlines the principles that should guide the design of teaching materials and modules for

integrating a variety of language skills. Similarly, as McCarthy (2001, p. p.54) contends, materials that help students transfer from one skill to another paves the way for "a greater integration of the traditional four skills in language teaching, where writing tasks might be 'spoken' in their mode, and vice versa."

Although ISA is promising in theory, it faces many challenges in practice. To start with, focusing on a single skill or component at a time is more conducive to focused teaching and intensive learning (McDonough & Shaw 2003). Moreover, in many contexts where the discrete skills method is favored, both teachers and learners resist skills integration (Richards & Rodgers, 2001). Another challenge in implementing ISA is that due to the dominance of discrete skills approach, many learners have developed language skills unevenly (Hinkel, 2003) and as such are unable and unwilling to participate in activities that involve skill integration.

Many qualitative studies have been conducted to explore learners' perceptions and the effect of this approach on language proficiency (Sue, 2007; Tajzad & Ostovar-Namaghi, 2014). . For instance, Su (2007) explored learners' views, and satisfaction with integrated-skill was examined. It was revealed that the majority of students wanted to implement the ISA in class for the next academic year. They strongly believed that four language skills should be integrated as an integral part of a language course. Similarly, Tajzad & Ostovar-Namaghi (2014) investigated the EFL learners' perceptions of the integrated skills approach to language teaching revealing that the participants positively perceive this approach as a way of teaching a foreign language. ISA shifts teachers' attention away from coverage of the textbook towards involving students in communication since integration saves time. Thus, their communicative competence will improve.

In addition to qualitatively exploring learners' perceptions of IAS, many experimental studies have also been conducted to test the effectiveness of this approach under experimental (Noyce & Christie, 1983; Joseph, 1984; Mekheimer, 2011; Mekheimer, 2013; Spada et al., 2014; Alhasan, 2018) and some other scholars

investigated the impact of semi-integrated models of the ISA (Steven, 1987; Simmons et al., 1994; Aljiffri, 2010; Durukan, 2011; Aldosari, 2011; Alqadi & Alqadi, 2013; Cho et al., 2015; Tajan, Sadeghi & Rahmany, 2015; Tajan, 2016; Mokhamar, 2016; Alqouran & Smadi; Aldosari, 2016; Mubarok & Sofiana, 2017; Erlidawati & Syarfuni, 2018).

Cheong et al. (2017) found that reading and listening have a significant positive but limited effect on integrated writing, but reading cognitive skills contributed more to the performance of the integrated writing task than what the listening cognitive skills did.In the same vein, Tavil (2010), who tested the effect of integrated listening and speaking on students' oral communicative competence, found that teaching listening and speaking tasks in integration is more effective and beneficial than teaching these skills separately. In another study (Alqadi & Alqadi, 2013) tested the effect on extensive reading on paragraph writing and found that integrating reading and writing not only significantly improvers learners in expressing their thoughts and feelings in writing but also decreases their stress in the writing process. In contrast with the previous studies, Hoang and Ngoc (2021) found that speaking practice has little impact on students' listening performance.

On quite a different vein, some studies have tried to test the effectiveness of integrating reading and writing task on reading and writing skills (Steven 1987; Simmons et al., 1993; Aljiffri, 2010; Durukan, 2011; Aldosari, 2011; Cho et al., 2015; Mokhamar, 2016; Alqouran & Samadi, 2016; Aldosari, 2016; Mubarok & Sofiana, 2017; Erlidawati & Syarfuni, 2018). They have reported that:

- the cooperative integrated reading and composition technique has a significant effect on the reading and writing skills of primary school students ( (Durukan, 2011);
- integrating reading and writing skills has a significant on students' paragraph writing (Mokhamar, 2016);
- reading-writing integration has a positive effect not only on students' overall writing performances but also on the writing sub-skills and reading comprehension (Alqouran & Smadi, 2016);

- integrated reading has a positive influence on the quality of writing (Aldosari, 2016);
- integrating literacy skills with content-based instruction led to better achievement gains in literacy development and Social Studies achievement (Aljiffri's, 2010);
- integrated listening activities greatly improves learners' speaking accuracy and fluency within the framework of task-based activities (Alhasan, 2018; Tajan et al., 2015; Tajan, 2016).

Another group of studies have tested the effect of ISA on language sub-skills and components (Cho & Brutt-Griffler, 2015; File & Adams, 2010; Joseph, 1984; Spada et al., 2014). The results of Joseph's (1984) study indicated that learners' performance in vocabulary has improved as a result of the integrated skills instruction whereas later in another context this result rejected. Conversely, File and Adams (2010) found that the ISA does not have such positive effects on learners' vocabulary, and segregated instruction is more effective. Similarly, Spada et al. (2014) studied the effect of this approach on learners' grammar and no significant differences between the instructional groups were indicated. They concluded that integrated and isolated instruction are complementary to each other and both of them positively affect L2 learning. Concerning proficiency differences, Cho and Brutt-Griffler (2015) investigated the differences between beginnings, intermediate, and advanced learners in terms of their achievement in integrated skills classrooms. He noticed learners in intermediate and advanced levels had significant improvement in integrated reading and writing; however, beginning learners did not show improvement.

A critical look at the experimental studies that aimed at testing the effectiveness of ISA on skills and components reveals that the results of these studies are inconclusive and at times contradictory. While some studies found that ISA has a significant effect on language skills and components ( e.g., Aldosari,

2016; Alhasan, 2018; Alqadi & Alqadi, 2013; Durukan, 2011; Mokhamar, 2016; Tavil, 2010 ) there are some other studies reporting that this approach has a limited or no effect on skills and components (Adams, 2010; Cho & Brutt-Griffler, 2015; Hoang & Ngoc; 2021; Spada et al.,2014). Based on the original philosophy of research, a study is justified if it helps policy makers make more informed decisions. Although ISA it theoretically far superior to discrete skills approach, and although many empirical studies have tested its effectiveness under experimental conditions, replacing the discrete skills approach with the ISA may be an unwarranted change because the results of the empirical findings are inconclusive. What is needed is a meta-analysis of the previous empirical findings to establish the overall effect of ISA. This study aims to fill in this gap by not only synthesizing the results of the previous experimental studies but also by exploring how moderator variables modify the effectiveness of this educational intervention.

## 3. Method

### *3.1. Sampling Procedure and Materials*

Since the participants of a meta-analysis are all participants of the previous studies in that specific subject (Green & Hall, 1984), the sampling procedure in meta-analysis involves selecting the relevant studies with the same problem which present the statistical data needed for meta-analysis and articles published in scientific journals rather selecting of the participants. But, prior to the selection of studies for meta-analysis, we defined a set of inclusion/exclusion criteria. Studies should meet the following criteria to be included in the sample:

1) Should be true or quasi-experimental in which L2 learners were provided with instruction on either true integrated or semi-integrated skills approach; 2) Should be published in English; 3) Should contain sufficient statistic data to compute the effect size such as means and standard deviations for pre- and post-test scores for

intervention and comparison group means, the sample size of each group, t-test values, and group sizes.

To search studies that meet the inclusion criteria, an in-depth search was done in Google Scholar and the Cross-Ref search engine and in four online databases (ending in June 2018) including ProQuest, ERIC, Science Direct, Elsevier. We obtained different experimental studies including articles published in refereed or non-freed journals, and MA theses to prevent publication bias which may not get published due to the lack or weak correlation with published research. We used a list of different combinations of keywords including the ISA, Integrating skills instruction, four skills instruction, Language skills integration, the effect of integrated skills, and the impact of integrating skills. We retrieved a comprehensive list of abstracts, so studies appearing to meet inclusion criteria would then be obtained and reviewed in full, and those not meeting the inclusion criteria would be excluded. If the item looked promising according to its abstract or title, we read the whole text to check other criteria such as the necessary statistical data. Once a document was obtained, the reference list was investigated to identify the other published studies. Of 29 documents collected, 22 studies on the ISA met inclusion criteria. Descriptive statistics of those 22 studies are presented in Table 1.

Table 1. *Characteristics of the included studies in the Meta-Analysis*

| **Characteristic** | | | | | | | | **Total** |
|---|---|---|---|---|---|---|---|---|
| **Publication year** | | 1983 | 1984 | 1987 | 1993 | 2010 | 2011 | N=22 |
| | N | 1 | 1 | 1 | 1 | 3 | 3 | |
| | % | 4.5 | 4.5 | 4.5 | 4.5 | 13.6 | 13.6 | |
| | | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 |
| | N | 1 | 1 | 2 | 4 | 1 | 2 | 1 |
| | % | 4.5 | 4.5 | 9 | 18.18 | 4.5 | 9 | 4.5 |
| **Type of research** | | MA thesis | | | Article | | | N=22 |
| | N | 2 | | | 20 | | | |
| | % | 9 | | | 90.9 | | | |
| **Country** | | USA | New Zealand | | Turkey | Iran | Canada | N=22 |
| | N | 5 | 1 | | 2 | 2 | 1 | |
| | % | 22.7 | 4.5 | | 9 | 9 | 4.5 | |
| | | Germany | Jordan | | Palestine | Saudi Arabia | Indonesia | |
| | N | 1 | 1 | | 1 | 5 | 2 | |
| | % | 4.5 | 4.5 | | 4.5 | 22.7 | 9 | |

## *3.2. Data collection*

The findings, the methodological and substantive characteristics of each study were evaluated and coded. Moreover, the moderator variables were classified since these variables can cause heterogeneity by affecting the strength of effect between studies. One risk that threatens Meta-analysis reliability is that the coder enters his personal judgment or bias in coding and classifying data. Thus, it is important to use a coding scheme and data extraction by multiple raters to establish inter-rate reliability (Bullock & Svyantek, 1985). Based on what was suggested by Plonsky and Oswald (2012), the coding reliability was assessed through measuring the inter-rater reliability. To do so, two coders coded separately all samples. The coders were MA students of TEFL in the Department of Applied Linguistics at Shahrood University of Technology. They were trained in one session to become familiar with the process of coding. Then, Cohen's kappa reliability coefficient between the coders was calculated by using SPSS

software. It was determined as 0.89 which is interpreted as the agreement between two independent coders is higher than 89% and it means that a perfect consistency exists between the raters. Finally, disagreements were discussed and resolved and coding sheets were checked and corrected based on the common agreement.

The statistical data were imported into a Microsoft Excel file. The specific driven data from retrieved studies were categorized as Name of the study, Name of author, Publication year, Country of study, Contact hour, Gender, Age, Publication type, Design of studies, Data collection instrument, L2 proficiency, Educational level, Dependent variable, Sample size, Control group sample size, Experimental group sample size, Mean of pretest and posttest in each group (control and experimental group), Standard deviation of pretest and posttest in each group, P-value, T-value, Pretest, and posttest correlation.

### *3.3. Data analysis*

The contributed effect size of each study was calculated. Among different measures for calculating the effect size, Hedges' g produces a more precise estimate and it is preferable to Cohen's d when the sample size is below 20 because it uses pooled standard deviations from both groups (Control and Experimental group) and it will give us a better estimate of the population standard deviation, therefore we used Hedges' g (Borenstein et al. 2009). In cases where a study examined the effect of the ISA on achievement on different groups or skills and reported the related data, more than one effect size was calculated for that study. Finally, the average effect size was calculated in order to combine the effect sizes of the studies to decide on the effectiveness of the ISAs and its effect direction. Having calculated the effect sized related to each study and synthesized them into an overall effect size, we then had to interpret its significance. Different scales have been proposed for the interpretation of the overall effect size. For instance, Cohen, Manion and Morrison (2007) interpret the effect size as follows:

- $0 \leq$ Effect size value $\leq 0.20$ insignificant;

- 0.21 ≤ Effect size value ≤ 0.50 small;
- 0.51 ≤ Effect size value ≤ 0.8 medium; and
- 0.81 ≤ Effect size values, large effect size.

There are various programs and software available to compute effect sizes and to do meta-analyses such as the Comprehensive Meta-Analysis (CMA), EPPI-Reviewer, and Review Manager (RevMan) program. In this study, CMA software was used to calculate effect sizes because it is the most appropriate one to do a meta-analysis. It provides the opportunity to calculate forest plot, funnel plot of both the observed study and imputed study, individual effect size, and main effect size.

### *3.3.1. Test of heterogeneity*

Prior to calculating the effect sizes of studies, we should select the appropriate analysis model. The degree of heterogeneity determines the analysis model and when statistical heterogeneity is recognized, it is normally recommended to use a random-effects model. A Q test and $I^2$ test were used to determine the degree of heterogeneity. If Q value and degree of freedom are equal, there is no heterogeneity (dispersion in observed effect).

### *3.3.2. Sub-groups analysis*

The potential effect of the different moderator variables of interest was measured to find statistical differences between sub-groups and between average effect sizes of the variables. It also allows for some comparison between groups (Little, Corcoran & Pillai, 2008). Publication type, Language skills and sub-skills, Educational level, and L2 Proficiency level were analyzed separately as moderator variables in this study.

### *3.3.3. Publication bias evaluation*

One potential threat of systematic reviews that researchers should consider is publication bias. Publication bias is originally defined as the direction and statistical significance of the results that determine publication or non-publication of studies (Rothstein et al., 2005). Consequently, a part of the literature on the subject will be lost and the literature will be unrepresentative of the population of completed studies. Since meta-analysis is based on the previous empirical studies and it aims to synthesize the result of all previous studies, losing some parts of the literature causes inflation in the estimation of effects. Funnel Scatter plot, Trim and Fill, and Classic Fail-Safe N were used to evaluate publication bias in this study.

## 4. Results

The results of the funnel plot (Figure1) and the Trim and Fill test presented in Table 2 show little evidence of potential publication bias on the ISA effect. As presented in Figure 1, the funnel plot is almost asymmetric and there is substantial variability in effect sizes, with a greater spread in the middle of the figure. Moreover, these effects are not distributed equally on both sides of the mean effect. In other words, since the funnel plot is symmetrical, there is no indication of publication bias. However, as shown in Table 3, fail-safe N test was also used to do away with the subjective interpretation of the funnel plot and present objective numerical information about publication bias. Moreover, as it is presented in the graph of Duval and Tweedie (figure 2), by adding 14 studies at the left side of the graph, it will be symmetric structure and unbiased. Trim and Fill shifts the most extreme small studies from the positive side of the funnel plot to make it a symmetric plot and give an unbiased estimate of the effect size (Borenstein et al., 2009). In other words, when the difference between observed and adjusted values is small, the result is considered to be valid.

The results show that the effect size value of the studies decreases from 1.18 to 0.81 regarding the random effect model of analysis. Based on Cohen's

(1987) effect size interpretation, both values show a large effect and there is not a big difference between observed effect size in a publication-biased case and adjusted effect size in a totally unbiased case.

**Figure 1.** *Funnel plot of observed studies*

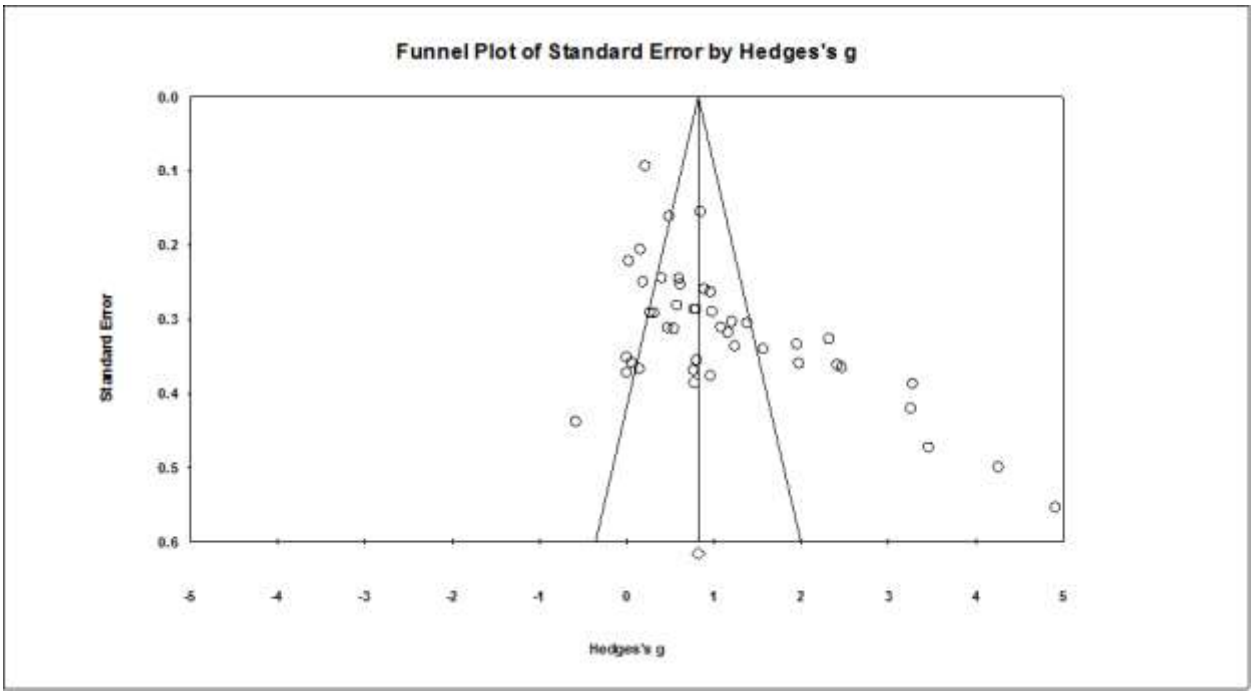
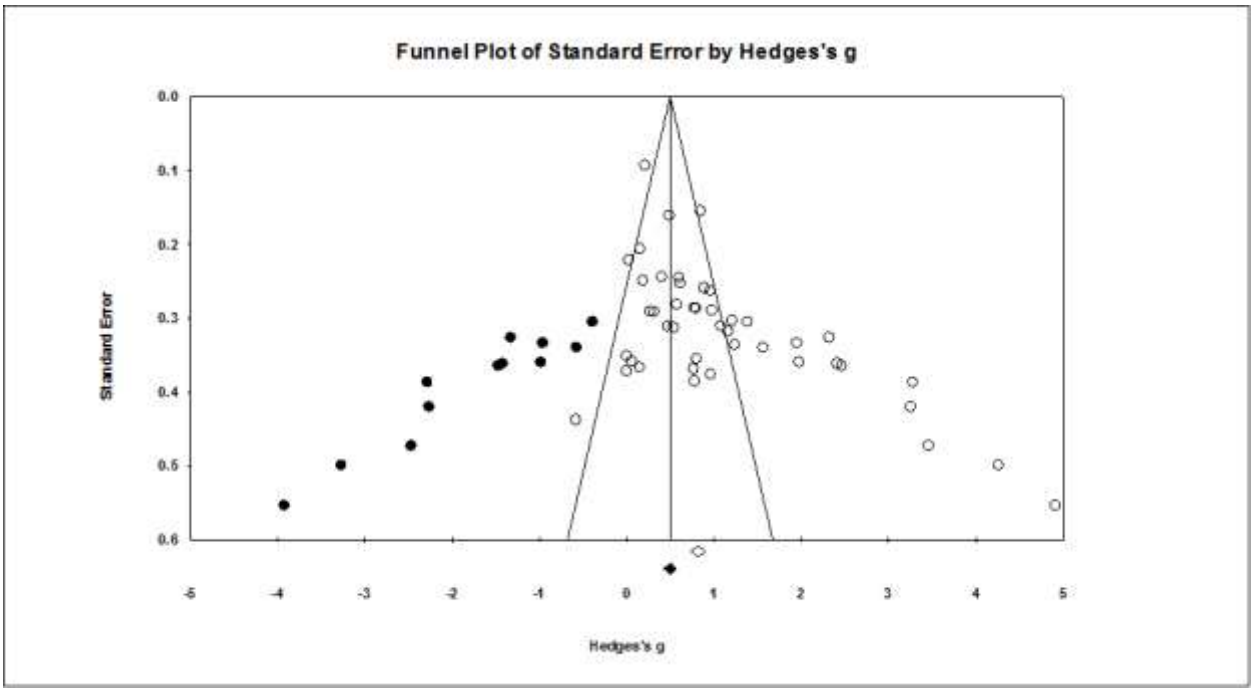



Table 2. *The result of Trim and Fill analysis*

| | | **Fixed Effects** | | | **Random Effects** | | | **Q Value** |
|---|---|---|---|---|---|---|---|---|
| | Studies Trimmed | Point Estimate | Lower limit | Upper limit | Point Estimate | Lower limit | Upper limit | |
| Observed values | 14 | 0.81 | 0.73 | 0.89 | 1.18 | 0.91 | 1.44 | 484.16 |
| Adjusted values | 14 | 0.49 | 0.43 | 0.56 | 0.81 | 0.61 | 1.23 | 1029.05 |

**Figure 2.** *Funnel plot on observed and imputed studies*




Since the results of the funnel plot are subjective, Fail-Safe N (Cooper, 1979) was also used to evaluate and adjust for the number of undiscovered or missing studies (insignificant effect sizes) that are required to nullify the observed effect size. The more the number of required studies, the less probability of publication bias exists. The fail-safe N test determined that a total of 6839 studies with null results would be needed in order to nullify the effect size (Table 3). This number goes beyond the criterion number (i.e., $5k + 10 = 240$ where $k = 46$ studies) (Rosenthal, 1991). The result of the test suggests that publication bias could not explain the significant positive effects observed across all studies. Therefore, the effect of unpublished studies which were not covered in this study does not have any effect on the results of this study.

Table 3. *Results of the classic fail-safe N*

| **Classic fail-safe N** | |
|---|---|
| Z value for observed studies | 23.97 |
| P value for observed studies | 0.00 |
| Alpha | 0.05 |
| Tails | 2.00 |
| Z for alpha | 1.95 |
| Number of observed studies | 46 |
| Number of missing studies that would bring p value to> *alpha* | 6839 |

In order to verify the effect of the ISA on students' proficiency, the overall effect size was measured. The Hedges' g was calculated for 46 independent samples from 22 studies (Figure 3). In line with the random model, the mean effect size, the result of the homogeneity test, P-value, and some other statistics are shown in Table 4. The estimated overall effect sizes for the random model is 1.18 of 95% confidence interval and is considered as large or incredibly effective according to Cohen et al. (1987) interpretation.

Table 4. *Fixed and random effect model statistic*

| **Model** | | **Effect size and 95% confidence interval** | | | | | **Test of null (2-tail)** | | **Heterogeneity** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Number studies | Point estimate | Standard error | Variance | Lower Limit | Upper Limit | Z value | P value | Q value | Df | P value | I squared |
| Fixed | 46 | 0.81 | 0.04 | 0.00 | 0.73 | 0.89 | 20.53 | 0.00 | 484.16 | 45 | 0.00 | 90.70 |
| Random | 46 | 1.18 | 0.13 | 0.01 | 0.91 | 1.44 | 8.72 | 0.00 | | | | |

Before we look at the mean effect size, we should know the magnitude of P-value. The corresponding P-value is p=0.00<0.001. Therefore, the test is

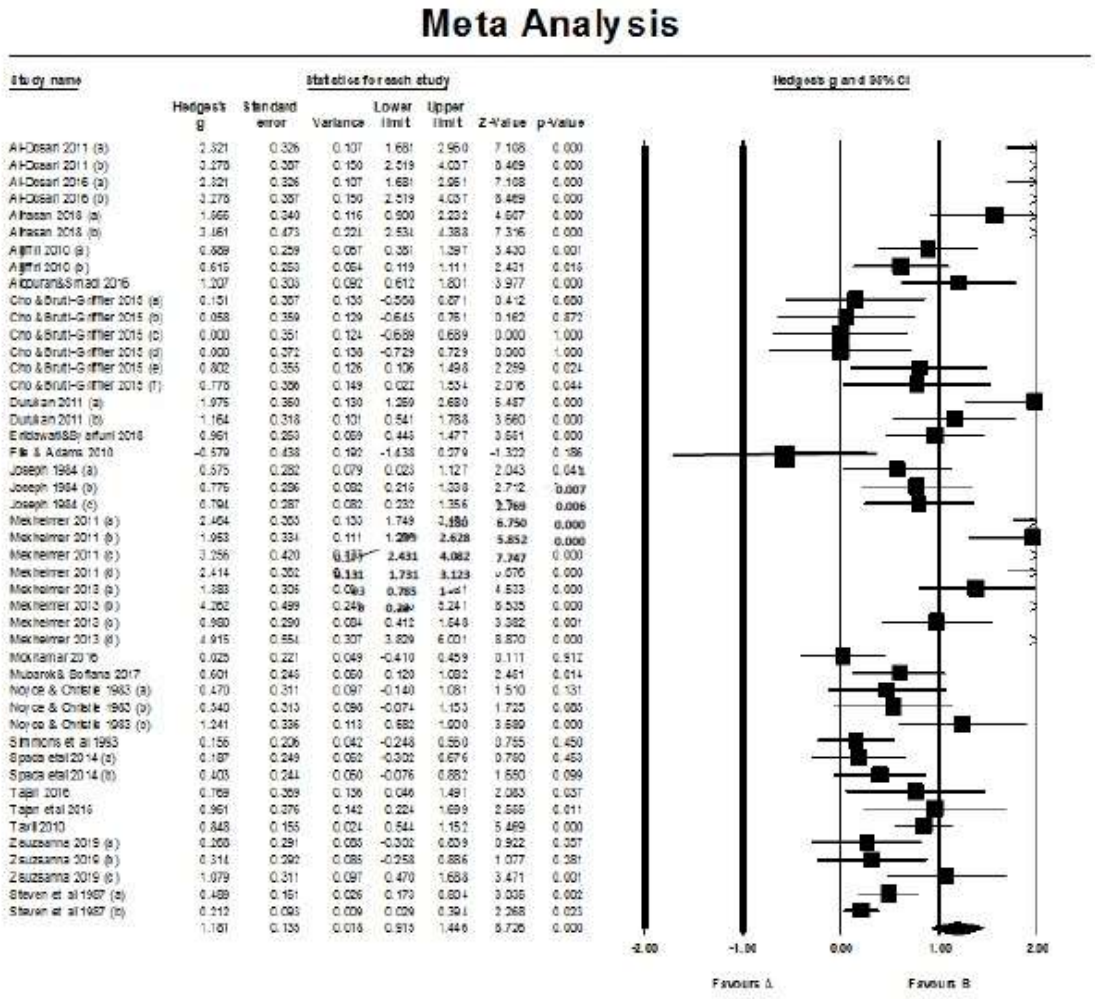
statistically significant and we reject the null hypothesis. Furthermore, the values of heterogeneity are presented in the right part of the table (Q, df, I squared). The significant Q statistic result (Q (46) =848.16, $p< 0.05$), shows heterogeneity of effect sizes. Moreover, that the $I^2$ value is bigger than 75% shows that the distribution of the effect values of studies on the ISA is highly heterogeneous. Therefore, the random effect model was used for subsequent analysis.

The forest plot which shows the distribution of the effect sizes of 46 studies together with the overall effect size in figure 3. In this figure, the box, depicting the point estimate of each individual study, is indicative of the weight of the corresponding study in generating the meta-analysis. The whiskers show the confidence interval the length of which indicating the precision of the study. The shorter the line, the more precise the study is. Finally, the vertical line in the middle of the graph shows the line of no effect and the diamond in the bottom row shows the overall effect of the ISA.

As shown in Figure 3, nearly all confidence intervals are entirely on the right side of the line. This clearly indicates that the ISA has significant positive effect on language skills and language components. Four confidence intervals overlap the line of no effect. These studies found that ISA does not have any significant effect on language skills and components. What is noticeable is that only one confidence interval is on the left side of the line of no effect. In other words, only one study found that ISA has significant negative effect. The diamond in the bottom row synthesizes the individual effect sized shown in the forest plot into the combined overall effect of ISA. The bottom row of the forest plot shows a diamond, which summarizes and combines the effect of individual studies into the weighted average effect or the combined effect and turns the plot into a meta-analysis. In short, the middle point of the diamond shown in the bottom line visualizes the overall effect of the ISA. The middle of the diamond in the bottom row shows the overall or the combined effect of the meta-analysis.

Since it is noticeably far to the right from the line of no effect, it clearly shows that the ISA has a significant positive effect on language skills and components.

**Figure 3.** *Forest plot of the effect sizes (k= 46)*




## *4.1. Moderator analysis*

The distribution of effect sizes was heterogeneous as indicated by Q=484.16 and $I^2$= 90.70. Consequently, the analysis of moderator variables as a source of variation was demanded to explain the observed variance in the effect size. As an attempt to explain this variance, we assessed the effect of key study variables on

effect sizes by conducting moderator analyses within six categories of study characteristics. The distribution of moderator variables and their corresponding effect size (g) of ISA on learners' proficiency are ***demonstrated in Table 5.***

Table 5. *Moderator analysis on the effectiveness of ISA*

| **Moderator variables** | **K** | **Variance** | **Z value** | **P value** | **Effect size** | **95% confidence interval** | | **Standard error** |
|---|---|---|---|---|---|---|---|---|
| | | | | | | **Lower** | **Upper** | |
| **Educational level** | | | | | | | | |
| Primary | 7 | 0.09 | 2.06 | 0.03 | 0.41 | 0.03 | 1.23 | 0.30 |
| Secondary | 14 | 0.05 | 3.44 | 0.00 | 0.78 | 0.33 | 1.22 | 0.22 |
| University | 18 | 0.04 | 9.97 | 0.00 | 1.99 | 1.60 | 2.38 | 0.20 |
| **Skills and components** | | | | | | | | |
| Listening | 3 | 0.34 | 4.15 | 0.00 | 2.43 | 1.28 | 3.57 | 0.58 |
| Reading | 14 | 0.07 | 4.52 | 0.00 | 1.19 | 0.67 | 1.71 | 0.26 |
| Writing | 15 | 0.06 | 4.94 | 0.00 | 1.26 | 0.76 | 1.76 | 0.25 |
| Speaking | 6 | 0.16 | 3.85 | 0.00 | 1.58 | 0.77 | 2.38 | 0.41 |
| Grammar | 3 | 0.31 | 0.71 | 0.47 | 0.40 | -0.70 | 1.50 | 0.56 |
| Vocabulary | 2 | 0.50 | 0.03 | 0.97 | 0.02 | -1.36 | 1.42 | 0.71 |
| General | 1 | 0.95 | 0.82 | 0.41 | 0.80 | -1.11 | 2.72 | 0.97 |
| Linguistic complexity | 1 | 0.96 | 0.27 | 0.78 | 0.27 | -1.64 | 2.19 | 0.98 |
| communicative competence | 1 | 0.89 | 0.89 | 0.36 | 0.85 | -1.00 | 2.70 | 0.94 |
| **Publication type** | | | | | | | | |
| Article | 43 | 0.02 | 8.16 | 0.00 | 1.17 | 0.89 | 1.45 | 0.14 |
| Thesis | 3 | 0.30 | 2.90 | 0.00 | 1.59 | 0.52 | 2.67 | 0.55 |
| **L2 Proficiency level** | | | | | | | | |
| Beginner | 2 | 0.44 | 0.11 | 0.90 | 0.07 | -1.23 | 1.38 | 0.66 |
| Intermediate | 22 | 0.04 | 6.40 | 0.00 | 1.27 | 0.88 | 1.66 | 0.19 |
| Advanced | 6 | 0.14 | 3.77 | 0.00 | 1.43 | 0.69 | 2.18 | 0.38 |

### *4.1.1. Educational Level*

The effect sizes of ISA for primary, secondary, and university levels were 0.41, 0.71, and 1.59 respectively. The results indicated that learners under integrated skills instruction made significantly the highest gains at the university level (g=1.09, SE=0.07, $CI_{95}$= 1.44, 1.74) and learners at the primary level made the lowest gain (g=0.63, SE=0.06, $CI_{95}$ =0.27, 0.54).

### *4.1.2 Skills*

To investigate whether the effectiveness of ISA on learners' proficiency is different according to skills, the studies are classified under eight groups according to the skills and components under instruction as listening, speaking, reading, writing, grammar, vocabulary, general proficiency, and communicative competence. The average effect size of these studies was calculated 2.18, 1.24, 0.68, 0.90, 0.40, 0.25, 0.80, and 0.85 respectively. The results indicated that the highest achievement was in listening skill (g=2.18, SE=0.21, $CI_{95}$=1.76, 2.60) and the lowest achievement was in vocabulary (g=0.25, SE=0.24, $CI_{95}$= -0.22, 0.72).

### *4.1.3. Publication Type*

The studies were classified under two groups according to their type as articles and MA theses. As a result, the average effect size of practices in articles and MA theses were calculated as 1.17 and 1.59, respectively, and they represent a strong effect on Cohen's scale.

### *4.1.4. L2 Proficiency level*

Studies were classified into three groups including Beginner, Intermediate, and Advanced, and as the results of Table 3 indicate, the average effect size of these groups was calculated as 0.07, 1.27, and 1.43 respectively. The learners in the advanced group gained the highest achievement (g= 1.43) and students in the beginner group gained the lowest achievement (g=0.07).

## 5. Discussion

Although the ISA is more line with theories and principles of second language acquisition, the results of the studies that tested the effect of this educational intervention under controlled conditions were inconclusive. While some studies reported that the ISA has significant positive effect, others reported that it has a limited or negligible effect. Moreover, individual studies did not show how moderator variables such as learners' level of education, proficiency level, and the type of skill the intend to develop modify the effect of this educational intervention; hence, this meta-analysis was conducted not only to synthesize the effect of the previous empirical findings into an overall combined effect size but also present empirical evidence as to how moderato variables modify the effect of the ISA. To this end, 22 studies primary studies and 46 effect sizes were reviewed and synthesized in which 4871 participants participated in all studies. Analysis revealed an overall effect size of 1.18, which shows that ISA is an effective intervention in developing EFL learners' proficiency.

The calculated overall effect size of this meta-analysis was 1.18 (95% CI 0.91 to 1.44); this effect size is considered a large effect with respect to both Plonsky and Oswald's (2010) and Cohen's (1987) scale. However, according to Cohen (1987) and Heges (2008), effect sizes are most accurately interpreted in comparison with other effect sizes. To this end, it is worth comparing the strength of this finding with the findings of Plonsky (2011) which indicated an overall average effect of 0.49. Norris and Ortega (2000) also synthesized the effect sizes of 78 studies in the domain of second language instruction and the result indicated an overall effect size of 0.96. Moving closer to the focus of this study, it is good to point Nakhaei and Ostovar-Namaghi's (2017) meta-analysis of CLIL efficacy that yielded an overall effect size of 0.81. Taking the overall effect sizes reported in these studies as a basis for comparison, the overall effect size found in this study is strong.

Despite the strength and magnitude of the overall effect size, we should not take it as a basis for decision making since the studies covered in this meta-analysis indicated a high degree of heterogeneity ($I^2$= 90.7) ; hence; as suggested by Borenstein et al (2009), subgroup analysis is needed to see how moderator variables modify the effect on ISA. Taking level as a moderator, a significant difference was observed in elementary, secondary, and university settings. That is, ISA had a medium effect for both the primary level (0.63) and the secondary level (0.71), but it had a strong effect for the university level (1.59). These figures clearly show that an increase in educational level is indicative of an increase in the effectiveness of ISA on learners' proficiency Plonsky (2011) found a similar result in his meta-analysis and explained this result by noting that because college students are usually a more homogenous group than other levels, then they produce less group variance which leads to larger effect sizes. However, this difference is best explained by Alexander et al. (1998) by noting that children and younger learners are less cognitively developed than adult learners, and as they become more experienced and competent in a subject, their strategic behavior changes, so they need more pure instruction than adult learners. These explanations can apply to the current study as well since adult learners are cognitively developed, they can more benefit from the ISA. Therefore, the implementation of this approach at the university is far more effective than its implementation at lower levels.

Taking language proficiency as the moderator, subgroup analysis revealed that advanced level learners (g=1.43, p=0) and intermediate learners (g=1.27, p=0) made a great gain and beginners (g=0.07, p=0.90) didn't make a significant gain. This difference may partially be explained by noting that learners in higher levels of proficiency might already use the techniques of the ISA. On the other hand, the ceiling effect may influence the results.

Just like other moderators, skill type modified the effect of ISA. Among different skills, it was found that ISA had a largest effect on listening (2.43) and

speaking (1.58) and a slightly smaller effect on reading (1.19), and writing (1.26). Compared with the effect of CLIL on skills and components reported in Ostovar and Nakhaei's (2019) meta-analysis, ISA is found to be more effective in enhancing the mentioned skills in this study. They found large to medium effect for grammar (1.32), listening (0.91), general (0.76), writing (0.81), and reading (0.72). As shown in the results section, Funnel Scatter Plot, Trim and Fill Method, and Fail-Safe N test showed a low level of publication bias. In other words, since the results of these tests show few unpublished lost studies, they do not change the calculated overall effect size. Therefore, the effect size reported in this study shows a high level of validity.

The large effect size reported in this study provides robust evidential support for previous theoretical perspectives. To start with, it supports Oxford's (2001) argument that ISA is far more effective than the discrete skills approach. It also present empirical evidence showing that the ISA is not only enthusiastically accepted by students but also has a positive effect on their attitude (Mitrofanova & Chemezov, 2011). Moreover, it shows that this approach helps learners carry over their skills and declarative knowledge from one skill to another which facilitates and simplifies the improvement of the other skills (Strang, 1972). What is more notable is that the large effect size reported here shows that the ISA develops learners' communicative competence rather than just help them master discrete language elements. The large effect size of the ISA also verifies the fact that helping students integrate language skills through materials that help learners transfer one skill to another, as stated by McCarthy (2001), turns the boring mechanical practice of discrete elements into authentic language use inside the classroom.

Since the ISA has a solid theoretical basis and also since it is supported by robust empirical evidence, i.e., a large effect size which is a synthesis of the results of dozens of previous empirical studies, it is essential that:

- Materials developers weave skill integration into language education syllabi, especially for higher levels of language proficiency, since as shown in the moderator analysis section, the effect size of the ISA is far larger for advanced learners than for beginners;
- Language teachers design classroom activities that involve learners in integrating skills, break away from the single skill instruction inculcated by standardized tests such as the IELTS and heed the fact that a single skill can be taught in isolation from other language skills but it is of no use in communication because communication involves integrating skills.
- Language learners use receptive skills such as reading and writing to develop productive skill such as speaking and writing because naturally the receptive skills give the language user the information and linguistic resources needed in speaking and writing.
- Researchers who are interested in testing the effect of an educational intervention such as the ISA be more precise in reporting basic statistics such as standard deviations, sample sizes, means, effect sizes, reliability, and confidence interval; a large number of studies were excluded from this meta-analysis merely because they did not report the statistics needed to compute the overall effect size.

## 6. References

**References marked with an asterisk mark indicate studies included in the meta-analysis.**

*Abrams, Z. I. (2019). The effects of integrated writing on linguistic complexity in L2 writing and task complexity. *System, 81*(2019),110-121.

* Aldosari, H. (2016). Effects of an integrated reading/writing approach on improving writing skills of EFL students. *Dirasat, Educational Sciences, 43*(1), 761-771.

*Aldosari, H. (2011), October. Effects of integrating reading and writing during reading process on EFL English writing in a Saudi college course: A quasi-experimental study. *First International Forume on Academic Teaching: Social Inclusion and Information and Communication Technology*, 325-345. Retrieved from http://www. forumdocente.progard.

Alexander, P. A. (1998). A perspective on strategy research: Progress and prospects. *Educational Psychology Review, 10*(2), 129–154.

*Alhasan, A. A. (2018). *The role of integrated skill approach in improving students' achievement in aural-oral skills: Case study of secondary school students* (Doctoral dissertation). Retrieved from http://www.repo.uofg.edu.sd.

* Aljiffri, I. H. (2010). Effects of the integrated approach to teaching English and social studies on acheivement in a Saudi private elementary school. *Journal of Educational & Psychological Science, 11*(4), 11-31.

*Alqadi, K. R., & Alqadi, H. M. (2013). The effect of extensive reading on developing the grammatical accuracy of EFL freshman at Al-Bayat University. *Journal of Education and Practice, 4*(6), 106-113.

*Alqouran, A. &. Samadi, O. (2016). The effect of reading-writing integration on Jordanian undergraduate students' writing performance. *Arab World English Journal, 7*(2), 177-189.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis.* London, UK: Wiley.

Bullock, R. J., & Svyantek, D. J. (1985). Analyzing meta-analysi: Potential problems, an unsuccessful replication, and evaluation criteria. *Journal of Applied Psychology, 70*(1), 108-115.

Cheong, C. M, Zhu, X., & Liao, X. (2017). *Differences between the relationship of L1 learners' performance in integrated writing with both independent listening and independent reading cognitive skills.* New York, NY: Springer.

* Cho, H., & Brutt-Griffler, G. (2015). Integrated reading and writing: A case of Korean English language learners. *Reading in a Foreign Language, 27*(2), 242–261.

Cohen, J. (1987). *Statistical power analysis for the behavioral sciences (2nd ed.).* Hillsdale, NJ: Erlbaum.

Cole, W. M. (2014). Speaking to read: Meta-analysis of peer-mediated learning for English language learners. *Journal of Literacy Research, 46*(3), 358-382.

Cooper, H. M. (1979). Statistically combining independent studies: A meta-analysis of sex differences in conformity research. *Journal of Personality and Social Psychology, 37*(1), 131-146.

Corder, S. P. (1978). Language-learner language. In J. C. Richards (Ed.), *Understanding second and foreign language learning* (pp. 71-92). Rowley, MA: Newbury House.

Dubin, F., & Olshtain, E. (1986). *Course design: Developing programs and materials for language learners.* Cambridge: Cambridge University Press.

*Durukhan, H. (2011). Effects of cooperative integrated reading and composition (CIRC) technique onn reading-writing skills. *Academic Journals, 6*(1), 102-109.

*Erlidawati1, & Syarfuni. (2018). The effect of cooperative integrated reading and composition on reading comprehension of IAIN Lhokseumawe, Indonesia. *Advances in Language and Literary Studies, 9*(4), 153-160.

* File, K. A., & Adams, R. (2010). Should vocabulary instruction be integrated or isolated? *Tesol Quarterly, 44*(2), 222-249.

Chela-Flores, B. (2001).Pronunciation and language learning:An integrative approach. *IRAL, 39*(2), 85-101.

Green, B. F., & Hall, J. A. (1984). Quantitative methods for literature reviews. *Annual Review of psychology, 35*(1), 37-54.

Harmer, J. (2007). *The practice of English language teaching (4th Ed.).* England: Pearson Education Ltd.

Hedges, L. V. (2008). What are effect sizes and why do we need them? *Child Dvelpoment perspectivs, 2*(3), 167-171.

Hikel, E. (2010). Integrating the four skills: Current and historical perspectives. In R. B. Kaplan (Ed.), *Oxford hanbook in applied linguistics* (pp. 110-126). Oxford: Oxford university Press.

Hinkel, E. (2003) Simplicity without elegance: Features of sentences in L1 and L2 academic texts. TESOL Quarterly, 37, 275–301.

Klimova, B. F. (2014). Detecting the development of language skills in current English language

teaching in the Czech Republic. *Procedia- Social and Behavioral Sciences,* 158, 85-92.

Littell, J. A., Corcoran, J., & Pillai, V. (2008). *Systematic reviews and meta-analysis.* New York, NY: Oxford University Press.

McCarthy, M. J. (2001). Issues in applied linguistics. Cambridge: Cambridge University Press.

McDonough, J., & Shaw, C. (2003) Materials and methods in ELT. 2nd ed. Maine: Blackwell Publishing.

*Mekheimer, M. (2011). Effectiveness of an integrated, holistic pedagogy of EFL skills in college students. *The Educational Research Journal, 25*(100), 41-74.

* Mekheimer, M., & Aldosari, H. (2013). Evaluating an integrated EFL teaching methodology in Saudi universities: A longitudinal study. *Journal of Language Teaching and Research, 4*(6), 1264-1276.

*Mokhamar, N. W. (2016). *The impact of integrating reading and writing skills on Palestine Technical College students' paragraph writing and attitudes* (Unpublished master's thesis). The Islamic University–Gaza: Gaza.

*Mubarok, H,. & Sofiana, N. (2017). Cooperative integrated reading and compsotion (CIRC). *Lingua Cultura*, *11*(2), 121-126.

Ostovar-Namaghi, S. A., & Nakhaei, S. (2019). The effect of CLIL on language skills and components: A meta-analysis. *Iranian Journal of Applied Linguistics (IJAL),* 22(1), 109-144.

Jing, W.U. (2006). Integrating skills for teaching EFL—Activity design for the communicative

classroom. *Sino-US English Teaching*, 3(12), 45-54.

*Joseph, N. (1984). *Integrated language skills: An approach to developmental studies.* (ERIC Document Reproduction Service No. ED241095).

Hoang, V. N. N., & Ngoc, T. H. (2021). The Effects of Integrating Listening and Speaking Skills Into Moodle-Based Activities. *Advances in Social Science, Education and Humanities Research*, 533, 166-174.

Norris, J. M. & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning, 50*(3), 417-528.

* Noyce, R. M. (1983). Effects of an integrated approach to grammar instruction on thirds grades' reading and writing. *The Elementary School Journal, 48*(1), 63-69.

Nunan, D. (1989). Designing tasks for the communicative classroom. Cambridge: Cambridge University Press.

Oxford, R. (1994). Integrating the language skills. *System*, 22(2), 257-268.

Oxford, R. (2001). Integrated skills in the ESL/EFL classroom. *ERIC Digest*. Retrived from http://www/cal.org/ericcll/0105oxford.html.

Oswald, F. L., & Plonsky, L. (2010). Meta-analysis in second language research: Choices and challenges. *Annual Review of Applied Linguistics, 30*(2010), 85-110.

Plonsky, L. (2011). The effectiveness of second language strategy instruction: A meta-analysis. *Language Learning, 61*(4), 993-1038.

Plonsky, L. & Oswald F. L. (2012). How to do a meta analysis. In A. Mackey & S. M. Gass (Eds.), *Research methods in second language acquisition: A practical guide* (pp. 275-296). UK: Wiley.

Plonsky, L. & Oswald F. L. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning, 64*(4), 878-912.

Richards, J., & Rodgers, T. (2001). Approaches and Methods in Language Teaching. New York: Cambridge University Press.

Robey, R. R. (1998). A meta-analysis of clinical outcoms in the treatment of aphasia. *Journal of Speech, Language and Hearing Research, 41*, 83-95.

Rothstein, H., Sutton, A. J., and Bornestein, M. (2005). *Publication bias and meta-analysis: Assessments and adjustments.* Chichester: John Wiley and Sons.

Stern, H. H. (1992). *Issues and options in language teaching*. Oxford: Oxford University

Press.

*Stevens, R. J., Madden, N. A., Slavin, R.E., & Farnish, A.M.(2016). Cooperative integrated reading and composition: Two field experiments. *International Reading Association, 22*(4), 433-454.

*Simmons, D. C., Kameeuni, E. J., Dickson, S., Chard, D., Gunn, B., & Baker, S. (1994). Integrating narrative reading comprehension and writing instruction for all learners. In C.k. Kinzer, D. J. Leu, J. A. Peter, L. M. Ayre, D. Frooman (Eds.), *Multidimensional aspects of literacy research, theory, and practice: Forty-third yearbook of The National Reading Conference* (pp. 572-582). Chicago, IL US: National Reading Conference, Inc.

*Spada, N., Jessop, L., Tomita, Y., Suzuki, W., & Valeo, A. (2014). Isolated and integrated form-focused instruction: Effects on differnt types of L2 knowledge. *Language Teaching Research, 18*(4), 453-473.

Su, Y. C. (2007). Students' changing views and the Integrated-Skills. *Asia Pacific Education Review, 8*(1), 27-40.

*Tajan, M. H. (2016). Integrated listening activities on EFL learners' speaking fluency. *International Journal of English Language Teaching*, *4*(7), 26-33.

*Tajan, M. H., Sadeghi, B., Rahmany, R. (2015). The effect of integrated listening activities on EFL learners' speaking accuracy. *International Journal of Language Learning and Applied Linguistics World, 9*(4), 74-82.

Tajzad, M., & Ostovar-Namaghi, A. (2014). Exploring EFL learners' perceptions of integrated skills approach: A grounded theory. *Canadian Center of Science and Education, 7*(11), 92-98.

*Tavil, Z. M. (2010). Integrating listening and speaking skills to facilitate English language learners' communicative competence. *Procedia Social and Behavioral Sciences*, *9*(2010), 765–770.

Wang, X., & Tragant, E. (2019). The effect of written text on comprehension of spoken English as a foreign language: A replication study. *International Review of Applied Linguistics in Language Teaching*.

Widdowson, H. G. (1978). *Teaching Language as Communication.* Oxford: Oxford University Press.

***Notes on Contributors:***

**Seyyed Ali Ostovar-Namaghi** is currently a full-time associate professor of TEFL at the department of applied linguistics, Shahrood University of Technology (SUT), Iran. He teaches both graduate and undergraduate courses including language teaching methodology, research methodology, materials development, and EAP. His chief research interest is language teacher education, grounded theory, and theories of practice. He has published in a number of leading peer-reviewed journals. He is also a member the editorial board of some journals in applied linguistics and language teaching.

**Fatemeh Ahmadnattaj** is an MA TEFL graduate from Shahrood University of Technology, Iran. Currently, she is a language teacher. She has been teaching language skills at private language schools in the past few years.