In the Name of God
The Beneficent, the Merciful

English Language Department
M.A. Thesis in English Language Teaching

Integrated Skills Approach: A Meta- analysis

By:
Fatemeh Ahmadnataj Kasgari

Supervisor:

Dr. Seyyed Ali Ostovar Namaghi

Advisor:

Dr. Fatemeh Mozaffari

October 2020

Dedication

This thesis is wholeheartedly dedicated to my loving family who encouraged and motivated me throughout the fulfillment of this thesis.

Acknowledgement

First and foremost, praises and thanks to God, the almighty, for his shows of blessing throughout my research work and helping me complete this research work successfully. Second, I would like to express my deep and sincere gratitude to my thesis supervisor Dr. Seyyed Ali Ostovar-Namaghi for his motivation, enthusiasm, immense knowledge, guidance, and instruction.

تعهدنامه

اینجانب فاطمه احمدنتاج کسگری دانشجوی دوره کارشناسی ارشد رشته آموزش زبان انگلیسی دانشگاه صنعتی شاهرود نویسنده پایان نامه «Integrated Skills Approach: A Meta-analysis» تحت راهنمایی آقای دکتر سید علی استوار نامقی متعهد می‌شوم:

- تحقیقات در این پایان نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است.
- در استفاده از نتایج پژوهش‌های محققین دیگر به مرجع مورد استفاده استناد شده است.
- مطالب مندرج در این پایان نامه تاکنون توسط خود یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ جا ارائه نشده است.
- کلیه حقوق معنوی این اثر متعلق به دانشگاه صنعتی شاهرود است و مقالات مستخرج با نام «دانشگاه صنعتی شاهرود» و یا «Shahrood University of Technology» به چاپ خواهد رسید.
- حقوق معنوی تمام افرادی که در به دست آمدن نتایج اصلی رساله تأثیرگذار بوده‌اند در مقالات مستخرج شده از رساله رعایت می‌گردد.
- در کلیه مراحل انجام این رساله، در مواردی که از موجود زنده (یا بافت‌های آنها) استفاده شده است ضوابط و اصول اخلاقی رعایت شده است.
- در کلیه مراحل انجام این رساله، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته یا استفاده شده است اصل رازداری، ضوابط و اصول اخلاق انسانی رعایت شده است.

تاریخ

امضای دانشجو

مالکیت نتایج و حق نشر

- کلیه حقوق معنوی این اثر و محصولات آن (مقالات مستخرج، کتاب، برنامه‌های رایانه‌ای، نرم‌افزارها و تجهیزات ساخته شده است) متعلق به دانشگاه صنعتی شاهرود است. این مطلب باید به نحو مقتضی در تولیدات علمی مربوطه ذکر شود.
- استفاده از اطلاعات و نتایج موجود در این رساله بدون ذکر مرجع مجاز نمی‌باشد.

Abstract

Integrated skills approach as a way of teaching language in which language skills are incorporated and interwoven with each other during teaching has become a dominant trend in language teaching. The goal of this research was to compute the overall effect size of integrated skills approach (ISA) as well as the sources and extent of variance in observed effects. To this end, the study carried out a meta-analysis of the effects of integrated skills approach on learners' language proficiency in which 22 experimental and quasi-experimental studies from peer-reviewed journals published from 1983 to 2019 of and 46 effect sizes were reviewed and synthesized. Results indicated an overall effect size of 1.18 demonstrating a strong effect size. Also, the effect sizes of moderator variables were calculated and it was reported that the integrated skills approach has the most effect in advanced learners and university level. These findings indicate that although they might improve their main language skills, vocabulary and grammar are not influenced significantly. The findings of this meta-analysis have implications for EFL/ESL teachers, researchers, policymakers and curriculum developers.

Keywords

Integrated skills approach (ISA), English proficiency, research synthesis, meta-analysis, effect size.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

EFL……English as Foreign Language

ESL……..……..English as Second Language

ISA…….……..……..Integrated Skills Approach

CLIL……Content and Language Integrated Learning

SLA…….……..Second Language Acquisition

SI…….……..Strategy Instruction

CHAPTER ONE: INTRODUCTION

1.1. Overview

Over the last decades, there has been a longstanding debate over the issue of whether the language should be taught as a set of discrete skills or should be taught as a whole. It is crucial to teach English in a way that not only improve their linguistic competence but also enables learners to communicate. For decades, taking the impact of the audio-lingual method, it had taken for granted that teaching language should be done in a segregated manner and component skills. This skill component view is rooted in audio-lingual (ALM) method and the rationale behind this view was that the advocates of ALM considered language proficiency as a set of different skills and learning the English language is equal with learning its vocabulary, grammar, pronunciation, and the four main skills: listening, speaking, reading, and writing. But, is it sufficient to be proficient in the English language? In the contemporary world of the second and foreign language teaching, most of the professionals consider it as a fact to teach language, it should be divided into discrete skill sets, typically reflecting speaking, listening, reading, and writing, and usually arranged in this order (Tajzad & Ostovar-Namaghi, 2014).

In a segregated-skills approach which is also called the language-based approach, the focus is on learning and becoming proficient at language skills not being able to use it in authentic communication (Oxford, 2001). The main principle of this segregated approach is "divide and rule" (Widdoson, 1978, p.144). In this approach, all components of language, including main skills such as speaking, reading, listening, writing, and sub-skills such as grammar, pronunciation were taught separately. In other words, this traditional mode of teaching was skill-oriented (Oxford, 1994). The problem is that in reality, it is rare to use a language skill in isolation of other skills (Hinkel, 2010). In real-life situations, they need to use a combination of skills to create and understand messages. English learners learned these skills appropriately, but they were not able to intermingle them and make a connection between these skills to communicate and use them to create messages. Because they didn't have the opportunity to learn and practice the language in an intermingled way and make a connection between these components. Consequently, they trained learners who were linguistically competent but communicatively incompetent (Tajzad & Ostovar-Namaghi, 2014).

Later this custom of teaching language was questioned. This segregated skills instruction was criticized by some scholars (Widdowson, 1978; Aldosari, 2016) and it was questioned whether language proficiency is divisible into components. They asserted that even if it is possible, this segregation of the skills is not compatible with the nature of first language acquisition. They said even in first language acquisition all language skills are interrelated. In this vein, even the way of assessing segregated skills was questioned. Harris (1968, p. 44) states:

"What evidence do we really have, for example, to justify the neat division of most language tests into listening-speaking-writing-grammar components as the most accurate and efficient means of evaluating language 'competence'?"

By the emergence of communicative language teaching, it was emphasized to teach language skills in order to enable learners to communicate by means of target language both inside and outside of the classroom. "CLT and its subsequent methodological offshoots have presently come to dominate integrated approaches to the teaching of the central four skills" (Hinkel, 2010, p. 6).

The introduction of the notion of *communicative competence* by Hymes (1972) changed the perspectives towards language teaching and emphasized the role of social context in communication. *Communicative competence* refers to the ability to interpret discourse and the learners' aim of language learning is acquiring this competence. According to Widdoson (1978) if the aim of language learning is to develop communicative competence, then an integrated approach is needed and instruction should be done in form of discourse and any approach which treats the different skills in isolation should be avoided. In addition to linguistic competence, communicative competence should be taken into account. According to Hymes (1979), there are rules of use without which the rules of grammar will be useless. On the other hand, developing communicative competence is possible in the context of interaction. Nevertheless, it was nearly impossible to use language as a means of communication in ALM classes because of the nature of ALM classes.

Considering deficits of the traditional approach and learners' communicative incompetence has led to the introduction of the integrated skills approach. Integrated skills approach is a more modern way of teaching in which teachers use two or three skills and sub-skills simultaneously to teach and improve another skill. It was introduced by Widdowson (1978) for the first time. In Widdowson's *Teaching Language as Communication*, it is mentioned that the most significant aim of learning language is to be able to communicate and interact with others. He stated that language uses take place in a social context and communication and in form of discourse. "We can talk of skills with respect to usage, but if we talk about language use, we need a different concept, and perhaps a different term" (Widdowson, 1978, p. 325). According to him, EFL learners need to learn language skills in an integrated manner to be able to mix them to use in communication.

1.2. Statement of the Problem

Considering EFL learners' communicative incompetence has led curriculum developers to seek a way of teaching which develops learners' both linguistic and communicative competence; so, integrated skills approach was introduced. But the problem is that there are lots of studies which have been conducted on the integrated skills approach both true integrated and semi-integrated and their impacts on learners' proficiency and different results are concluded and the results are inconclusive. It is because of the existence of different factors such as different sample size, time, and other conditions. Given the different factors which are intervening, it is difficult to ascertain precisely the overall effects of the integrated skills approach. 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.

1.3. Purpose of the Study

Prior researches on teaching different language skills called for integrated skills approach (Durukan, 2011; Tajan, Sadeghi, & Rahmany, 2015; Alqouran & Smadi, 2016; Oxford, 1994; Mokhamar, 2016; Aldosari, 2016; Cheong et al., 2017). Nevertheless, there is no consensus on the effect size of the integrated skills approach. The present study seeks to shed light on the previous

experimental researches and to synthesis their effect sizes to obtain a conclusive result, to determine the overall effect of the integrated skills approach (ISA). This study also examines how moderate variables associate with the effect of integrated skills approach on overall proficiency.

Research Questions:

For these aims the following research questions guided the present study:

1. What is the overall effect of integrated skills approach (ISA) on EFL learners' proficiency?
2. To what extent does the effect of integrated skills approach on learners' proficiency vary according to various moderator variables?
3. Is there any publication bias in this meta-analysis? If so, what is the level of bias in this study?

1.4. Limitations of the study

Although this study scrutinizes the rich body of previous studies and has precise design, like any other studies, has its own limitations and shortcoming. A large and growing body of literature has conducted over the effect of integrated skills approach on learners' proficiency, nevertheless some of the studies were not accessible because of the limitations of the databases. Also, some studies didn't present the required statistical data to meta-analysis. Thus, they were excluded. On the other hand, there were few studies that investigated the integration of all the four skills together. Most of the studies investigated the effect of reading and writing integration on written skills.

1.5. Delimitations of the study

In order to control intervening factors and limitations which would affect the results and generalizability of the study, some criteria were considered to include and exclude the studies. We select previous studies as our sample of study based on the inclusion and exclusion criteria. Studies had to meet the following criteria to be included in the sample:

• Studies that used a true or quasi-experimental design in which L2 learners were provided with instruction on either true integrated or semi-integrated skills approach because both of these designs lead to statistically similar results (Graham & Herbert, 2011)

- Studies that were published in English;
- Studies that contain the statistics necessary to compute the effect size;
- Studies that provide the sample size of the studied group.

CHAPTER TWO:

LITERATURE REVIEW

2.1. Overview

During the years of teaching English, there was a big shift away from the traditional skills component view (language-based teaching) towards a more modern approach, integrated skills approach, which was supposed to be more effective and more compatible with the nature of learning first language (Aldosari, 2016). The problem was that learners had acceptable knowledge of different parts of language but they had a problem in proficiency. So, there was a need for an innovative approach that could fulfill learners' desire of being proficient. Taking this problem into account and the need of an approach which provides the opportunity to EFL/ESL learners to learn English in a more authentic and effective way that makes them not only linguistically competent but also communicatively competent has led to the introduction *integrated skills approach* by some scholars (Widdoson, 1978).

2.2. Theoretical Perspectives

2.2.1. Definition of Integrated skills approach

The term Integrated skills approach is coined by Widdoson (1978) for the first time. He defines it as teaching language in a way that brings linguistic skills and communicative abilities into close association with each other. He states that even though a particular exercise may focus on a particular ability or skill, its effectiveness will require the learners to make reference to other aspects of their communicative competence and it should be consistent with the scheme of that skills and abilities.

Integrated language teaching and various offshoots of communicative language teaching are usually associated together in teaching the four main skills (Hinkel, 2010).However, several models are suggested for integrating the teaching of two or more language skills. They differ in the degree of complexity and the type of skills that are integrated to improve proficiency. In its simplest and most basic model, the skills are incorporated in the same language medium, either spoken to include listening and speaking or written to include reading and writing. A typical type of instructional program in ISA is one that deals with employing learners' receptive skills to provide input and modeling for productive skills. As an instance, listening selections may be

used as models for speaking and learners use the information, structure, pronunciation used in audio in their speaking and interaction. Or information-rich texts, reading input supplies, may be used as a model for writing. And in a more complex integrated curriculum, a combination of different language skills is used. In a more complex type, may a combination of two skills be used to improve one or two other skills. For example, listening and reading input can be used to promote speaking or writing, or to facilitate both speaking and writing. The main goal of integrated instruction is developing learners' language proficiency required for communication in various contexts through combining and making meaningful connections between different language skills.

Moreover, the integrated-skill approach is contrasted with the purely segregated approach. In this approach, English language learners are exposed to authentic language and they are challenged to interact naturally in the language. It helps them to rapidly gain a true picture of the richness and complexity of the English language as employed for communication (Oxford, 2001). Oxford (2001, p.5) declared that: "this approach emphasizes that English is not just an object of academic interest nor merely a key to passing an examination; instead, English becomes a real means of interaction and sharing among people". Integrated skills approach is consistent with communicative language teaching and is similar in that both of them focus on meaningful and authentic language use and oral and written language development improve hand in hand (Su, 2007).

Oxford (2001) claimed that integrated language teaching is supposed to be an effective strategy for language learning as a whole. According to her, plenty of threads such as vocabulary, syntax, pronunciation, spelling, and meaning should be woven well to create the integrated approach and "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" (p. 2).

2.2.2. Fundamental Theories behind Integrated Skills Approach

The current study is framed by some previous hypotheses such as unitary competence hypothesis and Communication competence model. Oller and Hinofotis (1980) stated that language proficiency can't be broken into separated and isolated parts. It can be conducted to teach propositions in a separate manner but it is not possible to create messages using propositions. The unitary competence hypothesis is contrary to the divisibility hypothesis. Divisibility hypothesis argues that language skills are discrete and separable into whether their linguistically defined components such as phonology, syntax, lexicon, or its' traditionally categorized skills such as listening, speaking, reading, and writing (Oller & Hinofotis, 1980). According to the unitary hypothesis, "the second language ability may be a more unitary factor such that once common variance on a variety of language tasks is explained, essentially no meaningful unique variance attributable to separate components will remain"(Oller & Hinofotis, 1980, p. 13).

According to Unitary competence hypothesis, linguistic competence is unitary and this unitary competence" is more than just a construct, that it really exists. This indivisible linguistic competence is the underlying principle of all language skills. All the process of comprehending and producing utterances is governed by this indivisible intellectual competence and it is common in both L1 and L2. In agreement with the unitary competence hypothesis, Widdoson (1978) has declared that a particular discrete exercise or skill will be needed to be connected and associated with other skills to be learned more effectively.

As early as the 1970s, many researchers and methodologists (Corder, 1971; Kaplan, 1970; Stern, 1992) claimed that teaching of language skills through isolated and segregated structural elements is not possible because skills are interwoven and in an authentic situation language skills are used together. "According to Howatt and Widdowson 2004: 299–300), the principles underlying language instruction between the 1950s and 1970s postulated that "all four language skills (listening, speaking, reading, and writing) should be taught, but the spoken skills should be given priority" (cited in Hinkel, 2010, p .4)." "Widdowson 1978 was one of the first linguists to call for integrating the four language skills in instruction to raise learners' proficiency levels and enable advanced language learning"(Hinkel, 2010, p. 6). "In his proposal for integrated and

communicative language teaching in general and in particular in English for specific purposes, Widdowson emphasized that virtually all language uses take place in the form of discourse and in specific social contexts" (cited in Hinkel, 2010, p. 6).

Schema theory is another theory that supports the idea of integration skills in language teaching. It proposes a cognitive basis for integrating skills instruction. It is claimed that learners must develop the background and structural schemata to compose and to comprehend texts or speak (Allen, 1987). It also argues that certain instructional strategies enhance the development of this schemata. When people read a text or write their own text during this process of making meaning in addition to thinking skills all language skills are involved to active their schemata.

Another theory is comprehensible input hypothesis. Krashen attempts to explain how second language acquisition takes place. The input hypothesis is only concerned with the acquisition", not "learning". based on this hypothesis, comprehensible input is necessary for second language acquisition and the learner improves and progresses along with the natural order 'input'(Krashen, 1985). This input should be one step beyond the learner's current stage of linguistic competence. For example, if a learner is at stage 'I' then acquisition takes place when he/she is exposed to comprehensible input that belongs to level 'I+1. Since not all of the learners can be at the same level of linguistic competence at the same time, Krashen suggests that natural communicative input is the key to design a syllabus ensuring in this way that each learner will receive some 'i+1' input that is appropriate for his/her current stage of linguistic competence (Krashen, 1985). Integrated skills approach provides a good context in which learners are provided with a great deal of comprehensible input which is a bit higher than their current level of linguistic competence.

Output hypothesis support integrated skills approach. Based on this hypothesis exposing learners to comprehensible input is not enough. It is necessary but not sufficient. They need to have the opportunity to produce comprehensible output either in spoken forms or written form and focusing on the accuracy in producing form. ISA provides a good opportunity in which students can receive comprehensible input on the one hand they are encouraged to produce comprehensible and accurate output.

Another theory is the interaction hypothesis (Long, 1981). Long (1981) in its first presentation, stated that for second language acquisition, it is necessary to participate in a conversation with native speakers or even more competent interlocutors, which is made possible through the modification of interaction. Long (1981) also clarified the definition of input and interaction: "*Input* refers to the linguistic forms used; by *interaction* is meant the functions served by those forms, such as expansion, repetition, and clarification" (p. 259). He emphasized the importance of interactive input and stated that it is far more important than input which is non-interactive. Long (1996) also noted the role of negative feedback obtained during negotiation work or elsewhere. He stated, "It may be facilitative of L2 development, at least for vocabulary, morphology, and language-specific syntax, and essential for learning certain specifiable L1-L2 contrasts" (p. 414). According to this theory interaction between learners provide a context in which ones' comprehensible output is another's comprehensible input.

2.2.3. Language Proficiency

Language proficiency is a concept that is defined differently. For example, according to Briere (1972) proficiency is "the degree of competence or the capability in a given language demonstrated by an individual at a given point in time independent of a specific textbook, chapter in the book, or pedagogical method" (p. 332). Thomas (1994) also defined it as persons' ability to perform in the second language and their overall competence in the second language. Another definition is provided by Clark (1972) as learners' ability to use language in real-life situations regardless of the manner in which that competence is acquired. But in their definition, the term competence is vague and not specified. Lado (1961) proposed a model of language proficiency including two dimensions: 1) linguistic knowledge consists of knowledge of lexis, morphology, syntax, and phonology; 2) four language skills (listening, reading, speaking, and writing). But these definitions were against Oller and Hinofotis' (1980) language proficiency model. According to his unitary competence model, language proficiency can't be broken into separated and isolated parts. Canale and Swain (1980) proposed a communicative competence model consisting of grammatical, sociolinguistic, and strategic competence to refer other aspects of proficiency. Later, this model became completed by Bachman's model in which *language ability*, is composed of organizational language knowledge (textual and grammatical knowledge), pragmatic language

knowledge (functional and sociolinguistic knowledge), and a component of strategic competence.

On the other hand, several researchers operationalized language proficiency in their studies and focused on different components of language proficiency. In one study it is defined the construct of language proficiency as "the knowledge of the language and ability to use the language in different modes (speaking, listening, reading, writing) in contextually appropriate ways" (Yamashita, 2002, p. 83). Lee and Shallert (1997) operationalized the knowledge of vocabulary and grammar structure and claimed they are the central components of language proficiency.

Considering all the above definitions in this study we operationalized language proficiency as communicative competence, the four main skills including listening, speaking, reading, writing, and components of vocabulary, grammar.

2.3. Empirical Findings

A great body of studies has been conducted on integrated skill presentations both in EFL and ESL contexts. Based on their goals, these studies could be classified into two main groups: studies seeking to investigate its effectiveness and studies that explore learners' and teachers' attitudes and perspectives.

Oxford (1994) studied the integration of language skills and noticed that it is common to use integrated language skills. Nevertheless, she found that single skill courses were still being taught. She stated that usually semi-integrated approach is used. It means skill integration occurred with two or perhaps three skills at a time, rather than all language skills at once. She found the most frequent combination of skills was reading and writing integration. "This was followed by other combinations: listening and speaking; listening, speaking, and grammar; writing and grammar; and reading and grammar and only 7.9% of the programs said they had total integration of all four language skills (reading, writing, speaking, and listening) plus grammar, 3.9% "(Oxford, 1994, p. 260).

Cheong et al. (2017) investigated how listening and reading affect the integrated writing when learners produce a piece of writing based on given information and idea form listening and reading. Integrated writing task needs participants to retrieve information and connect ideas from different two different mediums (reading and listening); therefore, comprehension of the given text and listening comprehension are para important. They found that reading and listening have a significantly positive but limited effect on integrated writing but reading cognitive skills contributed more towards the performance of the integrated writing task than what the listening cognitive skills did. They confirmed that "listening became a non- significant predictor when reading was entered in the regression models" (p. 21).

Some studies have been conducted to explore the effectiveness of integrated skills approach (Noyce & Christie, 1983; Joseph, 1984; Mekheimer, 2011; Mekheimer & Aldosari, 2013; Spada et al., 2014; Alhasan, 2018) and Some other scholars (Steven, 1987; Simmons et al, 1994; Aljiffri, 2010; Durukan, 2011; Aldosari, 2011; K. R. Alqadi & Alqadi, 2013; Cho and Brutt-Griffler, 2015; Tajan, Sadeghi, & Rahmany, 2015; Tajan, 2016; Mokhamar, 2016; Alqouran & Smadi; Aldosari, 2016; Mubarok & Sofiana, 2017; Erlidawati & Syarfuni, 2018) investigated the impact of semi-integrated models of integrated skills approach.

For example, in an experimental study (Mekheimer & Aldosari, 2013) the effectiveness of integrated EFL teaching was evaluated and it was found that skills integration had a significant effect on learners' communication proficiency especially their performance in language skills namely listening, speaking, reading, and writing.

A semi-integrated approach is used in which variations in the learners' written performance in terms of grammatical accuracy was examined. It is revealed that extensive reading considerably and positively affects the enhancement of the paragraph-writing grammatical accuracy (K. R. Alqadi & Alqadi, 2013). It is concluded that learners can more efficiently express their thought and feeling by using familiar and relevant structures and vocabulary derived from the rich reading source. Also, it is suggested that reading can decrease the degree of stress during writing because it offers the L2 writer the bulk repertoire of vocabulary, structures, idioms, discourse connectors, and stylistic devices.

Many researches (Steven, 1987; Simmons et al., 1993; Aljiffri, 2010; Durukan, 2011; Aldosari, 2011; Cho and Brutt-Griffler, 2015; Mokhamar, 2016; Alqouran & Smadi; Aldosari, 2016; Mubarok & Sofiana, 2017; Erlidawati & Syarfuni, 2018) have been conducted on the effectiveness of integrating reading and writing task on reading and writing skills. In one study, the effects of the cooperative integrated reading and composition technique on the reading and writing skills for primary school students was investigated and it is suggested that this way of teaching is more effective than the traditional method of teaching (Durukan, 2011). Durukan (2011) said there was a statistically significant difference between the reading and writing skills of the experimental and control groups in terms of academic achievement and retention.

Another researcher (Mokhamar, 2016) studied the impact of integrating reading and writing skills on students' paragraph writing including the coherent paragraph, paragraph development, unity, topic sentence, coherence, supporting sentences, cohesion, concluding sentence, and attitudes. He noted that since reading and writing are closely linked and mutually reinforce each other, therefore learning will be promoted when they are integrated into classroom activities. As a result, students become better readers, writers, and thinkers when they learn reading and writing together. He detected that integrating reading and writing skills had a great effect size and it significantly affects learners' writing performance. Also, this integration provides students with a better learning environment that positively reflected in their paragraph writing. Based on his work it is recommended that teachers as well as lecturers to integrate reading and writing activities in teaching writing skills in general and paragraph writing in particular, to develop and improve their paragraphs.

Alqouran and Smadi (2016) reported that the instructional program based on reading-writing integration has a positive effect not only on students' overall writing performance but also on the writing sub-skills. "Reading comprehension was also positively affected as a result of the integration of the two skills" (Alqouran & Smadi, 2016, p. 185). They asserted that reading writing integration is as an effective catalyst for writing performance.

In a similar vein, Aldosari (2016) noted that it is important to connect reading and writing. He used integrated instructional activities in various areas of writing skills and he found that "integrated reading has a direct influence on the quality of writing as it gives positive effects on literacy development which helps students efficiently consider the genre of writing in the learning process"(p. 761). His finding emphasized the importance of connecting reading and writing.

Aljiffri's (2010) findings indicated that the integrated approach has led to better achievement gains in literacy development and Social Studies achievement. "The gains of the integrated approach group in reading comprehension, writing, and social studies may be attributed to the emphasis placed on integrating literacy skills with content-based instruction"(p. 27).

Some researchers (Tajan, Sadeghi, & Rahmany, 2015; Tajan, 2016) investigated the effect of integrated listening activities on EFL learners' speaking accuracy and fluency within the framework of task-based activities. They confirmed that integrated technique has a great effect on teaching language skills to Iranian learners compared to the traditional approach and when these two skills, listening and speaking, are taught in an integrated manner consequently students' success increases.

Moreover, Tavil (2010) evaluated the effect of listening and speaking integration on English language learners' communicative competence. He used information-gap tasks and realized that the students were willing to actively participate in integrated tasks which created real life situations in the classroom. He found that listening and speaking integration enhance students' oral communicative competence.

Whereas some studies attempt to determine the effect of this new trend on language main skills, other studies investigated its effect on language sub-skills and components (Joseph, 1984; File & Adams 2010). The results of Joseph's (1984) study indicated that learners' performance in vocabulary has improved as a result of integrated skills instruction, whereas later in another context this result was rejected. File and Adams (2010) found that integrated skills approach has not such a positive effect on learners' vocabulary and segregated instruction is more effective. 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.

With regard to proficiency differences, Cho and Brutt-Griffler (2015) investigated the differences among 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 begging learners has not shown improvement.

Although a great number of studies (Noyce, 1983; Spada et al., 2014; Tajan, 2016; Aldosari, 2016) reported the positive effect of this approach on different skills, Gautam (2019) stated that "it is up to the syllabus designers and teachers that they should decide what is logical and illogical integration following some practical trails and testing upon the real students" (p.106).

A number of studies (Su, 2007; Tajzad & Ostovar-Namaghi, 2014; Almalki & Soomro, 2017) explored learners' and teachers' attitudes and perceptions of this approach. For example, Su (2007) explored learners' views and satisfaction with integrated skills and it was revealed that the majority of students wanted continuing to implement the ISA in class for the next academic year. They were strongly believed that the four language skills should be integrated and they are integral parts of a language course. In a similar vein, in another context, EFL learners' perceptions of the integrated skills approach were investigated (Tajzad & Ostovar-Namaghi, 2014) and it was revealed that the participants positively perceive this approach as a way of teaching the foreign language because it reduces their anxiety and increases their motivation. It is also asserted that, 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. Although segregated skills teaching may help students develop their knowledge of the language, but it does not enable them to use the knowledge in actual communication. Almalki and Soomro (2017) found that although teachers did not apply

integrated skills approach, they have positive perception of it and their problem is that they do not have a deep understanding of the way of applying skills integration.

As noted previously, the effectiveness of integrated skills approach has been investigated widely in various learning contexts and outcome variables. So, in line with this study's purposes of examining the overall effectiveness of Integrated Skills Approach research as well as determining the relationships between ISA and the moderator variables, the remainder of this literature review will present some comprehensive meta-analysis research in SLA area across the following dimensions: learning contexts, treatments, length of treatments, and outcomes. To start, it is worth to note that Plonskey (2011) Meta analyzed the effect of strategy instruction and the relevant moderator variables. He found that variables include type and number of strategies, learning context (second vs. foreign language), and length of intervention moderate the effectiveness of strategy instruction (SI). He also, examined speaking, reading, listening, writing, vocabulary, general, attitude, and pronunciation as outcome variables. In another context, Maeng (2014) investigated the effectiveness of strategy instruction on learners' reading comprehension. He found that among potential moderator variables Institution, Strategy Type, Number of Strategies, Test Type had a statistically significant moderating effect. Also, Norris and Ortega (2000) investigated the effectiveness of L2 instruction and found that the type of outcome measure affects the magnitude of the observed effect. In another study, Nakhaei (2017) Meta analyzed the effectiveness of CLIL and the findings indicated that language skills and components, subject matter, educational level, and publication type influence the impact of CLIL. Therefore, these contextual features can play a prominent role in the effectiveness of an independent variable.

2.4. Summary of Empirical Finding and Statement of the Gap

Different studies have investigated different aspects of integrated skills approach. For example, some scholars studied the effect of skills integration on writing skill (Mokhamar, 2016; Alqouran & Smadi, 2016; Steven, 1987; Simmons et al., 1994; Aljiffri, 2010; Durukan, 2011; Aldosari, 2011; Cho and Brutt-Griffler, 2015) and found that the instructional program based on reading-writing integration has a positive effect not only on students' overall writing performance but

also on the writing sub-skills.

Considering speaking skill some researchers found that (Tajan, Sadeghi, & Rahmany, 2015; Tajan, 2016) integrated technique has a great effect on teaching language skills to Iranian learners compared to the traditional approach and when these two skills, listening and speaking, are taught in an integrated manner consequently students' success increases.

The result also proved that learners' performance in vocabulary is under the positive influence of ISA (Joseph, 1984). By contrast, in another context, File and Adams (2010) found that integrated skills approach has not such a positive effect on learners' vocabulary and segregated instruction is more effective. In another study, it is proved that this approach had no significant effect on learners' grammar (Spada et al., 2014). They concluded that integrated and isolated instruction are complementary to each other and both of them positively affect L2 learning.

To date, however, there is no consensus among the overall effectiveness of integrated skills approach on overall proficiency, and also little is known about the role of moderator variables on the effectiveness of this method. The present study seeks to shed the light on the previous experimental researches and synthesize their effect sizes to obtain a conclusive result, to determine the overall effect of the integrated skills approach (ISA).

CHAPTER THREE:

RESEARCH METHOD

3.1. Overview

This chapter covers the procedure followed throughout the study and presents the methodology of this meta-analysis. In order to make readers cognizant with the process of meta-analysis, it introduces the philosophy of meta-analysis, and explains sampling procedure and materials, data collection, reliability of the study, data analysis, publication bias, and heterogeneity of the study.

3.2. The philosophy of Meta-Analysis

The term meta-analysis refers to a quantitative approach that statistically combines the results of primary empirical studies on the same subject and "provides a precise estimate of the population effect of a certain construct" (Ellis, 2015, p. 2). It aims to solve the disagreement among the literature on a subject and in this process it uses the analysis of variance and multiple regression procedures, with effect sizes treated as the dependent variable (Cooper & Hedges 2009). A Meta-analysis processes the processed data obtained from the previous analysis or studies. In other words, it is analysis of analyses (Glass, 1976). It determines how much difference an independent variable makes, therefore, the methodology of a meta-analysis is approximately similar to what happens in an empirical study, but it does not have the problem of ambiguous direction of causality and one of its interesting points is that it also calculate the effect of moderators variables (Glass, 1982), what has been neglected to examine in an original empirical study. Conducting a meta-analysis needs a prerequisite and it is existence of the primary research. According to Cooper and Hedges (2009), primary research on the topic must exist before a research synthesis can be conducted. To sum up, Meta-analysis is a "systematic review" (Rhodes, 2012, p. 25) which is based on the previous studies and tries to combine their effect sizes in order to reach a conclusive result and determine the effect direction of an independent variable.

Although meta-analysis is a type of systematic reviews, it is different from traditional narrative reviews in several ways, for example their purposes of the study are different . A mata-analysis is used to summarize the direction and magnitude the effects obtained from the previous empirical studies which investigated the same subject or phenomena (Lipsey & Wilson, 2001) whereas narrative reviews do not use a systematic search of the literature, and often focus on a subset of studies in an area chosen based on availability or author selection. They are primarily

descriptive and selective while a meta-analysis is inclusive rather than selective (Ellis, 2015). It increases the objectivity of the research literature (Smith et al., 1984). Thus narrative reviews are informative, they can have the selection bias problem, while a meta-analysis uses a systematic way to solve the problem of bias (Oswald & Plonsky, 2010). Narrative reviews can also be confusing in some cases, especially when similar studies led to divergent results and conclusions.

Considering deficiencies of empirical studies and traditional reviews there is a demand to use the new quantitative approach which provides more precise results (Glass et al., 1981; cook et al., 1992; Petticrew & Robert, 2006). So, the term of the meta-analysis was coined by Glass (1976) for the first time and it is considered as a substitution for narrative review. Meta-analysis removes subjective error and it is obtained by weighted and unweighted overall probabilities (Cooper & Rosenthal, 1980).

Different researchers suggested different steps in conducting a meta-analysis (Glass et al., 1981; Hunter et al., 1982; Cook et al., 1992; Fitz-Gibbon, 1984). For example, Cooper and Hedges (2009) suggested 6 stages model of meta-analysis: 1) Problem Formulation: start with the broadly defined problem area; 2) Literature Search: carry out a literature search and identify the set of studies to be analyzed; 3) Data Evaluation: Meta-analysis typically aims to be inclusive rather than selective in identifying the studies to be investigated on the grounds that criteria for evaluating and excluding studies are inevitably subjective; 4) Data Analysis: identifying categories to code the primary studies and aggregate all effect sizes to obtain an estimate of the population effect, followed by moderator analysis. The meta-analysts presented the results of their statistical analyses of the main and moderator variables without interpretative comments; 5) Interpretation of Results; 6) Public Presentation.

On the other hand, Uman (2011) suggested a more detailed process, including eight stages in meta-analysis:1) formulating the review question; 2) Defining exclusion and inclusion criteria; 3) Developing search strategy and locate studies; 4) Selecting studies; 5) Extracting data; 6) Disseminating studies; 7) Analyzing and interpreting results; 8) Disseminating findings.

In this study we took the eight steps model suggested by Glass et al. (1981) and Hunter et al. (1982):

> "1) Identify the variables for focus (independent and dependent);
> 2) Identify all the studies which feature the variables in which the researcher is interested;
> 3) Code each study for those characteristics that might be predictors of outcomes and effect sizes; (e.g. age of participants, gender, ethnicity, duration of the intervention);
> 4) Estimate the effect sizes through the calculation for each pair of variables (dependent and independent variable), weighting the effect-size by the sample size;
> 5) Calculate the mean and the standard deviation of effect-sizes across the studies, i.e. the variance across the studies;
> 6) Determine the effects of sampling errors, measurement errors, and range of restriction;
> 7) If a large proportion of the variance is attributable to the issues in Step 6, then the average effect- size can be considered an accurate estimate of relationships between variables;
> 8) If a large proportion of the variance is not attributable to the issues in Step 6, then review those characteristics of interest which correlate with the study effects" (cited in Cohen et al., 2007, p. 292).

3.3. Sampling Procedure and Materials

The participants of a meta-analysis are all participants of the previous study in the same subject (Little et al., 2008). Therefore, sampling procedure in meta-analysis involves selecting relevant studies and articles published in scientific journals with the same problem which present the statistical data needed in meta-analysis, rather the selecting participants.

In this study a wide and various sets of techniques were employed. An in-depth search was done in Google Scholar and Cross-Ref search engine. An electronic search was conducted (ending in June 2018) in four online databases including ProQuest, ERIC, and Elsevier. We, also, hand-searched Tesol, Wily, and Oxford peer-reviewed journals and we obtained different experimental studies, including articles published in non-freed journals, MA thesis, doctoral dissertations, and books are retrieved from these various online open-access databases. The rational behind including MA thesis and doctoral dissertations is that they "act as a brake on misleading, more

spectacular generalizations" (Glass et al., 1981, p. 5) and it causes the nullification of probability of publication bias. Because they may not get published due to the lack or weak correlation with published research.

It is obvious taht the results of the search depends on the keywords which are used to search. Given the importance of selecting an appropriate keyword, in searching relevant literature we used a list of different combinations of keywords including integrated skills approach, Integrating skills instruction, four skills instruction, Language skills integration, The effect of integrated skills on, and The impact of integrating skills on. Then, we retrieved and reviewed a comprehensive list of abstracts, anystudies appearing to meet inclusion criteria would then be obtained and reviewed in full and those don't meet the inclusion criteria 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, we investigate the reference list to identify other published studies. Our search revealed 29 studies as candidate studies to be included in the meta-analysis. Of the 25 documents collected, 22 studies on integrated skills approach met the inclusion criteria.

The previous studies were selected as our sample of study based on the inclusion and exclusion criteria. Studies should meet the following criteria to be included in the sample:

1) Should be true or quasi-experimental studies 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, for example, means and standard deviations for pre- and post-test scores for intervention and comparison group means, t-test values and group sizes, one-way ANOVA for the groups, F-values and sample sizes, one-way ANCOVA with error values and correlation (*r*) between the covariate, and dependent measure.

Table 3.1. Features of the Studies Included in the Meta-Analysis

Characteristic								Total
Publication year of Research		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			Journal Article			N=22
	N	2			20			
	%	9			91			
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	

In the next step, we read the full text of the selected studies. Finally, 7 articles failed to meet the inclusion criteria and were excluded because of the lack of needed data or other reasons, for example, they don't have the control groups and 22 articles were recognized as the appropriate items for data analysis. The remaining studies were analyzed precisely. Descriptive statistics of those 22 studies are presented in Table 3.1. Also, the studies as participants of the current study are depicted in Table 3.2.

Table 3.2. Studies as participants of the study

Authors	Country	Participants	Type of Documentation
Abrams (2019)	Germany	23	Article
Aldosari (2011)	Saudi Arabia	62	Article
Aldosari (2016)	Saudi Arabia	62	Article
Alhasan (2018)	Sudan	44	MA Thesis
Aljiffri (2010)	Saudi Arabia	64	Article
Alqouran & Samadi (2016)	Jordan	50	Article
Cho & Brutt-Griffler (2015)	USA	181	Article
Durukan (2011)	Turkey	45	Article
Erlidawati & Syarfuni (2018)	Indonesia	63	Article
File & Adams (2010)	New Zealand	20	Article
Joseph (1984)	USA	51	Article
Mekheimer & Aldosari (2013)	Saudi Arabia	52	Article
Mekheimer (2011)	Saudi Arabia	52	Article
Mokhamar (2016)	Palestine	80	MA Thesis
Mubarok & Sofiana (2017)	Indonesia	68	Article
Noyce (1983)	USA	41	Article
Simmons, Kameeuni, Dickson, Chard, Gunn & Baker (1994)	USA	93	Article
Spada, Jessop, Tomita, Suzuki & Valeo (2014)	Canada	129	Article
Stevens, Madden, Slavin & Farnish (2016)	USA	450	Article
Tajan (2016)	Iran	30	Article
Tajan, Sadeghi & Rahmany (2015)	Iran	30	Article
Tavil (2010)	Turkey	180	Article

3.4. Data Collection

After all studies were retrieved for analysis, the coding process was done to translate critical study information into coded form. In this step, the findings and the methodological and substantive characteristics of studies should be evaluated and coded. This step is time-consuming and needs more expertise and a close inspection (Smith et al., 1984). In addition to coding and identifying the characteristics of the study, any meta-analysis should be ended with computing moderator

effects since these variables (moderator variable) may affect overall effect size estimation through co-variation with the independent variables of interest (Cole, 2014). Having more than one research assistant to screen and select papers, ensure the reliability of study selection. Two coders independently coded all studies for four potential moderator variables that were thought to be theoretically relevant to the outcome of the studies. In order to facilitate the process, this step was done by computer-assisted. The statistical data was imported into Microsoft Excel file. This facilitates the analysis process because data will directly be imported into the software of analysis to calculate the effect size. It shortens the consumed time of analysis. It also leads information to become categorized and well organized. Coding forms capture identifying information on studies, characteristic of participants, ethnicity, groups information (1) year of publication; (2) publication source (book, thesis, journal); (3) subject taught (reading, math, etc.); (4) duration of instruction; (5) number of pupils in the study; (6) number of teachers in the study; (7) pupil ability; (8) pupil ages; (9) types of experimental control (random assignments, matching, etc.); (10) achievement measurement; (11) quantification of outcomes (Glass, 1982).

The specific extracted data from retrieved studies are categorized as follows:

- Name of the study
- Name of author
- Publication year
- Country of study
- Contact hour
- Gender of participants
- Age of participants
- Type of study
- Design of studies
- Data collection instrument
- L2 proficiency
- Educational level
- Dependent variable
- Independent variable
- Sample size
- Control group sample size

- Experimental group sample size
- Mean of pretest and posttest in each group (control and experimental group)
- The standard deviation of pretest and posttest in each group (control and experimental group)
- P-value
- T-value
- Pretest and posttest correlation

3.5. Reliability

One of the risks that threats meta-analysis reliability is that coder may enter his personal judgment or bias in coding and classifying data; as a result, the meta-analysis will be biased and suffer subjectivity. To guarantee the reliability of the meta-analysis and prevent data entry error, it is important to use a coding scheme and data extraction by multiple raters to establish inter-rate reliability (Bullock & Svyantek, 1985). As Stock and colleagues noted, such coding may have low reliability unless adequate precautions are taken to guarantee high reliability (Stock et al., 1982). Based on what was suggested by Plonsky and Oswald (2012) coding reliability was assessed through the measurement of inter-rater reliability. Therefore, two coders coded separately all samples in order to ensure inter-rater reliability. The coders were MA students of TEFL at applied linguistics department in Shahrood University of Technology. They were trained in two sessions to become cognizant of the process of coding. It is better if the coders are expert in that major because they can efficiently code and modify the variables. Then, Cohen's kappa reliability coefficient was calculated between the coders by using SPSS software. The Cohen's kappa reliability index was determined as 0.93 which is interpreted as the agreement between two independent coders is more than 0.93 % and it means a perfect consistency exists between the raters. Finally, disagreements were discussed and resolved and coding sheets were checked again and were corrected based on the common agreement.

3.6. Instrument

There are various programs and software available to compute effect sizes and to do meta-analysis such as, the Comprehensive Meta-Analysis (CMA), EPPI-Reviewer, and Review Manager (RevMan) program endorsed by the Cochrane Collaboration. In this study, we used CMA software

to calculate effect sizes because it is the most appropriate one to do a meta-analysis. It enables the researcher to enter various types of data such as odd ratio, standard deviation, pretest ad posttest correlation, and so on. Moreover, through the use of CMA3 forest plot, funnel plot of both observed study and imputed study, individual effect size, main effect size can be provided.

3.7. Data Analysis

3.7.1 Individual Effect Size and the Main Effect Size

Calculating the effect size of each individual study based on the extracted statistical information is the next step. It is obvious that different studies use different ways of presentation and interpretation of obtained differences between the experimental group and the control group. For instance, p-value, z value, F value, Mean, and standard deviation are the statistical terms may be used in reporting their results. They also may calculate the effect size of the study but since we do not know how they formulate and which formula they used to reach the value, we can't trust it and it is not reliable. So, we independently calculate the effect size. Through calculating the effect size of each study, the integration of all results in different studies to an identical scale can be obtained. Therefore, the process of comparison and combining the results of the studies with each other would be much more practical, rigorous, and easier. Effect size (standard mean difference), is a simple way of estimating the difference between means of the two groups, it tells you how much the experimental group of a study gained different achievement from its control group (Bornstine, Hedges, Higgin, & Rothestine, 2009). " Effect size is a measure of the degree to which a phenomenon is present or the degree to which a null hypothesis is not supported" (Cohen et al., 2007, p. 293). Standardized mean differences and correlations are the two most frequently used indices of effect size (Cohen et al., 2007).

Several different measures have been developed for estimating the effect size of experimental studies. The main measures for the effect size of studies are, Cohen's d, Hedges' g, and Glass's delta which in large samples (more than 20) they will be approximately equal. But when the sample size is small especially when it is below 20 they will be different because the type of standard deviation they use in the denominator is different (Decoster, 2009). Therefore, Hedges' g produces a more precise estimate and is preferable to Cohen's d when the sample size

is below 20. Hedges'g tells us how much one group differs from another, the difference between an experimental group and a control group. Hedges' g is preferable to Glass's delta because contrary to Glass's delta which uses only standard deviation of the control group, it uses pooled standard deviations from both groups (Control and Experimental group) and it will lead us to a better estimation of the population standard deviation. In cases where, the standard deviation is significantly different between the groups, Glass's delta will be recommended. It is also preferable to Cohen's d because it uses the unbiased least squares estimate of the pooled standard deviation. In other words, it uses pooled weighted standard deviation rather than pooled standard deviations and it is weighted based on the study's sample size and standard error. The greater the sample size the more weight study is. Furthermore, most of the analyses from which you will be deriving standardized mean difference effect sizes, such as t-tests and ANOVA, are based on the unbiased least squares estimate of the pooled standard deviation. This makes it much easier to calculate Hedges' g than Cohen's d from their statistics.

Taking the advantages of Hedges' g over other indices, in this study we used it to calculate effect size which its formula:

$$\text{Hedges } g = \frac{\bar{M}_1 - \bar{M}_2}{SD}$$

In this formula $\bar{M}_1$ is the mean of the experimental group, $\bar{M}_2$ is the mean of the control group, and S_p is pooled and weighted standard deviation.

Hedges' g (like Cohen's d) is biased upwards for small samples (under 50). The effect sizes have also been weighted and corrected for small sample sizes using the following formula:

$$g = \frac{M_1 - M_2}{SD^{*}pooled} \times \left(\frac{N-3}{N-2.25}\right) \times \sqrt{\frac{N-2}{N}}$$

Despite the little difference in Cohen's d and Hedge's g formula, they are interpreted in a similar way. An effect size can take the value between 0 to 1. According to Cohen the interpretation of results can be done in regarding the following rule (Cohen, Manion & Marrison, 2007):

- $0 \leq$ Effect size value (Cohen's d or Hedge's g) ≤ 0.20 insignificant effect;
- $0.21 \leq$ Effect size value (Cohen's d or Hedge's g) ≤ 0.50 small effect;
- $0.51 \leq$ Effect size value (Cohen's d or Hedge's g) ≤ 0.8 medium effect;
- $0.81 \leq$ Effect size value (Cohen's d or Hedge's g), strong effect

However, the following scale is suggested to interpret the effect sizes in L2 research (Plonsky & Oswald, 2014):

- $0 \leq$ Effect size value < 0.40 insignificant effect;
- $0.4 \leq$ Effect size value < 0.70 small effect;
- $0.7 \leq$ Effect size value < 0.1 medium effect;
- $0.1 \leq$ Effect size value, strong effect

The contributed effect size of each individual study was calculated. In cases that the study examined the effect of integrated skills approach on the achievement of different groups or skills and reported the related data, more than one effect size was calculated for these studies. In the end, the main effect size was calculated in order to combine the effect sizes of the studies to decide on the effectiveness of Integrated Skills approach and its effect direction. In other words, we determined whether this approach has a positive effect on students' overall proficiency and if it shows the positive effect, what is the magnitude of effect on their language proficiency.

3.7.2. Test of Homogeneity

The preceding step to calculating or analyzing the effect sizes of studies is selecting the analysis model for the calculation of effect size. The degree of heterogeneity determines which model should be used. Q statistic and I^2 are used to determine the degree of heterogeneity. Q is the sum of squares of all effects from the mean effect on the standardized scale and df (degree of freedom) is simply the number of studies in analysis minus 1. If all studies are based on the same population and all of the variance in observed effects is due entirely to sampling error, then the value of Q that would expect to see, is equal to the degree of freedom. In other words, if Q value and degree of freedom are equal, there is no heterogeneity (dispersion in observed effect).

The importance of selecting the appropriate model is that it will ensure the accuracy of estimation of various statistics because it affects the computations. Additionally, and more

fundamentally, the model serves to place the analysis in context. It affects the calculation of the effect size. It also, provides a framework for inferences and interpretation you draw from your results (Decoster, 2009). There are two statistical models for meta-analysis, fixed-effect model, and random effect model (Borenstein et al., 2009).

Fixed-effect model: The underlying assumption of the fixed-effect model is that there is one true or common effect size that underlies all the studies in the analysis. And the only source of variation is within-study variation thus the effect size is not affected by any other variable. Consequently, all differences in observed effects are because of this within-study variation which is sampling error. Except for sampling error, all studies are identical in treatment, method, measurement, sampling procedure, and participants which should provide identical results. As a result, the effect sizes from a homogenous set of studies can be combined and the mean effect size will be estimated for a single population. So, this model is recommended when there is homogeneity in the studies.

Random-effects model: The underlying assumption of this model is that either the true effect sizes can vary from study to study or all studies can share a common effect size. This model considers both within-study and study to study variation. It assumes some part of this study to study variation is because of measurement error and some of this variation is because of the actual differences among studies. This assumption implies that all studies involved in analysis share different but related, treatment effects, with the differences between these represented by random variation. For example, studies with the older participants or more educated participants may have higher or lower effect sizes. In other words, studies involved in the analysis are heterogeneous. In this model, we treat these two sources of variation differently when calculating the weighted mean effect size using a random-effects model. Therefore, meta-analysis imputes the mean and standard deviation of different effects.

Choice of Model

In cases that statistical heterogeneity is recognized, it is normally recommended to use random-effects models. “The impact of interventions is always likely to vary by chance, but testing for heterogeneity investigates whether there is more variation than would be expected by chance

alone"(Higgins et al., 2005, p.16). Taking these points into account, in this study, we used the random-effects model.

3.7.3. Moderator Analysis

Research Question 2 referred to variability in observed effects which is due to moderating variables. After computing the effect size of each individual study and the *mean effect size* (the average effect size) we investigated the potential effect of different moderator variables of interest through a random model. Because by moderator analysis, we can determine whether study characteristics significantly affect overall effect size estimates. Moderator variables can cause heterogeneity through moderating the strength of the effect between studies. They covariate with the independent variable of interest and consequently affect overall effect size. It is; therefore, necessary to analyze and calculate the effect of moderator variables. In this study, the moderator analysis is done to find statistical differences among sub-groups and among average effect sizes of the variables. Based on these differences we are able to compare groups. Considering the objective and procedure of the study, subgroup analysis is planned (Little, Corcoran & Pillai, 2008). Age, Publication type, Length of treatment (Contact hour), Language skills and sub-skills (listening, reading, writing, speaking, Grammar, communicative competence, general), Educational level, and Proficiency level (language level) were analyzed as moderator variables in this meta-analysis study. Each of these moderator variables was analyzed separately.

3.7.4. Publication Bias Evaluation

One potential threat of systematic reviews that researchers should consider is publication bias. Publication bias was originally defined as the direction and statistical significance of the results determine publication or non-publication of studies (Rothstein et al., 2005). In other words, studies showing a positive (beneficial) or statistically significant effect are more likely to be published than negative or inconclusive studies. Consequently, a part of the literature on the subject will be missed and literature on primary studies will be unrepresentative of the population of completed studies. Since meta-analysis is based on the previous empirical studies

and its aim is to synthesis the result of all previous studies, losing some parts of the literature causes inflation in the estimation of the effect. There are various statistical methods for determining the existence of publication bias in meta-analysis and assessing its effect on the analysis. Funnel Scatter plot and Classic Fail-Safe N were used to evaluate publication bias in this study.

3.7.4.1. Funnel Scatter Plot

A funnel plot is a graphical diagnostic tool which helps us visually and simply detect the existence of publication bias. Since it presents information visually, the results of the evaluation by this method are not objective and numerical. In a Funnel Plot, the X-axis shows the value of effect size and the Y-axis shows standard error. The studies with a greater sample size and more precise studies place at the top of the plot and less precise study with a smaller population place at the bottom of the plot. The interpretation of the Funnel Plot is based on being symmetric or asymmetric. If Funnel Plot is symmetric it shows there is no publication bias and our meta-analysis had taken all relevant studies. And if it is asymmetric it shows the existence publication bias. In other words, if the density of distribution around the main effect size is high and it is almost symmetric we conclude there is no publication bias in studies and there is no lost study.

3.7.4.2. Trim and Fill Method

According to Duval (2009), The Trim and Fill is a nonparametric technique and an explicit statistical test for the statistical significance of publication bias. This method is applied to estimate and adjust the number of missing studies (due to publication bias) in a meta-analysis. In other words, this procedure imputes these missing studies, adds them to the analysis, and then re-computes the summary effect size.

3.7.4.3. Fail-Safe N Test

Classical Fail-Safe N is another method to evaluate publication bias (Cooper, 1979).To estimate how many insignificant effect sizes (unpublished data) would be necessary to reduce the overall effect size to an insignificant level, the fail-safe N (i.e., classic fail-safe N) was used. In other

words, it evaluates the number of undiscovered or missing studies that are required to nullify the observed effect size, which means to reduce the main effect size value to 0.01 and to bring the significant level down to p = 0.5. The comparison criterion was 5n10, where n is the number of studies included in the meta-analysis. If the obtained fail-safe N is larger than 5n10, it means that the estimated average effect size of unpublished studies is unlikely to influence the effect size of the meta-analysis and H hypothesis is not rejected. The more the number of required studies, the less probability of publication bias exists.

CHAPTER FOUR:

RESULTS

4.1. Individual and Main Effect Size Analysis

Research Question 1: What is the overall effect of integrated skills approach (ISA) on EFL learners' proficiency?

To examine the benefit of integrated skills approach on students' proficiency we calculated the overall effect size. The mean effect size, degree of freedom (df), 95% confidence interval, P-value, and some other statistic based on the random model are presented in table 4.1. The estimated overall effect size for the random model is 1.18 of 95% confidence interval and it is considered as incredibly effective according to Cohen (1987) interpretation. The overall positive effect size demonstrates a positive effect of integrated skills approach on students' overall proficiency. Preceding to look at the mean effect size we should know the amount of P-value. The corresponding P-value is 0.00 (is less than 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). Q is the sum of squares of all effects from the mean effect on the standardized scale and df (degree of freedom) is simply the number of studies in analysis minus 1. If all studies are based on the same population and all of the variances in observed effects is due entirely to sampling error then, the value of Q that would be expected to see, is equal to the degree of freedom. In other words, if Q value and degree of freedom are equal, there is no heterogeneity (dispersion in observed effect). The significant Q statistic result (Q (46) =848.16, $p < 0.05$), shows heterogeneity of effect sizes. So, it suggests the need of using the random-effect model for subsequent analysis. Moreover, that I^2 value in table 4.1 is bigger than 75% shows that the distribution of the effect values of studies on integrated skills approach is highly heterogeneous.

Table 4.1. 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				

The distribution of the effect sizes of 46 studies is presented in Figure 4.1, the forest plot. In forest plot, the effect size for each individual study is calculated. Each individual effect size is represented visually by square; the squares and confidence interval bars illustrate the estimated precision of each study. Also, the mean effect size for the entire sample is visually represented by the diamond. According to Figure 4.1 among the effect sizes, Mekheimer (2013, d) had the largest effect size (4.915), and File and Adams (2010) had the smallest effect size (-0.579). Meanwhile, just one study had a negative effect size.

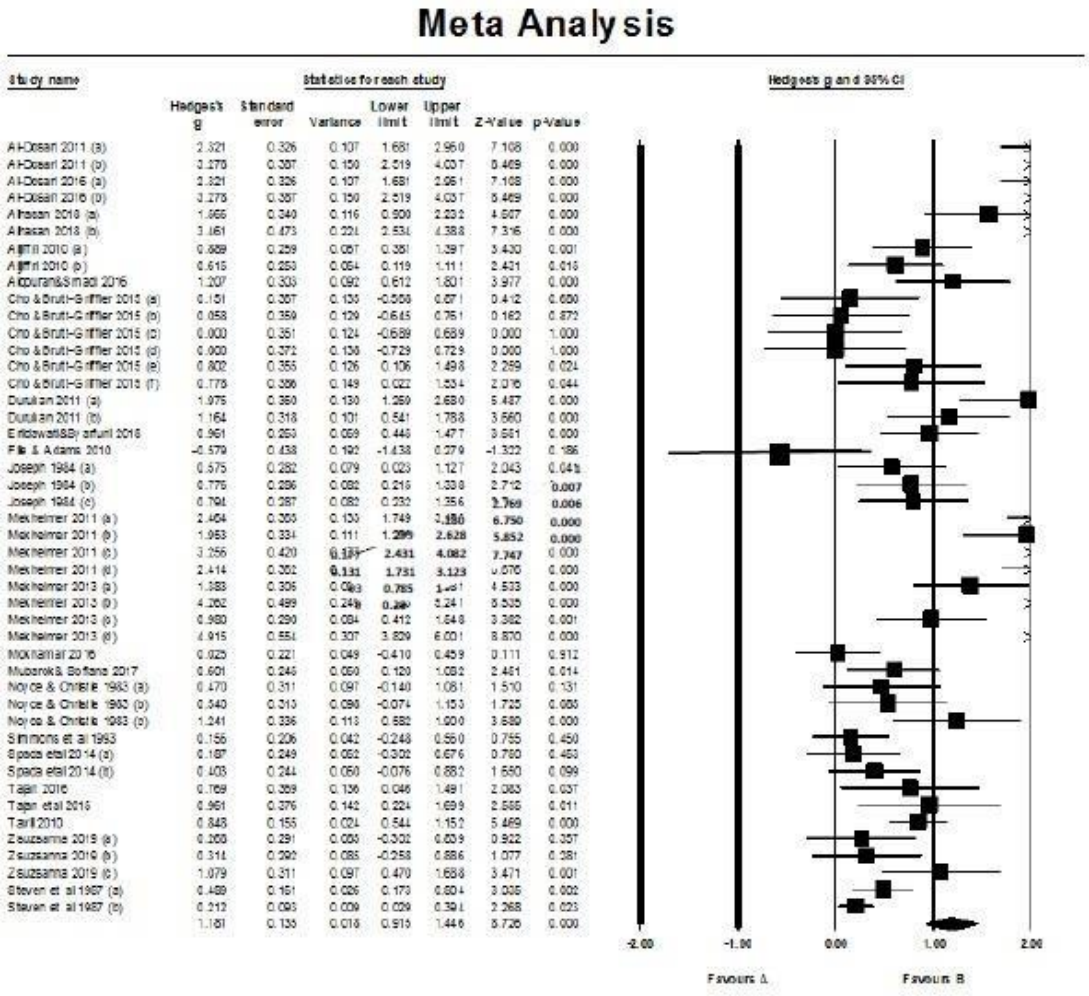

Figure 4.1. Forest plot of the effect sizes (k= 46)

4.2. Moderator Analysis (Sub-group Analysis)

Research Question 2: To what extent does the effect of integrated skills approach on learners' proficiency vary according to various moderator variables?

The distribution of effect sizes was heterogeneous as indicated by Q=484.16 and I^2= 90.70 in the previous section; consequently, there was a demand for analyzing moderator variables as a source of variation to explain the observed variance in the effect sizes. 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 including Age, Publication type, Length of treatment (Contact hour), Language skills and sub-skills (listening, reading, writing, speaking, Grammar, communicative competence, Linguistic complexity), Educational level, and L2 Proficiency level. In order to detect statistical differences between subgroups and between the average effect sizes of the variables, this analysis is done. The distribution of moderator variables and their corresponding effect size (g) of ISA on learners' proficiency is presented in Table 4.2.

Table 4.2. Moderator analysis on the effectiveness of ISA

Moderator variables	Number of calculated effect sizes	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	0.71	0.33	1.22	0.22
University	18	0.04	9.97	0.00	1.59	1.60	2.38	0.20
Skill								
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

4.2.1. Educational Level

In order to examine whether the effect levels of ISA on students' proficiency differentiate according to educational level, the studies included in this research were classified under three

groups as primary, secondary, and university level. The average effect size for different groups including primary, secondary, and university levels were 0.63, 0.78, and 1.99 respectively. Our results, presented in Table 4.2, indicated that learners under integrated skills instruction had significantly a large effect size at university level (g=1.59, SE=0.20, CI_{95}= 1.60, 2.38) and learners at secondary level had a medium effect size (g=0.71, SE=0.22, CI_{95}= 0.33, 1.22), and primary level (g=0.41, SE=0.06, CI_{95} =0.30, 1.22) had a small effect size on Cohen's scale.

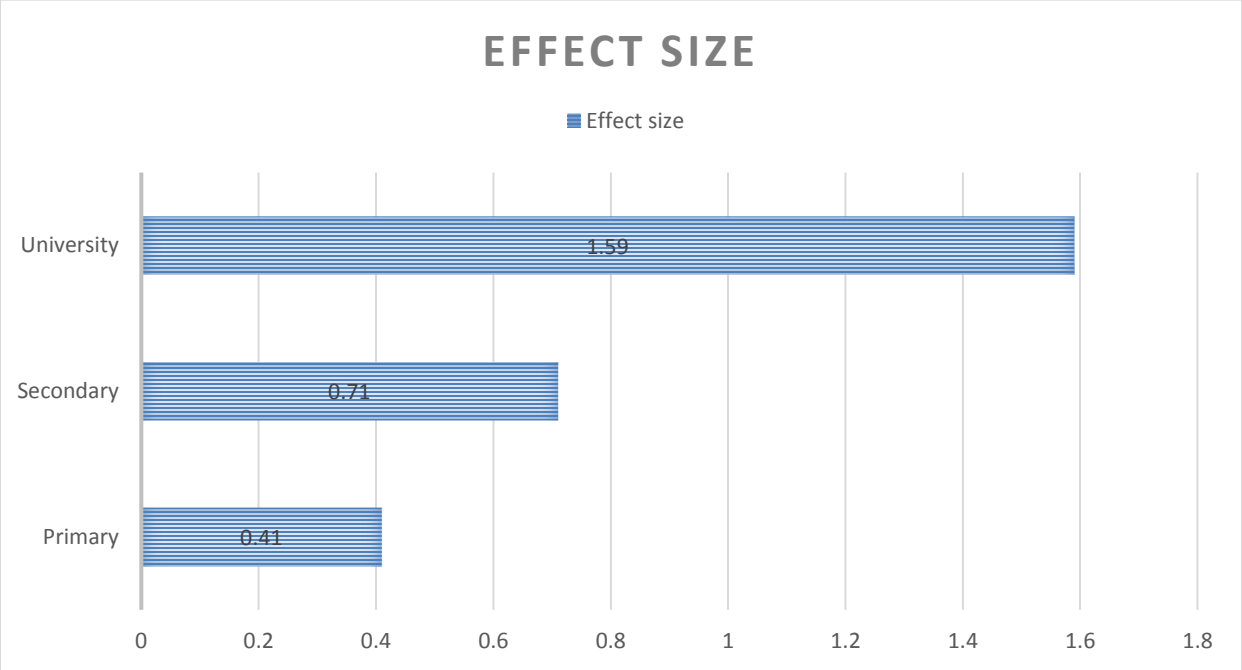

Figure 4.2. Comparison of effect sizes between different educational levels

4.2.2. Skills

Turning to the effects of integrated skills approach on different language 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 and the findings indicate that ISA is much more effective for some certain skills than others. According to presented result in Table 4.2 average effect size of these studies was calculated 2.43, 1.58, 1.19, 1.26, 0.40, 0.02, 0.80, and 0.85 respectively. Results indicate a medium to large effects were obtained for treatment groups over comparison groups in speaking, reading,

writing, listening, general proficiency, and communicative competence. A negligible effects were obtained for vocabulary and linguistic complexity. The maximum effect size was for listening skill (g=2.43, SE=0.58, CI_{95}=1.28, 3.57) and the minimum effect size was for vocabulary (g=0.02, SE=0.71, CI_{95}= -1.36, 1.42).

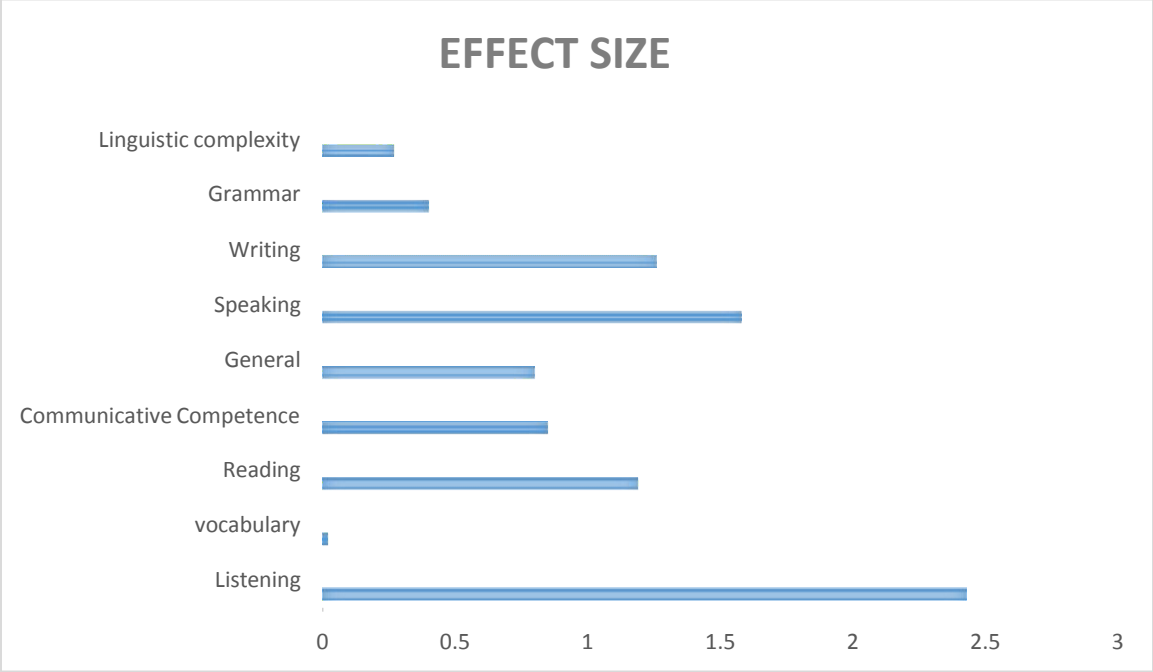

Figure 4.3. Comparison of effect sizes between different language skills and components

4.2.3. Publication Type

Table 4.2 shows the distribution of effect size based on the publication type in the study. To determine the effect of Integrated Skills Approach on students' proficiency, the studies were classified under two groups according to their type as articles and master's thesis. According to the result of the analysis presented in Table 4.2, the average effect size of practices in articles and master's thesis was calculated as 1.17 and 1.59 respectively. According to the result, we find a larger effect on the thesis rather than articles.

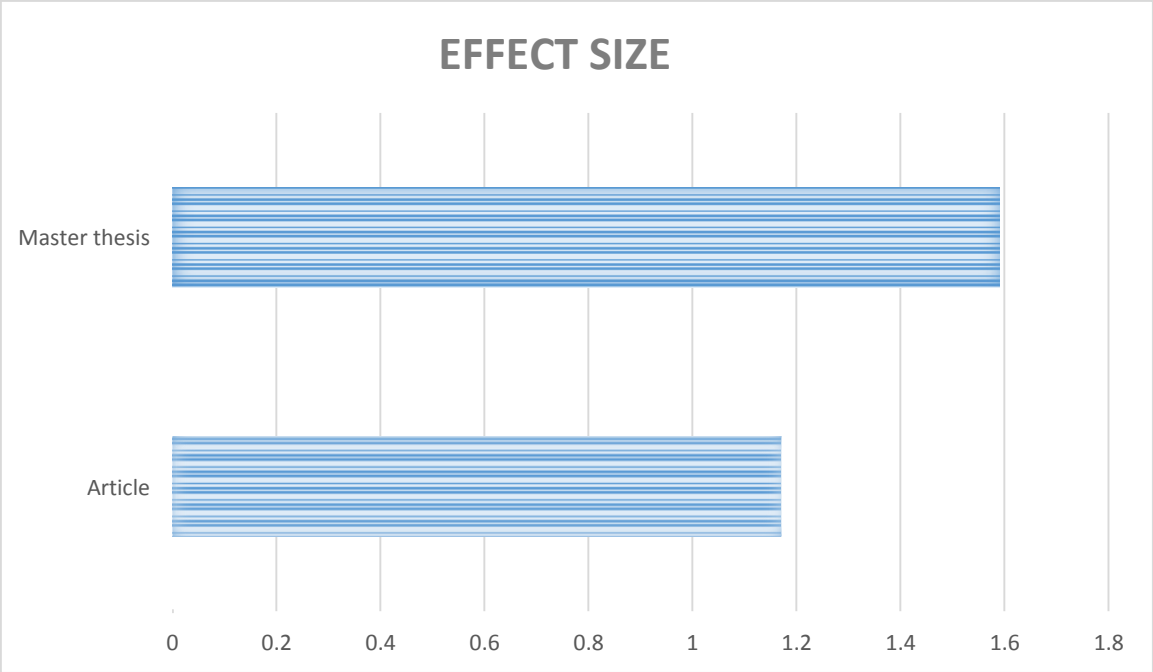

Figure 4.4 Comparison of effect sizes between different publication types

4.2.4. L2 Proficiency level

In order to investigate the potential moderating effect of L2 proficiency level on the effect of ISA on students' overall proficiency, the studies in this meta-analysis are classified according to their sample proficiency level. Studies were classified into three groups including Beginner, Intermediate, and Advanced. And as the results of table 4.2 indicate, the average effect size of these groups was calculated as 0.07, 1.27, and 1.43 respectively. The learners in the advanced and intermediate levels had the large effect size and students in the beginner level had the small effect size (g=0.07).

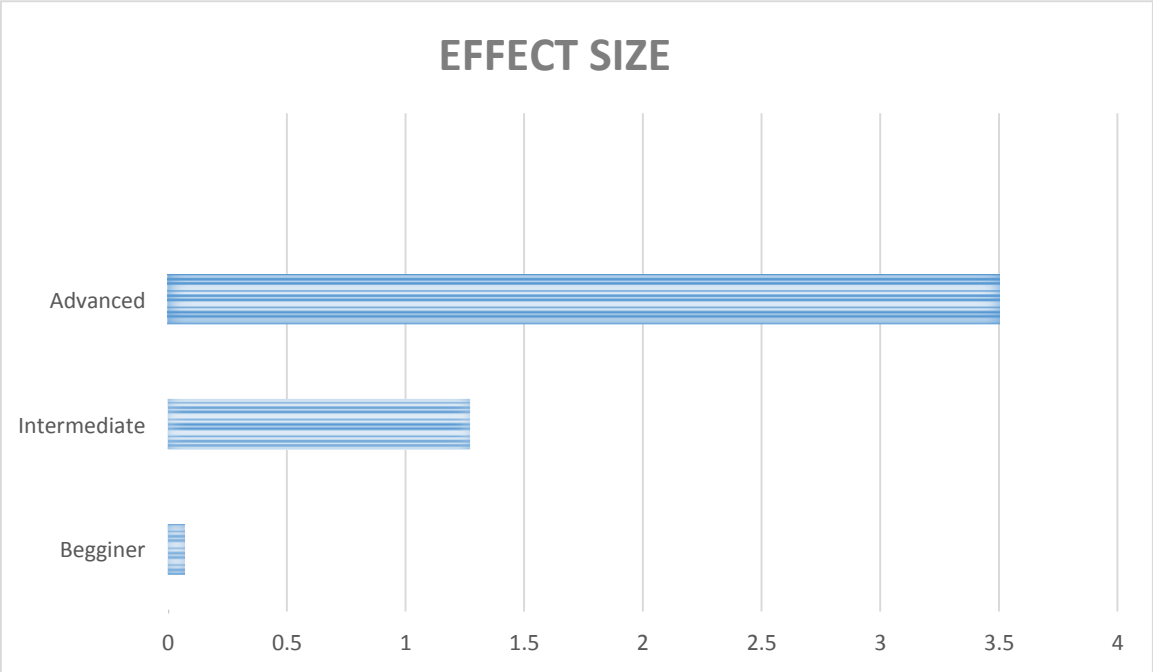

Figure 4.6. Comparison of effect sizes between different proficiency levels

4.3. Evaluation of Publication Bias

Research Question 3: Is there any publication bias in this meta-analysis? If so, what is the level of bias in this study?

One of the features of meta-analysis is that through it we can determine whether there is publication bias in the literature of a specific subject or not and if so what is the level of publication bias? And does it influence the result of the analysis? How much of the main effect size is due to publication bias?

4.3.1. Funnel Scatter Plot

The first diagnostic tool to detect the existence of publication bias is the funnel plot which is a graphical tool. As we described in the Methods, a funnel plot is used to inspect the data for the presence of publication bias over the effect of integrated skills approach. As it is illustrated in figure 4.6, the X-axis shows effect size value (Hedges's g) and the Y-axis shows standard error value in funnel scatter plot. At the bottom of the figure, are studies with smaller samples and larger sampling error. Whereas, studies with greater sample size and higher precision are spread at the

top of the figure around the mean effect size. If Funnel Plot is symmetric it shows there is no publication bias and our meta-analysis had captured all the relevant studies. And if it is asymmetric about the summary effect, it shows the existence of publication bias. In other words, if the density of distribution around the main effect size is high and it is almost symmetric we conclude there is no publication bias in the studies and there is no lost study. As presented in figure 4.6 the funnel plot is almost asymmetric and there is variability in effect sizes, with the greater spread in the middle of the figure an; moreover, these effects are not distributed equally on both sides of the mean effect. Therefore, it presents little evidence of potential publication bias.

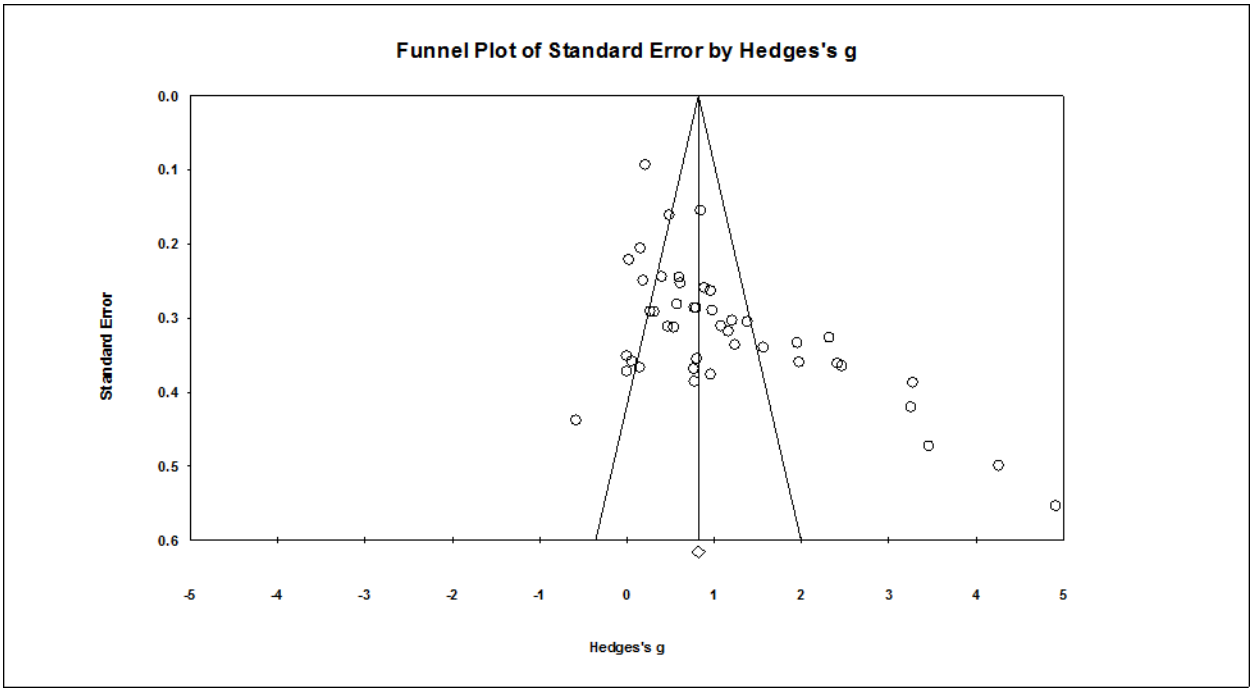

Figure 4.6. Funnel plot of observed studies

4.3.2. Trimm and Fill Method

We used Trimm and Fill method to estimate and adjust for the number of missing studies (due to publication bias) in a meta-analysis. As it is presented in the graph of Duval and Tweedie (figure 4.7), by adding 14 studies at the left side of the graph, it will be a fully symmetric structure and totally unbiased. The result shows 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 classification, both values show a large effect and there is not a big difference between observed effect size in publication-biased case and adjusted effect size in a totally unbiased case. Therefore, in case of publication bias, the studies which were not obtained will not change the calculated effect size for success.

Table 4.3. The result of Trimm 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

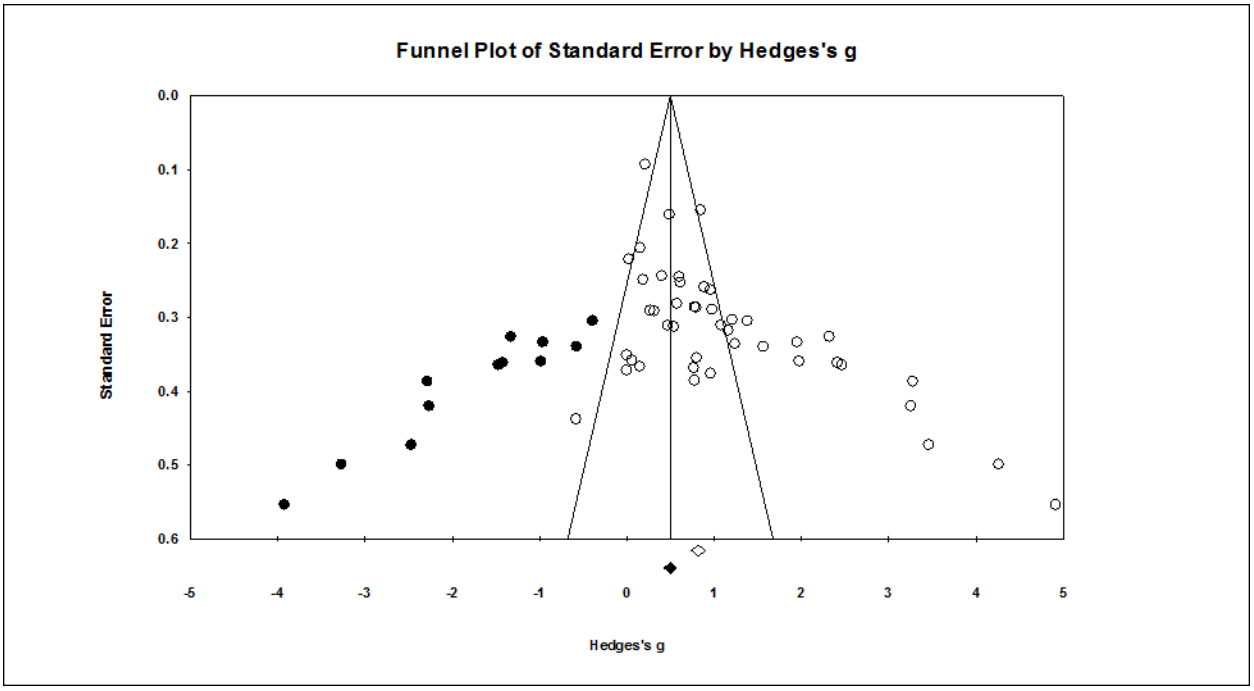

Figure 4.7. Funnel plot on observed and imputed studies

4.3.3. Fail-Safe N

Classic fail-safe N was the last method we used to evaluate publication bias in this study. This index was used to estimate the publication bias for the 46 selected studies in this research. As presented by the data in Table 4.4, the classic fail-safe N test determined that a total of 6839 studies with null results would be needed in order to nullify the effect size. 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.

Table 4.4. 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

CHAPTER FIVE:

DISCUSSION AND CONCLUSION

5.1. Overview

This chapter presents a summary of the findings of the study, outline a discussion of the findings concerning the previous studies in the field, provide the pedagogical implications of the study, and the suggestions for further study.

5.2. Discussion and Conclusion

Different studies have different conclusions about the effectiveness of integrated skills approach which is teaching language in combination with different language skills and components. One group of studies demonstrated the effectiveness of ISA, whereas other groups have not supported such a relationship. Consequently, this leads to disagreement on the positive effect of ISA. This meta-analysis examined the effects of integrated skills approach on learners' overall proficiency and effects of moderators on the effectiveness of this approach. This research carried out a meta-analysis of the effects of integrated skills approach on learners' language proficiency in which 22 studies on integrated skills approach and 46 effect sizes were reviewed and synthesized in which 1870 participants participated in all studies. As Robey (1998) states, "the products of meta-analysis, the average effects size and its confidence interval, estimate the degree to which particular null hypothesis is false on the basis of all available evidence" (p.173). Therefore, based on our findings the null hypothesis is rejected ($g=1.18$, $p=0$). Our results indicated that learners' proficiency benefited strongly from these interventions and a consistent pattern of the positive effect is found. This provided support for the first perspective and showed overall language proficiency benefited strongly from these interventions. This empirical finding supports Oxford's (2001) argument that integrated language teaching is supposed to be an effective strategy for language learning as a whole.

Findings indicated that learners' language proficiency benefited strongly from these interventions. 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 either Plonsky and Oswald's (2014) or Cohen's (1977) interpretation scale and a consistent pattern of the positive effect is found. However, according to Cohen (1988) 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 findings of Plonsky's (2011) work aggregated 95 effect sizes from 61 primary studies on strategy instruction (SI) and findings indicated an overall average effect of 0.49 and the result of outcome variables analysis is in favor of this intervention. Larger effects were obtained in strategy use (1.11), reading (0.74), speaking (0.97), and pronunciation (0.70), respectively. In another context, Maeng (2014) synthesized 45 effect sizes from 37 primary studies on strategy instruction and the result indicated an overall effect size of 0.58. Norris and Ortega (2000) also synthesized the effect sizes of 78 studies in the domain of second language instruction and the result indicates an average effect size of 0.96. Moving closer to the focus of this study, it is good to point Nakhaei's (2017) meta-analysis of CLIL efficacy that yielded an overall effect size of 0.81. She found that CLIL is most effective in enhancing listening (0.91) and grammar (1.32). Also, the lower level of education, elementary level, obtained the greater effect sizes.

Nevertheless, the findings of the current study should be interpreted with taking I^2 in to consideration. This value determines whether moderator analysis is needed or not. Our findings indicated a high degree of heterogeneity (I^2= 90.7) in variances and at the same time, it enabled us to do moderator analysis to find the source of differences among variances because as Borenstein et al (2009) suggested, the degree of heterogeneity determines whether we need to undertake moderator analysis or not.

Regarding theoretical and practical concerns, the variability in observed effect of ISA was investigated as function of four categories of potential moderator variables including educational level, publication type, L2 proficiency level, and skill. So, with respect to subgroup analyses and educational category, there was a significant difference among the results of the studies which were conducted in elementary, secondary, and university setting. ISA had an increased effect on learners at the university level (1.99) over the secondary and primary levels. This approach provides a small effect for the primary level (0.41), a medium effect for the secondary level (0.71), and a strong effect for the university level (1.59). The result has shown that with an increase in educational level, the effectiveness of ISA on learners' proficiency will also increase. 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 and it 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 Integrated Skills approach. Therefore, the implementation of this approach at the university level is prior to the lower levels and factors other than ISA should be considered for implementation at the elementary level. Also, studies on integrated skills have been conducted for different purposes such as MA thesis and research papers. Analysis of publication type as moderator variable indicated this approach had a large effect on both of these groups and they were not significantly different.

It is also worth comparing the differences between various levels of L2 proficiency level. Among L2 proficiency level, advanced level learners (g=1.43, p=0) and intermediate learners (g=1.27, p=0) made great gain and beginners (g=0.07, p=0.90) didn't make a significant gain. As discussed in the literature review, in some previous studies it is proved that advanced and intermediate learners had significant improvement in integrated reading and writing, however begging learners has not shown such improvement (Cho & Brutt-Griffler, 2015).

Among different skills, findings have shown, ISA had a large effect on listening (2.43), speaking (1.58), reading (1.19), and writing (1.26). These findings imply that language main skills are in favor of this intervention and their corresponding p values totally confirm this findings. On the other hand, considering the corresponding p values the effect of ISA on language sub-skills include vocabulary, grammar was not significant. Comparing these findings with Nakhaei's (2017) meta-analysis, integrated skills approach is more effective in enhancing the mentioned main skills. She found large to medium effect for grammar (1.32), listening (0.91), general (0.76), writing (0.81), and reading (0.72). Our findings are in line with Mekheimer's (2016) findings. ISA had a medium effect on communicative competence (0.85) and general proficiency (0.80) and a small effect on grammar (0.40). Its effect on vocabulary and linguistic complexity is not reliable and significant and these results support File and Adam's (2010) finding.

In terms of validity of the study, we used different tests, Funnel Scatter Plot, Trimm and Fill Method, and Fail-Safe N test. As a result, a low level of publication bias was detected among the primary studies. It means, there are few unpublished lost studies that may change the calculated main effect size. Therefore, it shows a high validity of this meta-analysis.

5.3. Implications for Practice

Finding a conclusive result in one hand support and validate the hypotheses behind the integrated skills approach and on the other hand, validate the effectiveness of the integrated skills approach. Obviously, the findings of this study yielded several implications for:

•EFL/ESL teachers, thus they can have better insight on this approach and will help them to implement and execute this way of teaching more effectively in their classroom. They can use this approach in their instructional courses. It helps them to organize an effective language teaching-learning environment. It also helps them diagnose the weaknesses and strengths of the integrated skills approach.

•Policy makers, because they can make an informed decision based on the finding of this study and using ISA as an approach to teach EFL/ESL in university and secondary school should be encouraged. Because the result of this study is more comprehensive and conclusive than individual studies. It is due to the great number of samples.

•The curriculum developers of second language institutions can design their educational programs based on the weakness and strengths of the integrated skills approach. They can better decide to employ this way of teaching in the instructional programs to appropriate levels.

• Researchers reach a conclusive result on the effectiveness of integrated skills approach. They don’t need to review the previous research to find the overall effect of this approach.

5.4. Suggestion for Further Research

This meta-analysis has attempted to provide the overall combined effect size of integrated skills approach coupled with its differential effect at different language levels, educational levels, publication types, and different language skills. Like any other secondary review, the completeness and precision of the present meta-analysis are contingent on the precision of the relevant primary studies. Therefore, this study will end by offering some suggestions for further study in ISA area to fill the gaps that exist in the literature review.

Although there is a rich literature in ISA area, a large number of studies are excluded from this meta-analysis because they did not report the statistics needed to compute the overall effect size. It is clear that our increasing knowledge of the effects of ISA depends on reported data and basic descriptive statistics needed to aggregate and compare findings of different studies, such as standard deviations, sample sizes, means, effect sizes, and measurement reliability, confidence interval. So, it is suggested researchers who are interested in ISA area present the statistical data to enable researchers to analyze the overall effect size.

As the forest plot indicated, the precision and weight of each study is demonstrated by the size of the boxes situated in line with effect sizes. Also, the lines which go through squares depicts the length of confidence interval and the loner the line, the less precise is the study. The length of the lines in this forest plot indicates that some of the primary studies included in this meta-analysis are not precise enough or their sample size is small. Therefore, it is suggested the researchers in ISA area assure the methodological rigor of their study and use the larger sample size to reach more precise results.

Moreover, rigorously reporting contextual data such as type of contexts (foreign language or second language), gender will enable researchers to do subgroup analysis more precisely, and also consumers of primary studies can contextualize and interpret findings more accurately. It is also suggested, researchers interested in the effectiveness of integrated skills approach present their treatment procedure, the length of treatment, and the instrument with more rigorous attention

to detail to enable researchers to compare with other studies or replicate the study; because, these factors have a prominent role in validity of the study (Plonsky, 2011).

Despite the rich literature on integrated skills approach testing its effect on different skills, there is a paucity of experimental studies testing the effectiveness of ISA on students' vocabulary, communicative competence, and grammar thus, more experimental and quasi-experimental research studies are needed in these domains for further investigations.

Finally, the value of integrated skills approach depends greatly on lasting its effect over time; however, only a few relevant studies to this meta-analysis used delayed posttests. Thus, additional measurements are needed to evaluate the persisting effects of ISA.

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 taskcomplexity. *System, 81*(2019), 110-121.

*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.

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

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.

Allen, T. C. (1987). *Using schema theory to integrate reading and writing process in composition*. (Master's thesis). Available from Theses Digitization Project database.

Almalki, M. S, & Soomro, A. F. (2017). Practitioners' perspectives on the application of integration theory in the Saudi EFL context. *English Language Teaching, 10*(6), 93-102.

*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.

Borensteina,M., Hedges, L.V., Higginsc, J.P.T., & Rothsteind, H.R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Wiley Online Library*, 197-111.

Briere, E. J. (1972). Are we really measuring proficiency with our foreign langauge tests? *Foreign Language Annals, 4*(4), 385-391.

Brown, H. D. (2001). *Teaching by principles: An interactive approach to language pedagogy (*2nd Ed). Whit Plains, NY: Addison Wesley Longman.

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.

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics, 1*(1), 1-47.

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.

Clark, J. L. D. (1972). *Foreign language testing: Theory and practice.* Philadelphia, PA: The Center for Curriculum Development.

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

Cohen, L., Manion,L., & Morrison, K. (2007). *Research methods in education.* New York, NY: Taylor & Francis.

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.

Cook. T. D., Cooper, H., Cordray, D. S., Hartman, H., Hedges, L. V., Light, R. J., Louis, T. A. (1992). *Meta-analysis for Explanation.* New York: Russell Sage Foundation.

Cook, D. J., Guyatt, G. H., Salena, B.J., & Laine, L. A. (1992). Endoscopic therapy for acute

nonvericeal upper gastrointestinal hemorrhage: A meta-analysis. *Gastroenterology, 102*(1), 139-148.

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.

Cooper, H. & Hedges, L. (2009). *The handbook of research synthesis and meta-analysis (2 nd).* New York, NY: Russel Sage Foundation.

Cooper, H. M., & Rosenthal, R. (1980). Statistical versus traditional procedures for summarizing research findings. *Psychological Bulletin, 87*(3), 442-449.

Corder, S. P. (1971). Idiosyncratic dialects and error analysis. *IRAL, 9*(2), 147-160.

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.

Decoster, J. (2009). *Meta-analysis notes.* Alabama: University of Alabama.

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

Duval, S. J. (2009). The trim and fill method. In H. R. Rothstein, A. J. Sutton, & M. Borenstein (Eds.) *Publication bias in meta-analysis: Prevention, assessment, and adjustments* (pp. 127–144). Chichester, England: Wiley.

Ellis, R. (2015). *Introduction: Complementarity in research syntheses.* Oxford: Oxford University Press.

*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.

Fitz-Gibbon, C. T. (1984). Meta-analysis: An explication. *British Educational Research Journal, 10*(2), 135-144.

Gautam, P. (2019). Integrated and segregated teaching of language skills: An exploration. *Journal of NELTA Gandaki, 1*(2676-1041), 100-107.

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *American Educational Research Association, 5*(10), 3-8.

Glass, G. V. (1982). Meta-analysis: An approach to the synthesis of research results. *Journal of Research in Science Teaching, 19*(2), 93-112.

Glass, G. V., Smith, M. L., & McGau, B. (1981). *Meta-analysis in social research.* Thousand Oaks, CA Sage Publication.

Graham, S. & Herbert, M. (2011). Writing to read: A meta-analysis of the impact of writing and writing instruction on reading. *Harvard Educational Review, 81*(4), 710-744.

Harris, D. P. (1968). *Testing English as a Second Language.* New York: McGraw-Hill.

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

Higgins, J. P., & Green, S. (2011). *Cohrance handbook for systematic reviews of interventions.* UK: Willy.

Higgins,s., Hall, E., Baumfield, V., & Moseley, D. (2005). *A meta-analysis of the impact of the implementation of thinking skills approaches on pupils.*London: Social Science Research Unit, Institute of Education, University of London.

Hinkel, 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.

Howatt, A. P. R., & Widdowson, H. G. (2004). *A history of ELT.* Oxford: Oxford University Press.

Hunter, J.E., Schmidt, F. L., & Jackson, G. B. (1982). *Meta-analysis: Cumulating research findings across studies.* Beverly Hills: Sage Publications.

Hymes, D. (1972). On communicative competence . In J. Pride & J. Holmes(Eds.), *Sociolinguistics: Selected readings.* Harmondworth, England: Penguin Books.

Hymes, D. (1979). Language in education: Forward to fundamentals. *Language, Children and Society*, 1-19.

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

Kaplan, R. (1970). *The mind system: A grammar-rule language*. Santa Monica, California: The RAND Corporation.

Krashen, S.D. (1985). *The input hypothesis.* London: Longman.

Lado, R. (1961). *Language testing.* New York, NY: McGraw-Hill.

Lee, J. W., Schallert, D. L. (1997). The relative contribution of L2 language proficiency and L1 leading ability to L2 reading performance: Atest of the threshold hypothesis in an EFL context. *Tesol Quarterly, 31*(4), 713-739.

Lipsey, M., & Wilson, D. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

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

Long, M. H. (1981). Input, interaction, and second language acquisition. *Annals of the New York Academy of Science, 379*(1), 259-278.

Long, M. H. (1996). The role of the linguistic environment in second language acquistion. In w. R. Bhatia (Eds.), *Handbook of second language acquistion.* San Diego: Academic Press.

Maeng, U. (2014). The Effectiveness of reading strategy instruction: A meta-analysis. *English Teaching, 69*(3), 105-127.

*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.

Nakhaei, S. (2017). *Content and language integrated learning: A meta- analys* (Unpublished master's thesis). Shahrood University of Technology: Shahrood.

Norris, J. M. (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.

Oller, J. W., & Hinofotis, F. B. (1980). Two mutually exclusive hypotheses about second language ablility: Factor analytic studies of a variety of language subtests. In J. W. Oller & K. Perkins (Eds.), *Research in language testing* (pp. 13-23). Rowley, MA: Newbury

House.

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

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.

Petticrew, M. & Robert, H. (2008). *Systematic reveiws in the social sciences: A practical guide.* UK: Wiley.

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.

Rhodes, W. (2012). Meta-analysis: An introduction using regression models. *Sage*, 36(1), 24-71.

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.

Rosenthal, R. (1991). *Meta-analytic procedures for social research.* Newbury Park, CA: Sage.

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

*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.

Smith, M. C., & Naftel, D. C. (1984). Meta-analysis: A perspective for research synthesis. *The Journal of Nursing Scholarship, 16*(1), 9-13.

*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.

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.

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. *English Language Teaching, 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.

Thomas, M. (1994). Assessment of L2 proficiency in second language acquisition research. *Language Learning, 44*(2), 307-336.

Uman, L. S. (2011). Systematic reviews and meta-analyses. *Journal of Canadian Academy of Child and Adolescent Psychiatry, 20*(1), 57-59.

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

Yamashita, J. (2002). Mutual compensation between L1 reading ability and L2 language reading ability and L2 language comprehension. *Journal of Research in Reading, 25*(1), 81-95.

چکیده

رویکرد تلفیق مهارتها به عنوان روشی برای آموزش زبان که در آن مهارتهای زبانی در حین آموزش با هم تلفیق و یکپارچه می شوند، به یک روند غالب در آموزش زبان تبدیل شده است.هدف این پژوهش محاسبه اثر کلی رویکرد تلفیق مهارتها و همینطور تشخیص منبع و میزان واریانس(پراکندگی) در تاثیر مشاهده شده است.برای این منظور, مطالعه فراتحلیل تأثیر رویکرد مهارتهای تلفیقی بر مهارت زبانی دانش آموزان انجام شده است که در آن 22 مطالعه تجربی و شبه تجربی از مجلات داوری شده منتشر شده در بازه زمانی 1983-2013 از و 46 اندازه اثر بررسی شدند. نتایج اندازه اثر کلی 1.18 را نشان داد که نشان دهنده اندازه اثر قابل توجهی است. اندازه اثر متغیرهای تعدیل کننده محاسبه و گزارش شد که رویکرد تلفیق مهارتها بیشترین اثر را در زبان اموزان سطح پیشرفته و سطح دانشگاهی دارد. این یافته ها نشان می دهد که اگرچه این روش ممکن است مهارتهای اصلی زبان را بهبود ببخشد ، اما روی واژگان و دستور زبان تأثیر معناداری ندارند. یافته های این فراتحلیل برای معلمان زبان انکلیسی، محققان، سیاستگذاران و برنامه ریزان اموزشی کاربرد دارد.

کلمات کلیدی : رویکرد تلفیق مهارتها، مهارت زبانی ، ترکیب پژوهش ، فراتحلیل ، اندازه اثر.

پایان نامه کارشناسی ارشد آموزش زبان انگلیسی

فراتحلیلی بر رویکرد تلفیق مهارتها

نگارنده

فاطمه احمدنتاج کاسگری

استاد راهنما

دکتر سید علی استوار نامقی

استاد مشاور

دکتر فاطمه مظفری

مهر ۱۳۹۹