The Effectiveness of English Language Teaching Methods: A Meta-Analysis in Pursuit for "the Best Method"

Esmaeel Saeedy Robat *   **Hossein Khodabakhshzadeh** **

Hamid Ashraf ***
Department of English, Torbat -e Heydarieh Branch, Islamic Azad University, Torbat -e Heydarieh, Iran
Majid Elahi Shirvan ****

Department of English, University of Bojnord

Abstract

What is "the best method" of language teaching? It has been one of the oldest questions of language teaching and learning during the last century. However, no comprehensive quantitative study has tackled the issue. In order to answer this question, the researchers meta-analyzed 56 studies with 7960 participants from many contexts. A coding scheme of 46 variables, in the form of four major moderator sets, including design characteristics, language characteristics, participant characteristics, and teaching characteristics, was developed. The overall effect size (g= 1.00) was found to be positive, strong, and significant for all language teaching

methods. The findings showed that all language teaching methods, irrespective of various contexts, were positively effective. The results of moderator analysis showed that most of the moderators, excluding language skills and place of the study, had no significant effect on language teaching methods. Implications for current theory and practice, for both method and postmethod, are discussed.

Keywords: Language Teaching Methods, The Best Method, Postmethod, Meta-Analysis

Research in language teaching and learning, like many other fields in social sciences and humanities, has proved to produce different results and findings. Many of these inconsistencies originate from a bunch of various factors, including research theories, data collection tools and instruments, researchers' background and experiences, data analysis, participants' classifications, and characteristics, and a plethora of other issues.

When we started searching the literature for the study for the first time in 2018, we asked the Google search engine, "What is the best method of language teaching, and why?" The results were almost countless, and it was not surprising, for it may happen to any other question. However, we found the same question that was posted on ResearchGate, a website, specially designed for verified researchers, and there were over 20 different answers. One answer was, "I think "Desuggestopedia" is an optimum method," from a university teacher from Malaysia. This diverse pool of data has both blinded the trend of the research in the past and obscured the path of future attempts. Robert and Lawrence (2006) pointed out that "no one seems to know what works."(p. 3). Concerning English language teaching methods, the numbers of studies are countless. We could dare to say that no other field of ELT has ever been under scrutiny and probing like teaching methods. Each one of these methods claims that it is best both in theory and practice.

The terms regarding teaching methods are mainly concerned with the how of language teaching profession and industry, and academic and scholarly endeavors. Dueñas Vinuesa (2002) believes that all methods include prescriptions for the teachers and the learners. That is, usually, all methods are pre-packaged sets of specifications of how the teacher should teach and how the learner should learn, derived from a particular theory of language and a theory of language learning. She continues that for the teacher, methods prescribe what materials and activities should be used, how they should be used, and what the role of the teacher should be. For learners, methods prescribe what approach to learning the learner should take and what roles the learner should adopt in the classroom. The amount of data and information on methods, styles and strategies in ELT are sometimes obscure and overwhelming as Robert and Lawrence (2006) say "no one seems to agree with anyone else's approach (p. 3).

Review of Literature

The sheer number of concepts, theoretical underpinnings, road maps, classroom practices, pedagogical approaches, methods, styles, strategies, techniques, and many other terminologies that have flourished in ELT in the last 70 years, is overwhelming. According to Cruz-Arcila (2013, p.82), "on many occasions, they are the modifications, complement or opposition to others." The point was also reiterated by Kumaravadivelu (2006) as a superficial view of the same phenomenon with many fundamentals the same. Both in the method and post method era, methods are alive and forceful, though with different mentalities and practices. Post-method is not a method itself as long as it has not restricted itself to the same limited framework of methods. We should argue that post-method pedagogy has provided teachers with the freedom of creating and crafting their methods, their styles,

techniques, etc. However, the ashes of method era have haunted many teachers and practitioners until today; language teaching shareholders have tried hard to put the heavy burden of the method down. With the Stern et al. (1983) declaration of method independence, we expected a real shift from the method era to the post method age. In most cases, some old concepts and transcriptions were replaced with fresh ones. Theory-driven, close-ended and prescriptive notions such as Audio-lingualism, natural approach, communicative language teaching, silent method, and community language learning, to name but a few, were substituted with theory-neutral, open-ended and descriptive concepts proposed by Kumaravadivelu (2006) like macro-strategic frameworks and the three-dimensional model by Stern (1992), and many other concepts. Waters (2009) argues that:

> The island of methodologia lies at the heart of the ELT world. It is here that classroom teaching skills are devised, tested, and popularized. However, many ELT practitioners are familiar only with the parts of the island where they grew up, even though there is much to be gained from an appreciation of the culture and history of methodologia as a whole. (p.1)

So long as the concept of the method is alive, though, with various names and alternatives, the question of more effective kinds of instruction in ELT is valid and worth responding.

Based on the review of the literature on different methods, styles, strategies, and other alternative or similar terminologies and concepts, we divided these terminologies and concepts into two major categories; the concepts of the method age and the concepts of the post method age. The former started from the grammar-translation method (GTM) to communicative language teaching (CLT) and task-based language teaching

(TBLT). The GTM was considered to be a theory-less method, and CLT and TBLT were considered to be theory-rich concepts in ELT literature. The latter started in the late 1980s and early 1990s movements of post-method activists like Kumaravadivelu and Stern and later by many others; it emerged as a dominant trend and force in ELT, which claimed to be theory-neutral Kumaravadivelu (1994). Bell (2003) argues that "contrary to this claim, some consider the term method to remain an apt description of what teachers do in classrooms" (p.325). Methods may be dead at the etic level (at scholars' level) but will remain as a quite valid and significant holding ground at the epic level (part of teachers and laypeople mentality) argued by Block (2001). The same point also was raised by Ellis (2010). It was pointed out that there is a gap between theory/research and language pedagogy. Teachers are out there teaching for themselves based on their cognition and many other concerns, as mentioned by Borg (2003); while researchers and theoreticians are also hard at work for their own sake. Each one has his or her own path, end and concern.

In a nutshell, both in the method and post method eras, the concept of the method is alive, though it may have gone through some metamorphic transformations. According to Block (2001), despite its shortcomings, the perception of methods as a prescription is still a salient one for classroom teachers. In other words, many teachers and students are always looking for the best or better methods, and the quest is on, just like before. For that reason, we have tried to answer the same old question of a more effective method or methods through conducting a thorough meta-analysis of language teaching methods. Several meta-analysis studies have been conducted in ELT concerning the effectiveness of instruction in the last two decades (see Table 1). Most of them have focused on a narrow scope of interventions in ELT instruction such as vocabulary learning, TBLT effectiveness, pronunciation,

	Journal of Teaching Language Skills **(JTLS)** 40(1), Winter 2021, pp. 187-235	**192** Esmaeel Saeedy Robat

THE EFFECTIVENESS OF ENGLISH LANGUAGE TEACHING METHODS

and grammar. Some recent studies have focused on instruction intervention, reported in the ELT literature (Table 1).

As depicted in Table 1, vocabulary learning and reading comprehension have received most of the attention. Concerning traditional methods, in the post-method sense, TBLT is probably the only teaching method that was investigated through a meta-analysis study. According to Bryfonski and McKay (2017), TBLT reported a high positive effect size (d=0.93) in many learning situations. Since this study included some moderators, such as type of the research design, time, place, and proficiency measures, in the final analysis, the findings were rather comprehensive.

Table 1.

Recent Meta-Analysis Studies in ELT

No	Title of the Paper	Author (s)	Focus the study
1	L2 reading comprehension and its correlates: a meta-analysis	(Hee & Junko, 2014)	Reading comprehension
2	The effectiveness of second language pronunciation instruction: a meta-analysis	(Lee et al., 2015)	Pronunciation
3	A meta-analysis of vocabulary learning strategies of EFL learners	(Nematollahi et al., 2017)	Vocabulary learning
4	The effects of corpus use on second language vocabulary learning: a multilevel meta-analysis	(Hansol et al., 2018)	Vocabulary learning
5	The effectiveness of processing instruction and production-based instruction on l2 grammar acquisition: a meta-analysis	(Shintani, 2015)	Processing instruction
6	Computer-mediated glosses in second language reading comprehension and vocabulary learning: A meta-analysis	(Abraham, 2008)	Reading comprehension

No	Title of the Paper	Author (s)	Focus the study
7	A synthesis of research on language of reading instruction for English language learners	(Slavin & Cheung, 2005)	Reading comprehension
8	A systematic review of the impact of summative assessment and tests on students' motivation for learning. Research Evidence in Education Library	(Harlen & Crick, 2002)	Motivation and assessment
9	TBLT implementation and evaluation: A meta-analysis	(Bryfonski & McKay, 2017)	TBLT
10	A meta-analysis and meta-regression of incidental second language word learning from spoken input	(De Vos et al., 2018)	Vocabulary learning
11	The associations between language aptitude and second language grammar acquisition: a meta-analytic review of five decades of research	(Li, 2015)	Grammar acquisition
12	The effects of task involvement load on l2 incidental vocabulary learning: a meta analytic study	(Huang et al., 2012)	Vocabulary learning
13	The effects of extensive reading on English vocabulary learning: a meta-analysis	(Liu & Zhang, 2018)	Vocabulary learning
14	Shared book reading interventions with English learners: a meta-analysis	(Fitton et al., 2018)	Reading comprehension

They found differences in effect sizes in various regions that implemented TBLT, from the Middle East (d=1.31), the highest, to East Asia and Europe (0.33) the lowest. The differences, on the one hand, support the applicability of TBLT in various socio-cultural settings, and on the other hand, pinpoint some incongruities in learning and teaching.

A similar meta-analysis study carried out by Li (2015) on the role of language aptitude in second language grammar acquisition. He included a

wide range of variables, including explicit and implicit instructions, age, language analytic ability, and memory, in the study to provide a full picture of grammar acquisition moderators. He pointed out:

"…. the importance of aptitude has been somewhat exaggerated, that it is predictive of initial L2 grammatical competence and less so of later stages of learning, and that it is a conscious construct that affects learning outcome in explicit conditions" (p.407)

The Rationale for Moderator Variables

For moderator selection, the researchers had two purposes in mind; first, we aimed to include as many variables as possible; second, we intended to present the variables as organized as possible. Since language teaching and learning exchange is a multifarious and complex web of many interwoven factors, and language knowledge itself is also multifaceted, as Nassaji (2020) pointed out, any single measure would only provide a partial picture of the nature of the issue. For the two mentioned purposes, we explored the theoretical dimensions of language teaching methods through review studies and carried out a preliminary literature search to identify the variables frequently reported in language teaching intervention studies. Finally, four moderator sets, design characteristics, language characteristics, participants' characteristics, and teaching characteristics were included in the final analysis.

The Rationale for Using a Meta-Analytical Approach

One problem in the area of language teaching is the overwhelming diversity of teaching methods. For that reason, a meta-analysis, with statistical and analytical power higher power than single studies as articulated by Blokdyk (2020), could provide us with a full picture of the language teaching methods. On the one hand, a meta-analytic approach would allow us to estimate the overall effect of language teaching methods with more precision,

but on the other hand, with moderator analysis, alternative covariates and potential intervening factors would be identified.

The following research questions guided the present meta-analysis:

Research Question1: What is the overall effectiveness of the English language teaching method(s) on L2 achievement?

Research Question 2: Are the effects of the English language teaching methods moderated by the features of intervention programs?

Method

Study Identification and Retrieval

To develop a comprehensive pooled data, based on the PRISMA (preferred reporting items for systematic reviews and meta-analyses) Moher et al. (2009) as seen in Figure 1, we thoroughly searched the electronic databases with the following search terms and phrases both separately and in combinations:

Language teaching methods, language teaching strategies, language teaching styles, English language teaching, reading, writing, speaking, listening, vocabulary teaching, vocabulary learning, strategic teaching, English learning and teaching, grammar teaching and learning, lexical teaching and learning, English lexicon, language teaching skills and, learning English, etc. We should also need to mention that we used advanced search tools and other academic search technologies, such as date intervals, the combination of terms, and subject and discipline categorization that electronic databases have provided.

Journal of Teaching Language Skills (JTLS) 40(1), Winter 2021, pp. 187-235	**196** Esmaeel Saeedy Robat

THE EFFECTIVENESS OF ENGLISH LANGUAGE TEACHING METHODS

Figure 1.

Study Retrieval Process
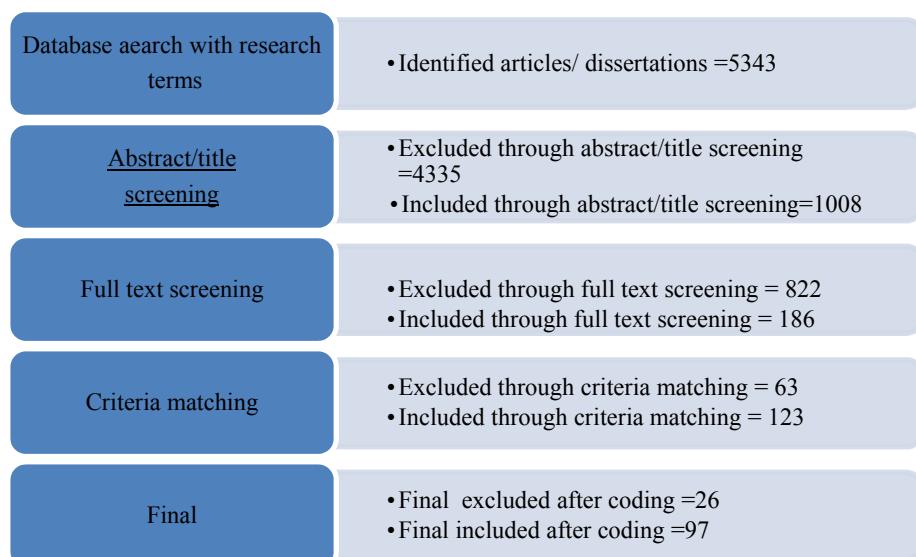
After identifying databases related to the current meta-analysis, we searched the following databases: ERIC - Education Resources Information Center, ScienceDirect, Sage Journals Online, SAGE Knowledge, ProQuest, Cambridge Core, Google Scholar, JSTOR, ProQuest Dissertations & Theses Global, PsycARTICLES, PsycINFO, SAGE Research Methods Online, Microsoft Academic Search, Academic Search Complete (EBSCO), Linguistics and Language Behaviour Abstracts (LLBA) (ProQuest), Project MUSE, Blackwell Reference Online, Scopus, Web of Science, Academic Search Premier, SpringerLink, Wiley Online Library, ResearchGate, iSEEK Education, RefSeek, Virtual LRC, Academic Index, Internet Public Library,

![JTLS]	Journal of Teaching Language Skills (JTLS) 40(1), Winter 2021, pp. 187-235	**197** **Esmaeel Saeedy Robat**

THE EFFECTIVENESS OF ENGLISH LANGUAGE TEACHING METHODS

Oxford Handbooks Online, Oxford Journals Digital Archive, and many other resources.

In the next stage of the study, the researchers identified major academic journals and publications in language, linguistics, and education studies. In order to avoid the inclusion of predatory journals or publishers in our analysis, we used scientific journal metrics and rankings such as impact factors and other assessments of scholarly publications. The researchers handpicked these academic journals, and their published studies were meticulously scrutinized one by one from the first issue to the latest ones as follows: Language Teaching Research from 1997, TESOL Quarterly from 1960, The Modern Language Journal from 1960, ELT Journal from 1960, Second Language Research from1985,Annual Review of Applied Linguistics from 1980, Language Teaching from 1960, Applied Linguistics from 1980,Studies in Second Language Acquisition from 1970, Language learning from 1960,Foreign Language Annals from 1967, System from 1973, Research in the Teaching of English from 1960, Journal of Second Language Writing from 1992, Journal of English for Academic Purposes from 2002, Language Acquisition from 1990, Innovation in Language Learning and Teaching from 2007, The Language Learning Journal from 1990, Asian-Pacific Journal of Second and Foreign Language Education from 2016, English Teaching: Practice & Critique from 2015, English Language Teaching from 2008, Language Awareness from 1992, Applied Psycholinguistics from 1980, In the next phase of the study, we checked ProQuest database to find unpublished doctoral dissertations. In the final part of the data retrieval and identification, we investigated the reference sections of the relevant papers and continued chasing for more studies on the topic. The study selection process is depicted in Figure 1.

Journal of Teaching Language Skills (JTLS) 40(1), Winter 2021, pp. 187-235	**198** Esmaeel Saeedy Robat

THE EFFECTIVENESS OF ENGLISH LANGUAGE TEACHING METHODS

Inclusion and Exclusion Criteria

A study was included in this meta-analysis if it met all the criteria listed below.

1. The study investigated English language teaching methods and their effect on language learning.

2. The studies were reported in a refereed, peer-reviewed journal, thesis, or dissertation in English between 1970-2020.

3. The study used an experimental or quasi-experimental design.

4. The study contained required quantitative and statistical data to perform the meta-analysis.

5. The independent variable was the treatment or intervention through the English language teaching method(s).

6. The dependent variable or variables are one or all of the following:

A. Learners scores through standardized tests and/or any other established, well-documented assessment procedures.

B. Proficiency levels in English language skills (listening, reading, writing, and speaking) discretely of holistically.

C. Proficiency levels in English language components like vocabulary, grammar, and pronunciation.

7. The participants were children, young or adult learners of English as a second or foreign language.

A (part of a) study was not included if at least one of the exclusion criteria was met:

1. The study did not report enough statistical data to perform the analysis.

2. The study was conducted before 1960.

3. The participants had disabilities like hearing-impaired persons or any other forms of disabilities.

Journal of Teaching Language Skills **(JTLS)** 40(1), Winter 2021, pp. 187-235	**199** **Esmaeel Saeedy Robat**

THE EFFECTIVENESS OF ENGLISH LANGUAGE TEACHING METHODS

Effect Size Calculation

For the main effect sizes used in this study, we employed a standard mean difference, which is a weighted estimate of the difference between treatment and comparison groups as mentioned by Cooper et al. (2009). We calculated the effect sizes as Hedge's g, which is the effect size measure that represents the standardized difference between means and is considered less biased than Cohen's d. The two statistics are similar except, in the case where the sample sizes are below 20, Hedge's g is preferable according to Durlak (2009) . The positive effect size in this study suggests that language learners in a particular setting (i.e., experimental groups) outperformed language learners in another context (i.e., control groups). For the calculation of effect sizes, we used adjusted post-test means that accounted for potential pre-treatment differences. We also calculated effect sizes based on a pre-test post-test control group design using Comprehensive Meta-Analysis software (CMA, version 3; ©2014, Biostat, Inc., Englewood, NJ). For studies that reported using multiple measurement instruments for a certain variable, such as multiple choice and cloze tests, the average score from the variable was calculated.

Coding of Moderator Variables

Besides the computation of mean effect sizes, we included a series of moderator analyses. Since moderating variables are factors that may affect the general effect size estimation through covariation with the key independent variables, we added four major moderators in the analysis in the form of four major sets or clusters. Figure 2. depicted moderators' major sets.

	Journal of Teaching Language Skills (JTLS)	**200**
JTS	40(1), Winter 2021, pp. 187-235	**Esmaeel Saeedy Robat**

THE EFFECTIVENESS OF ENGLISH LANGUAGE TEACHING METHODS

Figure 2.
Moderators' Major Sets.
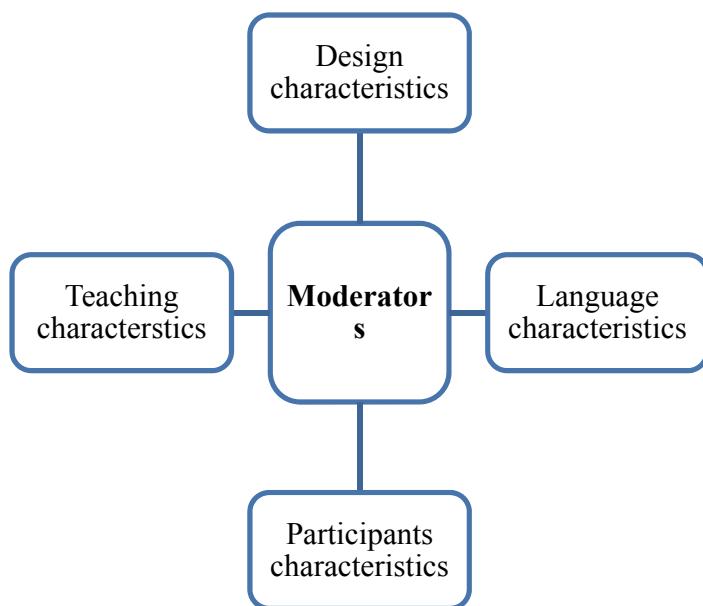
Under these sets of moderators, we included 46 different covariates, which might influence outcomes. A separate meta-regression was conducted for each set of moderators and their subset, including design characteristics, language characteristics, participants' characteristics, and teaching characteristics. For detailed information, the coding scheme is presented in Table 2.

Table 2.

Moderator' Coding Scheme

Variables	Value	Definition
Author(s) name	Name of the author(s)	The name of the reported author(s) is mentioned
Publication date	Publication year	Reported year of publication
Place of the study	Middle east = 1 Asia = 2 North and south America = 3 Europe and Oceania = 4 Mixed = 5	Reported place of the study
Study type	Journal Article =1 Unpublished Ph.D.=2	The study was a journal article or an unpublished Ph.D. dissertation
Study ID code	001- 00…	A unique number assigned to each student
Native language	Persian = 1 Arabic = 2 Chinese and Taiwanese = 3 Japanese and Korean = 4 European languages = 5 Mixed and other languages = 6	The reported native language of the majority of participants
Language major focus	Reading = 1 Writing = 2 Speaking = 3 Listening = 4 Grammar = 5 Vocabulary = 6 General = 7 Pragmatics=8	Self-explanatory
L2 proficiency level	Beginner to low intermediate =1 Intermediate =2 High intermediate to advance=3 Not reported = 4 Mixed =5	Reported variables representing the L2 proficiency levels of participants
	Elementary school = 1	

	Journal of Teaching Language Skills (JTLS) 40(1), Winter 2021, pp. 187-235	**202** Esmaeel Saeedy Robat

THE EFFECTIVENESS OF ENGLISH LANGUAGE TEACHING METHODS

Variables	Value	Definition
Instructional level	Junior high school = 2 Senior high school = 3 University = 4 Not reported =5 Mixed = 6	The variables represented the instructional level of the participants
Study design	Experimental = 1 Quasi-experimental = 2	The reported design of the study
Sample size	The number of participants	The reported number of participants
Proficiency measurement	Standardized test = 1 Institutional test =2 Researcher's judgment=3 Self-designed = 4 Not reported =5	Proficiency measurement instruments

Missing Data

During the coding process, we detected examples of missing data in some studies. In some instances, we contacted the corresponding author or primary investigator through email for missing data to compute the effect sizes required. In some other studies, we tried to calculate the effect size via raw data. However, this didn't yield more studies to include in our analysis.

Reliability

Coding reliability was measured through the assessment of inter-coder reliability. Following the development of the coding scheme, we met with two other independent raters. The initial discussions led to independent coding of 50 percent of the studies. The primary inter-rater reliability of 91% was calculated based on each study's features to be included in the analysis. The team met again in cases of discrepancies and tried to make the inclusion criteria as transparent and explicit as possible, and a consensus was established on final reliability of 94%.

Publication Bias

In order to avoid or decrease publication bias in meta-analysis studies, three distinct solutions are suggested. First, funnel plots are visual tools that provide us with some sense of data distribution and help us evade overestimated or non-existing effects. Symmetric distribution about the mean effect size indicates the absence of publication bias. Second, the Trim and fill method developed by Sue Duval and Richard Tweedie (2000) is another technique to remove the most extreme small studies and employ a re-computing to adjust the asymmetric funnel plot. Third, according to Cole (2014), "a simple difference in means between published and unpublished studies provided a preliminary estimate of the magnitude of differences in average effect sizes" (p.11). None of the mentioned tests and techniques can prove the presence or absence of publication bias; employed together, they may minimize the risk of bias in the observed results and invalidity of generalization. For this purpose, we used CMA tailored for this purpose, with a lot of utilities and calculations. In using the funnel plot, in the absence of publication bias, the underlying assumption is that small studies are more prone to publication bias than large ones, and these studies will be scattered around the mean evenly, while the studies with high precision tend to cluster near the mean effect size. Figure 3, illustrates the funnel plot of the present meta-analysis.

The funnel plot shows that most of the studies are clustered around the mean and five studies at the lower right end of the funnel. With the normal visual inspection, it seems that both large and small studies are presented in the plot; however, the bottom of the graph is more crowded with smaller studies.

| JTLS | Journal of Teaching Language Skills **(JTLS)**
40(1), Winter 2021, pp. 187-235 | **204**
Esmaeel Saeedy
Robat |

THE EFFECTIVENESS OF ENGLISH LANGUAGE TEACHING METHODS

Figure 3.

Funnel Plot of Precision by Effect Sizes for the Observed and Imputed Studies

Funnel Plot of Precision by Hedges's g
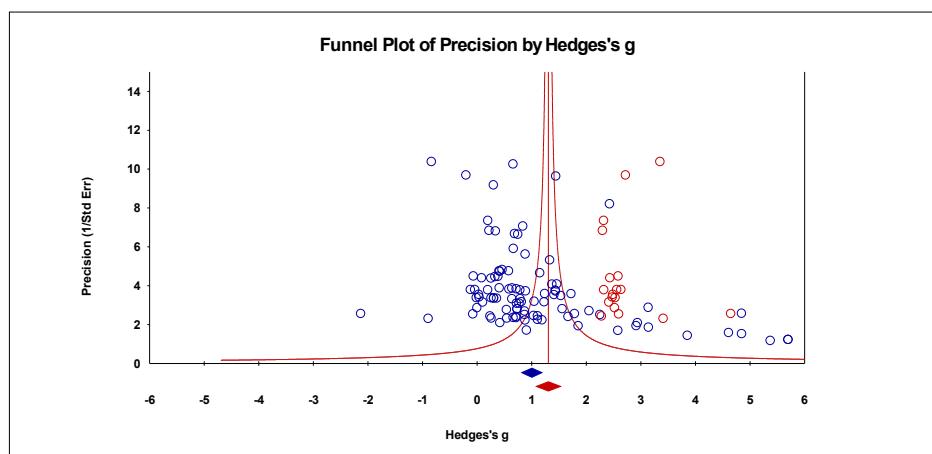
On the right side, there seem to be missing studies and a sign of publication bias. To address the issue of missing studies and publication bias, we performed some methods for detecting the presence of bias and assessing its impact on the analysis using CMA.

We also conducted a Trim-and-Fill analysis by Duval and R. Tweedie (2000) to find the missing values that would change the mean effect size if these values were imputed. It showed that under the fixed-effect model, the point estimate, and 95% confidence interval for the combined studies was 0.67 (0.62, 0.72). Using Trim and Fill, the imputed point estimate was 1.10 (1.06, 1.14). Under the random-effects model, the point estimate and 95% confidence interval for the combined studies was 1.00 (0.81, 1.18). Using Trim and Fill, the imputed point estimate was 1.30 (1.08, 1.52). In both models, as seen in Table 3., since the shift was trivial, we were more confident that publication bias was negligible.

Table 3.
Duval and Tweedie's Trim and Fill

	Studies trimmed	Point estimate	Fixed Effects (Lower, Upper)	Point estimate	Random Effects (Lower, Upper)	Q Value
Observed values		0.67	(0.62, 0.72)	1.10	(0.81, 1.18)	1365.67
Adjusted values	18	1.10	(1.06, 1.14)	1.30	(1.08, 1.52)	3027.27

Results

Descriptive Statistics of the Included Sample

The descriptive statistics of the primary studies included the kind of language teaching method, language skills or components, publication year, instructional levels, study design, effect size, sample size, and all other relevant information, see coding scheme in Table 2. Each effect size was counted as an independent study. Studies that yielded more than one effect size were listed separately in the tables with a. b. c. etc. labeling.

Overall Meta-Analysis Results

The first research question aimed at finding the overall effectiveness of each English language method and the combined effect of all methods on language learning achievement. Out of 56 primary studies, from 1970 to 2020, with 7960 participants from almost all continents, 97 effect sizes (Hedges' g) were collected for meta-analysis. In Table 5, the results of the average weighted Hedges'g, the 95% prediction intervals, the between-study variance, the Q-test for heterogeneity, the two-tailed test of null, and the percentage of variation between studies due to heterogeneity rather than sampling error, are presented for all categories. The effect sizes vary widely between –2.13 and 5.70. The overall effect size was found to be 0.67, with a standard error of

	Journal of Teaching Language Skills (JTLS)	**206**
JTLS	40(1), Winter 2021, pp. 187-235	**Esmaeel Saeedy Robat**

THE EFFECTIVENESS OF ENGLISH LANGUAGE TEACHING METHODS

0.02, a z-value for a test of the null of 28.24, a corresponding p-value of less than 0.001 for the fixed model, and a mean of 1.00, a standard error of 0.09, a z-value for a test of the null of 10.64 and a corresponding p-value of less than 0.001 for the random model. For both models, we concluded that the mean effect size was significant. For the random model (g > 0.8) based on Hedges'g interpretation, as stated by Cooper et al. (2009), it was found to be a large effect. However, the Q statistic on the heterogeneity of effect sizes was 1365.67, df=96, and p <.001. It indicates that all the variance is unlikely to be due to sampling error, and we also conclude that the true effect size is likely to vary from study to study. Therefore, the fixed model is violated and does not match the data. For that reason, we used the random effect model. (Borenstein et al., 2013)

The between-studies variance (τ^2) was estimated as 0.74. Some observed variance is due to actual differences in the size of the effect, while some are due to the errors of sampling. The I^2 statistic (92.97) represents the proportion of variation due to real differences and could be possibly explainable by covariates. In this case ($I^2 = 92.97\%$), the results indicate that almost all (over 90%) of the observed variances reflect real differences in study effects. As part of a sensitivity analysis, we also conducted the one-study-removed analysis via CMA. It was found that the average effect size does not seem to be heavily affected by outliers. It was also found that the average effect of language learning outcomes without the most prominent study was g = 1.00, p <.001, 95% CI [0.82, 1.19] $I^2 = 93\%$.

In Table 5, the results of separate calculations for each language teaching method on language learning achievement are also presented. The average effect size of communicative language teaching (g=1.02) was found to be large and significant. The average effect size of explicit instruction (g= 0.83) was also significant and considered a strong effect size (g = more than 0.80)

JTLS	Journal of Teaching Language Skills **(JTLS)** 40(1), Winter 2021, pp. 187-235	**207** Esmaeel Saeedy Robat

THE EFFECTIVENESS OF ENGLISH LANGUAGE TEACHING METHODS

according to Lakens (2013). Form-focused instruction also produced a significant (g= 1.19) and large effect size. Implicit instruction was the only teaching method that produced the smallest (g= 0.56) and no significant outcomes. The effect size for input-based instruction was found to be both significant and large (g=1.10). For meaning-focused instruction, a medium effect size (g=0.75) was detected, which was significant. The average effect size for output-based instruction was found to be the highest (g=2.22) and significant. The mean effect size for task-based instruction was also found to be high and significant (g= 0.94). Finally, for traditional language teaching, the average effect size was significant and strong (g=1.28). The Q-test was significant for all teaching methods, and it seems that the distribution of effect sizes is considered heterogeneous with $I^2 > 75\%$, indicating that a large proportion of the variability appears to be true variance. In a nutshell, 8 out of 9 language teaching methods produced significant effects on language learning achievement; however, the variance of the effects (0.74) needs to be addressed.

Table 4.

Results of the Univariate Random-Effects Meta-Analyses of English Language Teaching Methods on Language Learning Achievement

Dependent variable	Independent variables	N	K	g	SE	95% CI	Test of null		Heterogeneity				Tau-Squared		
							Z	P	Q	df	p	I^2	τ^2	se	τ
	Communicative language teaching	1589	6	1.02	.38	[.27,1.77]	2.66	.01	517.66**	5	.00	99.03	2.10	1.62	1.45
L2 learning outcomes	Explicit instruction	1694	18	.83	.23	[.38,1.27]	3.65	.00	95.52**	17	.00	82.20	.22	.11	.47
	Form focused instruction	577	14	1.19	.26	[.67,1.71]	4.50	.00	60.44**	13	.00	78.49	.42	.22	.65
	Implicit instruction	1071	9	.56	.32	[-.07,1.18]	1.75	.08	69.53	8	.00	88.49	.31	.22	.56
	Input based instruction	717	13	1.10	.28	[.56,1.64]	3.97	.00	188.97**	12	.00	93.65	1.29	.65	1.14
	Meaning focused instruction	772	12	.75	.28	[.19,1.30]	2.63	.01	29.47*	11	.00	62.68	.13	.11	.36
	Output based instruction	281	5	2.22	.46	[1.33,3.12]	4.86	.00	94.57**	4	.00	95.77	3.00	2.67	1.73
	Task based instruction	906	10	.94	.31	[.33,1.54]	3.04	.00	140.38**	9	.00	93.59	.94	.65	.97
	Traditional language teaching	353	10	1.28	.32	[.64,1.92]	3.95	.00	92.54**	9	.00	90.27	1.26	.75	1.12
Overall	All studies combined	7960	97	1.00	.09	[.82,1.19]	10.64	.00	1365.67**	96	.00	92.97	.74	0.17	.86

Note: N= total number of participants, k =number of effect sizes, g= mean weighted effect size in Hedges' g, SE = standard error, CI = confidence interval, Z = Z value, P = P value, Q = Cochran's heterogeneity test; df = degrees of freedom Q-test, τ^2 = between-study variance; I^2 = percentage of variation between studies that is due to heterogeneity rather than sampling error. V= variable

Moderator Analyses

For the second research question, it was aimed to investigate the moderating factors of the intervention programs on language learning achievement. We conducted meta-regression analysis for each group of moderator variables independently through sets of characteristics. Q-Statistic was used to evaluate if a particular variable was a significant moderator. We analyzed four major sets of variables, which included design characteristics, language characteristics, participants' characteristics, and teaching characteristics. We did not conduct meta-regression for each language teaching method separately since the analyses were not considered to be meaningful, as the total number of studies for each group of language teaching methods was not enough. Therefore, the majority of moderator categories could not be significantly compared. We investigated the impact of moderating variables on the overall effectiveness of language teaching methods on language learning. The detailed information is presented in Tables 4-8.

Design Characteristics

Since one of the crucial moderating factors concerning the effects of the intervention on language learning achievements seemed to be research design characteristics, we included three major groups such as type, place, and reach design separately and in combination as a set in the analysis. The three major groups included a total of 9 variables, and the Q statistic was used to examine the significance of the heterogeneity of the effects. A significant Q value indicates that the studies are not drawn from the same population, while a non-significant Q value shows the opposite (Higgins et al., 2019).

The first moderating group was the type of study, journal articles ($g = 1.04$), and Ph.D. dissertations ($g = 0.85$), the results of the Q test, $Q^b = 0.68$,

df = 1, p^b = 0.41, τ^2 =0.74, I^2 =92.89, R^2 =0.00, was found to be non-significant. Studies published as journal articles and those published as Ph.D. dissertations also indicated a strong positive effect size. The second group of moderating factors was the place of the study. The four regions included in the analysis, Asia (g = 1.00), Europe and Oceania, (g = 0.50) Middle East (g=1.37) North and South America (g =0.80). Studies from Asia, Middle East, North, and South America demonstrated a strong positive effect, and Europe and Oceania showed medium positive effects. The results of the Q statistics, Q^b =9.09, df = 3, p^b = 0.02, τ^2 =0.74, I^2 =92.76, R^2 =0.00 was found to be significant however, based on R^2 =0.00 and I^2 =92.76, the amount of explained variance by place of the study was zero. In other words, the place of the studies did not predict the effect of teaching methods on language learning achievement. We also examined the moderating effect of the study's design on learning achievement in our analysis. The results, Q^b =0.01, df = 1, p^b = 0.90, τ^2 =0.76, I^2 =93.03, R^2 =0.00, were not significant concerning the possible moderating effect of the design of the study on variables. Both groups of studies, the experimental (g =0.99) and Quasi-experimental (g =1.01) showed a largely positive effect. We also measured the combined effect of the design set, which included all the design groups as one set. The result of the Q test, Q^b =9.79, df = 5, p^b = 0.08, τ^2 =0.74, I^2 =92.66, R^2 =0.00, was found to be non-significant. More information could be consulted in Table 5.

Table 5.

Moderator Analysis of Design Characteristics

Moderator	N	K	G	95%CI	Q^b	df	p^b	τ^2	I^2	R^2
Type of the study					0.68	1	.41	.74	92.89	.00
Journal article	6522	78	1.04	[0.83,1.25]						
PhD dissertation	1438	19	0.85	[0.43,1.26]						
Place of the study					9.09*	3	0.02	0.74	92.76	00
Asia	2950	45	1.00	[0.73,1.27]						
Europe and Oceania	1846	14	0.50	[0.02,0.97]						
Middle East	2872	26	1.37	[1.02,1.73]						
North and South America	292	12	0.80	[0.26,1.34]						
Design of the study					0.01	1	0.90	0.76	93.03	0.00
Experimental	3725	44	0.99	[0.71,1.27]						
Quasi-experimental	4235	53	1.01	[0.76,1.27]						
Design characteristics set					9.79	5	0.08	0.74	92.66	0.00

Note: N = total number of participants, k = number of effect sizes, g = mean weighted effect size in Hedges' g, CI = confidence interval, P = P-value, Q^b = Q-between, df = degrees of freedom, τ^2 = between-study variance; I^2 = percentage of variation between studies that is due to heterogeneity rather than sampling error. R^2 = the proportion of the original variance explained by the covariates.

Participants Characteristics

The second set of moderating factors (Table 6) was the participants' characteristics. Four moderating groups, including participants' native language, L2 proficiency level, L2 proficiency measurement scale, and

Journal of Teaching Language Skills (JTLS)	**212**
40(1), Winter 2021, pp. 187-235	**Esmaeel Saeedy Robat**

THE EFFECTIVENESS OF ENGLISH LANGUAGE TEACHING METHODS

instructional level, were meta-analyzed both separately and in combination as a set. For participants' native language, we had Arabic (g = 1.91) Chinese and Taiwanese (g = 0.86) European Languages (g = 0.69) Japanese & Korean (g = 1.00) mixed and other languages (g = 0.85) Persian (g = 1.32). Except for European languages, with medium positive effect sizes, the effect sizes were found to be positive and large for other languages. The result of the Q test for participants' native language, Q^b =9.13, df = 5, p^b = 0.10, τ^2 =0.75, I^2 =92.76, R^2 =0.00, was found to be nonsignificant. For the second group of moderating factors, participants' L2 proficiency level, including beginner to low intermediate (g =0.91), high intermediate to advance (g =1.39), intermediate (g =0.83), mixed (g =0.76) and not reported (g =1.48), the results of the meta-regression were all positive and strong based on Hedge's g interpretation (Durlak, 2009). However, the results of the Q test for participants' L2 proficiency level, Q^b =8.47, df = 4, p^b = 0.07, τ^2 =0.76, I^2 =92.81, R^2 =0.00, were non-significant. The other group of moderating factors we included in the meta-regression were participants' proficiency measurement scales. Five variables, including institutional tests (g =1.01) researcher's judgment (g =1.05) self-designed tests (g =1.53) standardized tests (g =0.84) and the not-reported (g =1.08) were meta-analyzed. The results for all moderating variables in this category were positive and strong. The Q statistics for proficiency measurement scales, Q^b =5.08, df = 4, p^b = 0.27, τ^2 =0.78, I^2 =93.17, R^2 =0.00, was not significant. Our last group of moderating factors in this set was the participants' instructional level. Six levels of instruction, including elementary school (g =1.92) junior high school (g =0.94), senior high school (g =0.96), university-level (g =0.93) and Mixed level (g =0.67) were included in the analysis. We had positive and strong effect sizes for elementary school, junior high school, senior high school, university level,

| JTLS | Journal of Teaching Language Skills **(JTLS)**
40(1), Winter 2021, pp. 187-235 | **213**
Esmaeel Saeedy
Robat |

THE EFFECTIVENESS OF ENGLISH LANGUAGE TEACHING METHODS

and medium effect size for the mixed level. The Q test in the category, Q^b =8.43, df = 4, p^b = 0.07, τ^2 =0.78

Table 6.

Moderator Analysis for Participants' Characteristics

Moderator	N	K	g	95%CI	Q^b	df	p^b	τ^2	I2	R^2
Native language					9.13	5	0.10	0.75	92.76	0.00
Arabic	338	5	1.91	[1.1,2.72]						
Chinese & Taiwanese	837	10	0.86	[0.27,1.44]						
European Languages	1795	15	0.69	[0.22,1.15]						
Japanese & Korean	1534	31	1.00	[0.67,1.33]						
Mixed	2597	20	0.85	[0.44,1.26]						
Persian	859	16	1.32	[0.85,1.78]						
L2 Proficiency level					8.47	4	0.07	0.76	92.81	0.00
Beginner to low intermediate	1228	32	0.91	[0.58,1.25]						
High intermediate to advance	535	6	1.39	[0.63,2.16]						
Intermediate	2252	18	0.83	[0.40,1.25]						
Mixed	2699	21	0.76	[0.36,1.15]						
Not reported	1246	20	1.48	[1.06,1.89]						
L2 Proficiency Measurement					5.08	4	0.27	0.78	93.17	0.00

	Journal of Teaching Language Skills (JTLS) 40(1), Winter 2021, pp. 187-235	**214** **Esmaeel Saeedy Robat**

THE EFFECTIVENESS OF ENGLISH LANGUAGE TEACHING METHODS

Moderator	N	K	g	95%CI	Q^b	df	p^b	τ^2	I2	R^2
Institutional test	1578	16	1.01	[0.54,1.47]						
Not reported	200	6	1.08	[0.30,1.87]						
Researcher's judgment	1738	19	1.05	[0.63,1.48]						
Self-designed	447	13	1.53	[0.99,2.07]						
Standardized test	3997	43	0.84	[0.55,1.12]						
Instructional level					8.43	4	0.07	0.78	93.12	0.00
Elementary school	299	9	1.92	[1.26,2.58]						
Junior high school	2768	21	0.94	[0.53,1.35]						
Mixed	180	5	0.67	[-0.16,1.51]						
Senior high school	1642	10	0.96	[0.40,1.53]						
University	3071	52	0.93	[0.68,1.19]						
Participants characteristics set					31.10	17	0.01	0.86	93.17	.00

Note: N = total number of participants, k = number of effect sizes, g = mean weighted e ect size in Hedges' g, CI = confidence interval, P = P-value, Q^b = Q-between, df = degrees of freedom, τ^2 = between-study variance; I^2 = percentage of variation between studies that is due to heterogeneity rather than sampling error. R^2 = the proportion of the amount of heterogeneity accounted for.

I^2 =93.12, R^2 =0.00, was also found to be not significant. In order to see the differences of the participants set in general, concerning the moderating effect on language teaching methods on language learning, we calculated the Q test for all variables as one set. The results, Q^b =30.10, df = 17, p^b = 0.01, τ^2 =0.86, I^2 =93.17, R^2 =0. 00, were significant.

Language Characteristics

For the third set of moderating factors, Table 7, we calculated the effects of language skills and components through meta-regression analysis. Eight variables, including general language ability (g =1.18) grammar (g =0.88) listening (g =1.07) pragmatics (g =1.20) reading (g =0.91) speaking (g =0.47) vocabulary (g =1.76) and writing (g =0.25) were met-analyzed for possible moderating effects. For general language ability, grammar, pragmatics, reading, and vocabulary, we found strong positive effects. For the speaking scores, the effect was found to be medium, and for writing, we found a small effect size. The results of the Q test, Q^b =19.95, df = 7, p^b =0.00, τ^2 =0.71, I^2 =91.70, R^2 =0.04 were significant and based on ($R2$ =0.04) statistic at least part of the between-group variance is because of different language skills and components.

Table 7.

Moderator Analysis for Language Characteristics

Moderator	N	K	g	95%CI	Q^b	df	p^b	τ^2	I^2	R^2
					19.95	7	0.00	0.71	91.70	0.04
General	1247	8	1.18	[0.57,1.79]						
Grammar	2065	37	0.88	[0.59,1.18]						
Listening	591	3	1.07	[0.09,2.06]						
ragmatics	799	8	1.20	[0.58,1.82]						
Reading	1123	7	0.91	[0.26,1.57]						

Moderator	N	K	g	95%CI	Q^b	df	p^b	τ^2	I^2	R^2
Speaking	663	12	0.47	[-0.04,0.97]						
Vocabulary	883	17	1.76	[1.30,2.21]						
Writing	589	5	0.25	[-0.54,1.04]						

Note: N = total number of participants, k = number of effect sizes, g = mean weighted effect size in Hedges' g, CI = confidence interval, P = P-value, Q^b = Q-between, df = degrees of freedom, τ^2 = between-study variance; I^2 = percentage of variation between studies that is due to heterogeneity rather than sampling error. R^2= the proportion of the amount of heterogeneity accounted for.

Teaching Characteristics

For the last set of moderating factors, as seen in Table 8., the researchers calculated the effect sizes of language teaching methods including

Table 8.

Moderator Analysis for Teaching Characteristics

Moderator	N	K	g	95%CI	Q^b	df	p^b	τ^2	I2	R^2
Teaching methods					11.86	8	0.15	0.83	93.17	0.00
CLT	1589	6	1.02	[.27,1.77]						
EI	1694	18	.83	[.38,1.27]						
FFI	577	14	1.19	[.67,1.71]						
II	1071	9	.56	[-.07,1.18]						
IBI	717	13	1.10	[.56,1.64]						
MFI	772	12	.75	[.19,1.30]						
OBI	281	5	2.22	[1.33,3.12]						
TBI	906	10	.94	[.33,1.54]						
TLT	353	10	1.28	[.64,1.92]						

Note: N = total number of participants, k = number of effect sizes, g = mean weighted effect size in Hedges' g, CI = confidence interval, P = P-value, Q^b =

	Journal of Teaching Language Skills (JTLS) 40(1), Winter 2021, pp. 187-235	**217** **Esmaeel Saeedy Robat**

THE EFFECTIVENESS OF ENGLISH LANGUAGE TEACHING METHODS

Q-between, df = degrees of freedom, τ^2 = between-study variance; I^2 = percentage of variation between studies that is due to heterogeneity rather than sampling error. R^2 = the proportion of the amount of heterogeneity accounted for.

communicative language teaching (CLT), explicit instruction (EI), form-focused instruction (FFI), implicit instruction (II), input-based instruction (IBI), meaning-focused instruction (MFI), output-based instruction (OBI), task-based instruction (TBI), and traditional language teaching (TLT) on language learning scores of the learners. We were particularly interested in between-study differences. Almost all English teaching methods produced significant positive large effect sizes. The effect sizes of the univariate analyses are presented in Table 8. The results of the Q statistic, Q^b =11.86, df = 8, p^b = 0.15, τ^2 =0.83, I^2 =93.17, R^2 =0.00, were found to be nonsignificant. In other words, none of the teaching methods produced superior negative or positive results in comparison with others.

Discussion

What is the Overall Effectiveness of English Language Teaching Method(s) on L2 Achievement?

The first research question we addressed in the meta-analysis was about the overall effectiveness of English language teaching methods on L2 achievement. We found an overall positive significant large effect of g= 1.00 based on Hedge's g interpretation (g > 0.80;) according to Durlak (2009). The mean effect size for the eight language teaching methods that we calculated separately ranged from a positive medium-size effect of g = 0.56 for explicit instruction to a strong positive effect of g =2.22 for output-based instruction. No small or negative overall effects were found in our analysis. The aim of the study was simply to find the answer to the old question of the so-called

JTLS	Journal of Teaching Language Skills (JTLS) 40(1), Winter 2021, pp. 187-235	**218** Esmaeel Saeedy Robat

THE EFFECTIVENESS OF ENGLISH LANGUAGE TEACHING METHODS

"the best method?" of language teaching based on available data of almost a century of literature. Our final analysis was based on studies we found from the 1970s. However, our literature search started in 1900, and we could not find qualified studies before 1970. The original studies that we included in our analysis were also trying to answer the same mentioned question. We have some meta-analytic studies in recent years that tackled the same question. A similar meta-analysis study was conducted by Bryfonski and McKay (2017). They investigated the effect of task-based language teaching (TBLT) on L2 learning and found an overall large effect of d = 0.93 which was very close to our effect estimation g =0.94. Another meta-analysis study was conducted by Shintani (2015) concerning the effectiveness of processing and production-based instruction on L2 grammar acquisition. He found positive and strong effect sizes for both modes of teaching; however, the results were significantly different in favor of the processing instruction. The same issue was also evident in many primary studies which we included in other analyzes. Almost all the studies we investigated through our literature searching and screening showed positive effects of the indented language teaching method. In our analysis, traditional language teaching showed a strong effect size, which means it was very effective in a practical sense. However, it may not be justifiable from a theoretical perspective. In fact, traditional language teaching was more effective than communicative language teaching. The results of the study, and our experience of the primary and similar synthesis studies that we explored in the process of data screens, indicate two crucial issues. First, all teaching methods were effective, and second, the contexts were very heterogeneous. The two points seem to be contradictory, however, "the messy practice of crossing boundaries" as Canagarajah (2006, p. 30) stated is and probably was the nature of unity and diversity of language teaching over time. It was rightly pointed out by Larsen-Freeman (2012, p. 1) as "Diversity within

| | Journal of Teaching Language Skills (JTLS) | **219** |
| **JTLS** | 40(1), Winter 2021, pp. 187-235 | Esmaeel Saeedy Robat |

THE EFFECTIVENESS OF ENGLISH LANGUAGE TEACHING METHODS

unity" and "unity to diversity. She argued that language "teaching field had moved from unity to diversity" however, diversity continues to exist in our field, but it has a different relationship to unity. Diversity is not opposed to be unity but can be located within it. The results of the present meta-analysis once again reiterate the unity and diversity at the same time that is coexisting. The diversity side is demonstrated by different methods, places, times, levels of instruction and proficiency, etc., while the unity side is depicted through consistent significant results of all language teaching methods irrespective of their heterogeneous contexts. More than 85% of our original studies that were included in our analysis were from the 2000s onwards. It was almost a decade after it was said that TESOL methodology has moved "beyond methods" as pointed out by Richards (1990) to the "postmethod conditions" (Kumaravadivelu, 1994). It could be argued that methods may be considered dead in the minds of theorizers; however, it is undoubtedly not dead in the minds of real teachers in Pakistan, China, or Iran, etc. For example, in Iran, as mentioned in a study by Leather and Motallebzadeh (2015, p. 2), "the primary objective of English language teaching is "observing the principle of stabilization and enforcement of the Islamic–Iranian identity." In Pakistan, as stated by Manan et al. (2016), "English is used in the domains of power such as government, education, law, corporate sector, research, and media. The language hierarchy is based on power in which English stands as the most powerful" (p.227) or as mentioned by Haidar and Fang (2019) English language teaching in Pakistan is still linked to the country's British colonial background. In China, English language teaching is part of the national curriculum and official policy from 2001 mainly due to rapid Chinese economic and communication developments on a global scale as reported by Bolton and Graddol (2012). In European Union and the US, with regard to language learning and teaching policies, according to Jeffery & Van

	Journal of Teaching Language Skills (JTLS) 40(1), Winter 2021, pp. 187-235	**220** **Esmaeel Saeedy Robat**

THE EFFECTIVENESS OF ENGLISH LANGUAGE TEACHING METHODS

Beuningen (2020, p. 186), paradox and parallel trends coexist. Both regions reiterate the importance of globalization and regional mobilities and communications which require a more unified code of language learning and teaching programs. At the same time, issues such as "multilingual competencies", immigrant and minority learners, and "cultural and linguistic diversity" highlight the need for diversity and multilingualism. Each one of these countries or educational systems have been seeking their own policies and agenda of language teaching and learning which have other stakeholders like teachers, learners, researchers, parents, and many others on a national scale and at the same time, other players like international content providers, textbook producers, and assessment and testing institutions in a global scale. This intricate and interconnected web of massive players is both diversified and unified concerning nature and practice at the same time; "diversity within unity".

In a nutshell, the multiplicity and unity of language teaching were confirmed by the results of our study; language teaching methods produced almost unified effective outcomes within their diversified contexts.

Are the Effects of English Language Teaching Methods Moderated by the Features of Intervention Programs? Or what are Mediating Factors?

The second research question of the study was the analysis of moderating factors. Four major sets of moderators, including design characteristics, language characteristics, participants' characteristics, and teaching characteristics and 46 subsets, were meta-analyzed in our study.

The first set we addressed was about the possible moderator effects of design characteristics which included the type of the study, journal article and Ph.D. dissertation, place of the study, Asia, Europe and Oceania, Middle East, North, and South America, and the design of the study, Quasi-experimental

and experimental, on the e ectiveness of language teaching. For type and design, we did not find significant heterogeneity in e ect sizes between studies. However, for the place of the study, heterogeneity was found to be significant. A large effect size (g = 1.37) was found in the Middle East and Asia for all language teaching methods while it was a medium-size for Europe and Oceania and America. One possible explanation for the similarities between Asia and the Middle East, as compared with Europe and America, could be traced back to their sociocultural backgrounds as well as the distinct agenda of language teaching and learning in western and eastern counties. For example, a case study of communicative language teaching in China, Hu (2002) showed that:

…. CLT and the Chinese culture of learning are in conflict in several important respects, including philosophical assumptions about the nature of teaching and learning, perceptions of the respective roles and responsibilities of teachers and students, learning strategies encouraged, and qualities valued in teachers and students. In view of such fundamental differences, the paper contends that it is counterproductive to take an 'autonomous' attitude, rather than an 'ideological' one, to pedagogical innovations developed in a different sociocultural milieu (p.1).

We believe the same argument could be viable in the Middle East, particularly in Iran. Here, as mentioned by Leather and Motallebzadeh (2015) and Mirhosseini and Khodakarami (2016), language teaching and learning are too ideologically value-laden. The English language is considered a tool that imposing imperialistic and capitalistic evil intentions on our supposedly pure traditional Islamic values. Until recently, the dominant language teaching method was the GTM, and the principal purpose was to empower learners to be able to read and understand scientific texts at universities. Therefore, CLT or TBLT was not appropriate for the implementation of the state-imposed

JTLS	Journal of Teaching Language Skills (JTLS) 40(1), Winter 2021, pp. 187-235	**222** Esmaeel Saeedy Robat

THE EFFECTIVENESS OF ENGLISH LANGUAGE TEACHING METHODS

perspectives. However, with the rapid boom of new telecommunication technologies, internet access and speed, artificial intelligence, big data, learning analytics, massive open online courses (MOOC), smart apps, online translation services, etc., on the one hand, and active participation and competition of the private sector in the language teaching industry, on the other hand, as well as forces from globalization trends have led to tremendous changes of language teaching practices towards more homogeneity worldwide.

The second moderator set the researchers investigated in the analysis was participants' characteristics. In this set, we included four subsets, native language, L2 proficiency level, L2 proficiency measurement, instructional level, and 21 variables to provide a full picture of possible moderating effects (Table 6.). The results were found to be non-significant in all four sets separately; however, the heterogeneity of the combined effects was significant. We had some variation, from 0.67 to 1.92, in effect sizes of moderator sets, but not up to a significant level. All language teaching methods, we included in our analysis, were effective in their respective contexts irrespective of variations. In a broader context, the results are also in line with a meta-analysis on the TBLT implementation by Bryfonski and McKay (2017, p. 20). They concluded that TBLT was "effective pedagogy in a variety of contexts for learners at a variety" of contexts. Input-based practice and metapragmatic instruction were the topics of a meta-analysis that was conducted by Shirinbakhsh et al. (2018). They also found significant effects for both modes of instruction. Some other meta-analysis studies such as Liu and Zhang (2018), Jeon and Day (2016) and Hee and Junko (2014) on reading comprehension, Bakhshandeh and Jafari (2018) on grammar instruction through explicit instruction, Shintani et al. (2013) on productive and receptive knowledge, and five studies, Nematollahi et al. (2017), Hansol et al. (2018),

De Vos et al. (2018), (Huang et al., 2012; Liu & Zhang, 2018) on vocabulary teaching and learning in different contexts, also found significant effect sizes in their respective studies. Because of the nature of meta-analysis studies that are prone to publication bias, the findings of these studies, as well as our results, need to be interpreted with utmost caution.

The third set of moderating factors that we analyzed in our study was language characteristics including general language ability, grammar instruction, listening comprehension, pragmatic instruction, reading comprehension, speaking, vocabulary, and writing. The range of effect sizes (from g= 0.25 for writing to g =1.76 for vocabulary) and heterogeneity were found to be significant. One of the reasons could be due to the number of studies in each category. For writing, we had 6 studies, and for vocabulary 17, and for grammar 37. The findings indicate that the effects of all language teaching combined were stronger on vocabulary than other language skills and components. For speaking skills, a rather small effect was found. For all other skills and components (see Table 7), the effect sizes were found to be large and significant. Integrating or isolating language skills is not a recent concept. One of the pioneer studies on this issue was conducted by Selinker and Tomin (1986). They discussed five case studies on underlying theories and practices of integration and/or separation schemes of language skills. They pointed out that the concept is "grounded on a practical extension of theoretical prescriptions of the nature of language and language learning, historical tradition and practical constraints" (p.334). We should argue that the same concepts could be viable in today's English language teaching and learning practices. The results of our study also show that there was significant heterogeneity among teaching methods when we calculated the moderating effects of language skills. However, in general, the issue of separation and integrations of language skills was not a significant factor in determining the

superiority of one to another. What seems to be a more logical recommendation is the combination of both trends at different levels of teaching and learning. This perspective of blending language skill through the "macro and micro or in other words, skill based or an integrated skills" curriculum was discussed in a recent argument by Anderson (2019):

A skill based curriculum for language teaching isolates the macro language skills with the expressed purpose of breaking them into micro-skills for explicit instruction. An integrated   skills curriculum provides the opportunity to use the four macrolanguage skills within the same instructional context in order to accomplish a language learning task. (p.2).

The concept of the macro language skills and their integration in the curriculum development process was proposed by Kumaravadivelu (1994) as part of his "postmethod" concept.

Language teaching characteristics were the last set in our analysis. It was intended to find the differences in the effects of language teaching methods on language learning ability. In other words, which one of these methods (Table 8.) is "the best" method? The response was simple and concise; none of them. Our data, almost unanimously, revealed that all language teaching methods were positively effective, and the heterogeneity among leaching methods was not significant. What we found based on the literature of almost a century of language teaching is in line with Prabhu (1990). In his influential paper, he argued that the concept of "no best method" based on a broad interpretation is justifiable for three reasons. 1. "Different methods are best for different contexts. 2 All methods are partially true or valid. 3 The notion of good and bad methods is itself misguided" (p.161). In our analyses, we included many factors to see their possible effects on outcomes. What we found was that each language teaching context is unique and when the issue of THE best method is raised, it "all depends" as rightly mentioned by Prabhu

(1990), Kumaravadivelu (1994), Richards (1990), Bell (2003), Pennycook (1989) and many other shareholders in ELT.

Conclusion

Review studies, in general, provide us with the big picture of the issues. Systematic reviews with meta-analysis and moderator regression, in particular, are recent research tools in ELT that focus on areas where a considerable number of studies are available, and through the aggregation of findings, they provide us rigorous quantitative understandings, with more generalizability and external validity. (Borenstein et al., 2013) This meta-analysis included 56 studies, around 50 variables, 7960 participants, and 97 effect sizes into account. It was aimed to find the overall effectiveness of language teaching methods and their moderating factors on language learning ability. What we found is in line with current theories of language teaching in the postmethod era. However, language teaching and learning practices in the real world, in the classroom, are method-bound. The point was argued by Block (2001):

> while the method has been discredited at an etic level (that is, in the thinking and nomenclature of scholars) it certainly retains a great deal of vitality at the grassroots, emic level (that is, it is still part of the nomenclature of laypeople and teachers). (p. 72)

According to Block (2001, p. 3), "whether postmethodologists like it or not, methods have not gone away, nor are they likely to." Our findings also reiterate the fact that as long as specific methods fit certain contexts, those methods, with any kinds of meaning ascribed to the term method, whether theory- less/theory-neutral, or theory- bound or not, for those contexts, those

	Journal of Teaching Language Skills (JTLS) 40(1), Winter 2021, pp. 187-235	**226** Esmaeel Saeedy Robat

THE EFFECTIVENESS OF ENGLISH LANGUAGE TEACHING METHODS

methods are the "best methods." In a nutshell, our findings indicate that the best method is the one that works for you.

Meta-analysis studies are inherently prone to certain types of limitations. Comprehensive data search is an arduous and time-consuming task that could be expensive for non-academic researchers. For that reason, publication bias is a significant threat to the validity and generalizability of meta-analysis results in similar contexts. We employed most major techniques, available, to eliminate or at least to decrease the degree of publication bias in our analysis. The other major problem of meta-analysis studies based on Jak (2015, p. 2) "is the so-called 'apples and oranges' problem; that is, the differences among individual studies make the reliability of the meta-analysis questionable". We tried to diminish the impact of these issues through a rigorous coding book that we developed to organize and categorize moderators and variables both in terms of their fundamental consideration as well as technical and logistical characteristics. Finally, for some of our moderators, the small number of studies may also have been blamed for the lack of significance. For the same reason, we were not able to compare the effects of each language teaching method separately with other moderators. Therefore, we used the combined effect of methods and the moderating factors. Future research could tackle this issue and compare the effects of each method independently.

References

Abraham, L. B. (2008). Computer-mediated glosses in second language reading comprehension and vocabulary learning: A meta-analysis. *Computer Assisted Language Learning, 21*(3), 199-226.

Abu Radwan, A. (1999). *Focus -on -form instruction and the acquisition of English dative alternation: Does noticing help?* Unpublished doctoral dissertation. Georgetown University.

Akakura, M. (2012). Evaluating the effectiveness of explicit instruction on implicit and explicit L2 knowledge. *Language Teaching Research, 16*(1), 9-37.

	Journal of Teaching Language Skills **(JTLS)**	**227**
JTLS	40(1), Winter 2021, pp. 187-235	**Esmaeel Saeedy Robat**

THE EFFECTIVENESS OF ENGLISH LANGUAGE TEACHING METHODS

Al Muhaimeed, S. A. (2015). *Task-based language teaching vs. traditional way of English language teaching in Saudi intermediate schools: A comparative study.* Unpublished doctoral dissertation. Kent State University College.

Alamri, K., & Rogers, V. (2018). The effectiveness of different explicit vocabulary-teaching strategies on learners' retention of technical and academic words. *The Language Learning Journal, 46*(5), 622-633.

Amer, A. A. (1992). The Effect of story grammar Instruction on EFL students' comprehension of narrative text. *Reading in a Foreign Language, 8*(2), 711-720.

Ando, J. (1992). The Effects of two EFL (English as a Foreign Language) teaching approaches studied by the cotwin control method: A comparative study of the communicative and the grammatical approaches. *Acta geneticae medicae et gemellologiae: twin research, 41*(4), 335-352.

Bai, B. (2015). The effects of strategy-based writing instruction in Singapore primary schools. *System, 53*, 96-106.

Bakhshandeh, S., & Jafari, K. (2018). The effects of input enhancement and explicit instruction on developing Iranian lower-intermediate EFL learners' explicit knowledge of the passive voice. *Asian-Pacific Journal of Second and Foreign Language Education, 3*(1), 18.

Bangert-Drowns, R. L.-R., Lawrence M. (1991). *Meta-analysis in educational research.* ERIC clearinghouse on tests measurement and evaluation Washington DC.

Bataineh, R. F., Al-Qeyam, F. R., & Smadi, O. M. (2017). Does form-focused instruction really make a difference? Potential effectiveness in Jordanian EFL learners' linguistic and pragmatic knowledge acquisition. *Asian-Pacific Journal of Second Foreign Language Education, 2*(1), 17.

Bejarano, Y. (1987). A cooperative small-group methodology in the language classroom. *TESOL Quarterly, 21*(3), 483-504.

Bell, D. M. (2003). Method and postmethod: Are they really so incompatible? *TESOL Quarterly, 37*(2), 325-336.

Benati, A. (2005). The effects of processing instruction, traditional instruction and meaning—output instruction on the acquisition of the English past simple tense. *Language Teaching Research, 9*(1), 67-93.

Block, D. (2001). An exploration of the art and science debate in language education. In *Reflections on language and language learning: In honour of Arthur van Essen* (pp. 63–74). John Benjamins.

Blokdyk, G. (2020). *Meta-analysis a complete fuide - 2020 edition*. Emereo Pty Limited.

Bolton, K., & Graddol, D. (2012). English in China today. *English Today, 28*, 3-9.

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2013). *Comprehensive meta Analysis* In (Version 3) Biostat.

Borg, S. (2003). Teacher cognition in language teaching: A review of research on what language teachers think, know, believe, and do. *Language Teaching, 36*(2), 81-109.

Bryfonski, L., & McKay, T. H. (2017). TBLT implementation and evaluation: A meta-analysis. *Language Teaching Research, 0*(0), 1362168817744389.

Canagarajah, a. S. (2006). TESOL at forty: what are the issues? *TESOL Quarterly, 40*(1), 9-34.

Chan, M. (2019). The role of classroom input: processing instruction, traditional instruction, and implicit instruction in the acquisition of the English simple past by Cantonese ESL learners in Hong Kong. *System, 80*, 246-256.

Chang, A. C.-S., & Read, J. (2006). The effects of listening support on the listening performance of EFL learners. *TESOL Quarterly, 40*(2), 375-397.

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

Colovic-Markovic, J. (2013). *The effects of explicit instruction of formulaic sequences on second-language writers*. Unpublished doctoral dissertation. The University of Utah.

Cooper, H., Hedges, L. V., & Valentine, J. C. (2009). The handbook of research synthesis and meta-analysis 2nd edition. In *The Hand. of Res. Synthesis and Meta-Analysis, 2nd Ed.* (pp. 1-615). Russell Sage Foundation.

Cruz-Arcila, F. (2013). Accounting for difference and diversity in language teaching and learning in Colombia. *Educación y Educadores, 16*, 80-92.

De Vos, J. F., Schriefers, H., Nivard, M. G., & Lemhöfer, K. (2018). a meta-analysis and meta-regression of incidental second language word learning from spoken input. *Language Learning, 68*(4), 906-941.

*Doughty, C. J. S. (1988). *The effect of instruction on the acquisition of relativization in English as a second language*. Unpublished doctoral dissertation. The University of Pennsylvania.

Dueñas Vinuesa, M. (2002). *Diane larsen-freeman. 2000. Techniques and principles in language teaching. Second edition.* (Vol. 3).

JTLS	Journal of Teaching Language Skills (JTLS) 40(1), Winter 2021, pp. 187-235	**229** Esmaeel Saeedy Robat

THE EFFECTIVENESS OF ENGLISH LANGUAGE TEACHING METHODS

Durlak, J. A. (2009). How to select, calculate, and interpret effect sizes. *Journal of Pediatric Psychology, 34*(9), 917-928.

Duval, S., & Tweedie, R. (2000, Jun). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56*(2), 455-463.

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel–plot–based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56*(2), 455-463.

Ellis, R. (2010). Second language acquisition, teacher education and language pedagogy. *Language Teaching, 43*(2), 182-201.

Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of l2 grammar. *Studies in Second Language Acquisition, 28*(2), 339-368.

Farsani, H. M., Tavakoli, M., & Moinzadeh, A. (2012). The effect of task-based instruction on the acquisition and use of English existential constructions by Iranian EFL learners. *Innovation in Language Learning and Teaching, 6*(1), 45-67.

Fitton, L., McIlraith, A. L., & Wood, C. L. (2018). Shared book reading interventions with English learners: A meta-analysis. *Review of Educational Research, 88*(5), 712-751.

Fordyce, K. (2014). The differential effects of explicit and implicit instruction on efl learners' use of epistemic stance. *Applied Linguistics, 35*(1), 6-28.

Fukuta, J. (2016). Effects of task repetition on learners' attention orientation in L2 oral production. *Language Teaching Research, 20*(3), 321-340.

approach for the development of second language oral fluency, comprehensibility, and accentedness. *TESOL Quarterly, 51*(1), 115-142.

Ghavamnia, M., Eslami-Rasekh, A., & Vahid Dastjerdi, H. (2018). The effects of input-enhanced instruction on Iranian EFL learners' production of appropriate and accurate suggestions. *The Language Learning Journal, 46*(2), 114-131.

Haidar, S., & Fang, F. (2019). English language in education and globalization: a comparative analysis of the role of English in Pakistan and China. *Asia Pacific Journal of Education, 39*(2), 165-176.

Hansol, L., Mark, W., & Ho, L. J. (2018). The effects of corpus use on second language vocabulary learning: A multilevel meta-analysis. *Applied Linguistics.*

Harlen, W., & Crick, R. D. (2002). *A systematic review of the impact of summative assessment and tests on students' motivation for learning. Research Evidence in Education Library.*

Hee, J. E., & Junko, Y. (2014). L2 reading comprehension and its correlates: A meta-analysis. *Language Learning, 64*(1), 160-212.

Heidari-Shahreza, M. A., & Tavakoli, M. (2016). The effects of repetition and L1 lexicalization on incidental vocabulary acquisition by Iranian EFL Learners. *The Language Learning Journal, 44*(1), 17-32.

Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (2019). *Cochrane Handbook for Systematic Reviews of Interventions.* Cochrane, 2019.

Hinenoya, K., & Lyster, R. (2015). Identifiability and accessibility in learning definite article usages: A quasi-experimental study with Japanese learners of English. *Language Teaching Research, 19*(4), 397-415.

Hu, G. (2002). Potential cultural resistance to pedagogical imports: The case of communicative language teaching in China. *Language, Culture and Curriculum, 15*(2), 93-105.

Huang, S., Willson, V., & Eslami, Z. (2012). The effects of task involvement load on l2 incidental vocabulary learning: A meta-analytic study. *Modern Language Journal, 96*(4), 544-557.

Jak, S. (2015). *Meta-analytic structural equation modeling.* Springer.

Jeffery, J. V., & Van Beuningen, C. (2020). Language education in the EU and the US: Paradoxes and parallels. *PROSPECTS, 48*(3), 175-191.

Jeon, E., & Day, R. (2016). The effectiveness of ER on reading proficiency: A meta-analysis. *Reading in a Foreign Language, 28*(2), 246–265.

Jiang, X. (2007). *The impact of graphic organizer instruction on English -as-a -foreign -language college students' reading comprehension.* Unpublished doctoral dissertation. Northern Arizona University.

Jones, N. B. (1993). *Comparison of instructional strategies for teaching English in a Taiwan college: Traditional Chinese approach vs. Natural language approach* Unpublished doctoral dissertation. Washington State university

Kim, J.-W. (2008). *A comparative study of the effects of explicit-inductive/cooperative instruction versus explicit-deductive/individualistic instruction on the second language acquisition of English relative clauses in Korean university-level EFL learners.* Unpublished doctoral dissertation. Allian International University.

Kim, Y. (2013). Effects of pretask modeling on attention to form and question development. *TESOL Quarterly, 47*(1), 8-35.

Kim, Y., & McDonough, K. (2011). Using pretask modeling to encourage collaborative learning opportunities. *Language Teaching Research, 15*(2), 183-199.

Kumaravadivelu. (1994). The postmethod condition: (e)merging strategies for second/foreign language teaching. *TESOL Quarterly, 28*(1), 27-48.

Kumaravadivelu. (2006). *Understanding language teaching: From method to postmethod*. Routledge.

Kupferberg, I., & Olshtain, E. (1996). Explicit contrastive instruction facilitates the acquisition of difficult l2 forms. *Language Awareness, 5*(3-4), 149-165.

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in Psychology, 4*(863).

Larsen-Freeman, D. (2012). From unity to diversity... to diversity within unity. *English Teaching Forum, 50*(2), 6.

Leather, S., & Motallebzadeh, K. (2015). *Effecting methodological change through a trainer-training project: a tale of insider-outsider collaboration.* British Council Iran.

Lee, J., Jang, J., & Plonsky, L. (2015). The effectiveness of second language pronunciation instruction: A meta-analysis. *Applied Linguistics, 36*(3), 345-366.

Li, S. (2015). The associations between language aptitude and second language grammar acquisition: A meta-analytic review of five decades of research. *Applied Linguistics, 36*(3), 385-408.

Lightbown, P. M., & Spada, N. (1990). Focus-on-form and corrective feedback in communicative language teaching: Effects on second language learning. *Studies in Second Language Acquisition, 12*(4), 429-448.

Lindblad, T., & Levin, L. (1970). *Teaching grammar. An experiment in applied psycholinguistics. Assessing three different methods of teaching grammatical structures in English as a foreign language*.

Liu, J., & Zhang, J. (2018). The effects of extensive reading on English vocabulary learning: A meta-analysis. *English Language Teaching, 11*(6), 1-15.

Macdonald, D., Yule, G., & Powers, M. (1994). Attempts to improve english l2 pronunciation: The variable effects of different types of instruction. *Language Learning, 44*(1), 75-100.

| | Journal of Teaching Language Skills **(JTLS)** | **232** |
| JTS | 40(1), Winter 2021, pp. 187-235 | **Esmaeel Saeedy Robat** |

THE EFFECTIVENESS OF ENGLISH LANGUAGE TEACHING METHODS

Manan, S. A., David, M. K., & Dumanig, F. P. (2016). English language teaching in Pakistan: Language policies, delusions and solutions. In R. Kirkpatrick (Ed.), *English Language Education Policy in Asia* (Vol. 11). Springer.

McGuthrie, M. E. L. (2015). *An interventional study: Adult ESL beginners and advanced learners on acquiring and producing pragmatic requests.* Unpublished master's thesis. California State University.

Mirhosseini, S.-A., & Khodakarami, S. (2016). Aspects of English language education policies in Iran: 'Our own beliefs' or 'out of who you are'? *Journal of Multicultural Discourses, 11*(3), 283-299.

Mitsugi, M. (2017). Schema-based instruction on learning English polysemous words: Effects of instruction and learners' perceptions. *Journal of Pan-Pacific Association of Applied Linguistics, 21*(1), 21-43.

Mizumoto, A., & Takeuchi, O. (2009). Examining the effectiveness of explicit instruction of vocabulary learning strategies with Japanese EFL university students. *Language Teaching Research, 13*(4), 425-449.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009, Jul 21). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med, 6*(7).

Muranoi, H. (1997). *Effects of interaction enhancement on restructuring of interlanguage grammar: A cognitive approach to foreign language instruction.* Unpublished doctoral dissertation.Georgetown University, Ann Arbor.

Muranoi, H. (2000). Focus on form through interaction enhancement: Integrating formal instruction into a communicative task in EFL classrooms. *Language Learning, 50*(4), 617-673.

Nakatani, Y. (2005). The effects of awareness-raising training on oral communication strategy use. *The Modern Language Journal, 89*(1), 76-91.

Namaziandost, E., Dehkordi, E. S., & Shafiee, S. (2019). Comparing the effectiveness of input-based and output-based activities on productive knowledge of vocabulary among pre-intermediate EFL learners. *Asian-Pacific Journal of Second Foreign Language Education, 4*(1), 2.

Nassaji, H. (2020). The importance of using multiple measures or data sources in L2 instructional research. *Language Teaching Research, 24*(2), 131-135.

Nematollahi, B., Behjat, F., & Kargar, A. (2017). A meta-analysis of vocabulary learning strategies of EFL learners. *English Language Teaching 10*(5).

	Journal of Teaching Language Skills **(JTLS)**	**233**
JTLS	40(1), Winter 2021, pp. 187-235	**Esmaeel Saeedy Robat**

THE EFFECTIVENESS OF ENGLISH LANGUAGE TEACHING METHODS

Niu, R., & Helms-Park, R. (2014). Interaction, modality, and word engagement as factors in lexical learning in a Chinese context. *Language Teaching Research, 18*(3), 345-372.

Park, J. K. (2000). *The effects of forms and meaning-focused instruction on ESL learners' phonological acquisition.* Unpublished doctoral dissertation. University of Pennsylvania, Ann Arbor.

Peckham, D. W. (2000). *Attention and consciousness in second language acquisition: An investigation into the effects of instruction on noticing.* Unpublished doctoral dissertation. The University of Pittsburgh.

Pennycook, A. (1989). The concept of method, interested knowledge, and the politics of language teaching. *TESOL Quarterly, 23*(4), 589-618.

Prabhu, N. S. (1990). There is no best method—why? *TESOL Quarterly, 24*(2), 161-176.

Qi, X. E. (1994). *Consciousness-raising in ESL writing: Its effects on the syntactic skills and writing proficiency of university students.* Unpublished doctoral dissertation.State University of New York at Buffalo.

Qin, J. (2008). The effect of processing instruction and dictogloss tasks on the acquisition of the English passive voice. *Language Teaching Research, 12*(1), 61-82.

Rassaei, E. (2012). The effects of input-based and output-based instruction on l2 development. *TESL-EJ: Teaching English as a Second or Foreign Language, 16*(3).

Richards, J. C. (1990). *The language teaching matrix*. Cambridge University Press.

Robert, L., & Lawrence, M. (2006). Meta-analysis in educational research. *Practical Assessment, Research, and Evaluation, 2*(8).

Seiba, Z. (2002). *Classroom instruction and second language acquisition: The effect of explicit form-focused instruction on l2 learners' linguistic competence.* Unpublished doctoral dissertation. The University of Mississippi.

Selinker, L., & Tomin, R. S. (1986). An empirical look at the integration and separation of skills in ELT. *ELT Journal, 40*(3), 227-235.

Shabani, K., & Hosseinzadeh, S. (2013). A comparative study of the effects of teacher-initiated planned preemptive and reactive focus on form on L2 learners' accuracy in narrative writing. *Innovation in Language Learning and Teaching, 13*(1), 76-92.

| JTLS | Journal of Teaching Language Skills (JTLS)
40(1), Winter 2021, pp. 187-235 | **234**
Esmaeel Saeedy Robat |

THE EFFECTIVENESS OF ENGLISH LANGUAGE TEACHING METHODS

Shintani, N. (2011). A comparative study of the effects of input-based and production-based instruction on vocabulary acquisition by young EFL learners. *Language Teaching Research, 15*(2), 137-158.

Shintani, N. (2012). Input-based tasks and the acquisition of vocabulary and grammar: A process-product study. *Language Teaching Research, 16*(2), 253-279.

Shintani, N. (2013). The effect of focus on form and focus on forms instruction on the acquisition of productive knowledge of l2 vocabulary by young beginning-level learners. *TESOL Quarterly, 47*(1), 36-62.

Shintani, N. (2015). The effectiveness of processing instruction and production-based instruction on l2 grammar acquisition: A meta-analysis. *Applied Linguistics, 36*(3), 306-325.

Shintani, N., Li, S., & Ellis, R. (2013). Comprehension-based versus production-based grammar instruction: A meta-analysis of comparative studies. *Language Learning, 63*(2), 296-329.

Shirinbakhsh, S., Rasekh, A. E., & Tavakoli, M. (2018). Metapragmatic instruction (6Rs) versus input-based practice: a comparison of their effects on pragmatic accuracy and speed in the recognition and oral production of English refusals. *The Language Learning Journal, 46*(4), 514-537.

Sims, J. M. (1997). *A comparative study of improvements in reading comprehension of skill-based instruction and extensive reading for pleasure with Taiwanese freshman university students.* Unpublished doctoral dissertation. The Florida State University.

Slavin, R. E., & Cheung, A. (2005). A synthesis of research on language of reading instruction for English language learners. *Review of Educational Research, 75*(2), 247-284.

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

Stern, H. H., Tarone, E. E., Stern, H. H., Yule, G., & Stern, H. (1983). *Fundamental concepts of language teaching: Historical and interdisciplinary perspectives on applied linguistic research*. OUP Oxford.

Tammenga-Helmantel, M., Arends, E., & Canrinus, E. T. (2014). The effectiveness of deductive, inductive, implicit and incidental grammatical instruction in second language classrooms. *System, 45*, 198-210.

![JTLS logo]	Journal of Teaching Language Skills **(JTLS)**	**235**
	40(1), Winter 2021, pp. 187-235	**Esmaeel Saeedy Robat**

THE EFFECTIVENESS OF ENGLISH LANGUAGE TEACHING METHODS

Tode, T. (2007). Durability problems with explicit instruction in an EFL context: the learning of the English copula be before and after the introduction of the auxiliary be. *Language Teaching Research, 11*(1), 11-30.

Waters, A. (2009). A guide to Methodologia: past, present, and future. *ELT Journal, 63*(2), 108-115.

West, A. J. (2016). Adaptation of communicative language teaching methodology to an English textbook for English language learning of Nida students. *PASAA: A Journal of Language Teaching & Learning in Thailand, 52*, 25-52.

Yoshimura, F. (2006). Does manipulating foreknowledge of output tasks lead to differences in reading behavior, text comprehension, and noticing of language form? *Language Teaching Research, 10*(4), 419-434.