

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/335565589>

# Connecting language proficiency to teaching ability: A meta-analysis

Article in *Language Teaching Research* · September 2019

DOI: 10.1177/1362168819868667

CITATIONS

18

READS

5,171

3 authors:



**Farahnaz Faez**

The University of Western Ontario

35 PUBLICATIONS 567 CITATIONS

[SEE PROFILE](#)



**Michael Karas**

The University of Western Ontario

12 PUBLICATIONS 79 CITATIONS

[SEE PROFILE](#)



**Takumi Uchihara**

Waseda University

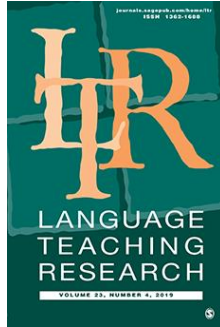
21 PUBLICATIONS 314 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Deep learning for Information Extraction [View project](#)



## Connecting language proficiency to teaching ability: A meta-analysis

Farahnaz Faez, Michael Karas, and Takumi Uchihara

University of Western Ontario

### Abstract

Most English language teachers around the world speak English as an additional language, and their level of English proficiency is often a matter of concern for them and their employers who associate higher levels of language proficiency with more effective teaching skills. To this end, several studies have examined the relationship between language proficiency and teachers' beliefs about their pedagogical capabilities, commonly known as self-efficacy. While generally studies show a positive relationship between language proficiency and self-perceived teaching ability, findings regarding the strength of the relationship, the role of specific language skills (e.g., speaking, listening), and how they interact with different teaching abilities (e.g., classroom management) are inconsistent. By combining data from 19 studies, this meta-analytic study examined the relationship between language proficiency and teaching self-efficacy and analyzed the role of various moderators such as teaching degree, teaching experience, and type of self-efficacy/proficiency measures. Findings reveal a moderate relationship ( $r = .37$ ) between language proficiency and teaching self-efficacy, with some moderator variables showing significant differences across correlations. The results indicate that only a small percentage of the variance in self-efficacy can be accounted for by teachers' language proficiency, suggesting that while language proficiency is important, there is more to self-efficacy than just language proficiency.

The unique role of English as the world's Lingua Franca and language of international communication has resulted in an increased demand for highly qualified and effective English language teachers—even more so with the early introduction of English in state curricula in several countries around the world. Teaching is a highly contextual activity and there are many factors that contribute to teacher expertise; thus, there are no universally acceptable criteria for identifying expert teachers (Tsui, 2003). However, teacher language proficiency has often been a main consideration in judgments related to teacher ability. The attention given to the English proficiency of teachers is a valid concern: Teachers require language-specific competencies such as the ability to provide good models of English, maintain fluent target language use, identify student errors, provide appropriate feedback, and engage in improvisational teaching (Medgyes, 2001; Richards, 2010). Equally important, teachers' levels of language proficiency impact their confidence in their teaching abilities (Eslami & Harper, 2018; Reves & Medgyes, 1994) and perceptions of their professional legitimacy (Seidlhofer, 1999).

### **Teacher Language Proficiency**

Defining proficiency is not an easy or straightforward task. Proficiency is contextually bound, and different levels and types of proficiency are required for different contexts and purposes. Compounding this complexity is the existence of different varieties of English. Proficiency in one variety does not necessarily mean proficiency in all varieties (Mahboob & Dutcher, 2014). Hence, the issue of teacher language proficiency and the question of what level of proficiency is required for teachers to be effective is a complicated matter.

Much of the research in the area of teacher language proficiency has emphasized general proficiency for teachers (e.g., Chacon, 2005; Eslami & Fatahi, 2008), often a source of concern for non-native English speaking teachers (NNESTs) (e.g., Kamhi-Stein, 2009; Murdoch, 1994). Richards (2010) argued that teachers need to reach a certain proficiency threshold in order to teach effectively. Emphasizing the notion of a threshold level, Tsang (2017) argued that teachers' general proficiency plays a substantial role in the classroom but only to a certain extent: Once a certain proficiency threshold is met, other factors such as teachers' pedagogical skills and personality play a more vital role. However, considering the various contexts, tasks, contents, and cultures teachers are required to perform in, this threshold remains an elusive notion (Elder & Kim, 2014).

This emphasis on teacher language proficiency, while a necessary discussion, has resulted in undesired outcomes. For example, native speakers are often favoured over non-native speakers in hiring practices because some contexts associate 'nativeness' with effective teaching (Freeman, 2016). Referred to as native-speakerism (Holliday, 2006), this remains a highly pervasive ideology in ELT, and subsequently 'nativeness' continues to be listed as a hiring criteria (Mahboob & Golden, 2013).

However, many have argued that native-like mastery of the English language is not necessary for teaching it well (Canagarajah, 1999; Richards, 2017). Richards (2017) argued that most English teachers around the world are not native speakers and "do not have nor need native-like ability" (p. 9) to teach well. Freeman (2017) also challenged the idea that general proficiency is needed for classroom purposes. Based on ideas from Language for Specific Purposes, he argued that teachers need *English-for-teaching* (Freeman, Katz, Garcia Gomez, & Burns, 2015)—a specific language set that highlights common words and phrases used by

teachers in the English language classroom. The English-for-teaching approach, though still developing, is one approach that researchers argue can help the field of ELT better prepare English teachers for specific tasks enacted in the classroom (Freeman, 2017; Richards, 2017).

While some have acknowledged that native-like language ability is not necessary for teaching, researchers have sought to understand how teacher's language proficiency may impact their confidence in their teaching capabilities. To investigate this relationship, researchers have looked at the concept of teacher self-efficacy.

### **(Language) Teacher Self-efficacy**

Self-efficacy has been investigated across numerous domains including healthcare, athletics, business, and education (Bandura, 1997). Most contemporary self-efficacy research is based in Bandura's (1997) sociocognitive perspective. Confidence in our abilities has long been seen as important, but research into self-efficacy has proven especially significant, with Bandura (1997) proclaiming "people's beliefs in their efficacy affect almost everything they do: how they think, motivate themselves, feel, and behave" (p. 19). In general education, teachers' self-efficacy beliefs have proven especially impactful. Defined as "teacher's individual beliefs in their capabilities to perform specific teaching tasks at a specified level of quality in a specified situation" (Dellinger, Bobbett, Olivier, & Ellet., 2008, p. 752), research has shown that teachers with high self-efficacy are often more motivated and persist longer when faced with adversity (Tschannen-Moran, Woolfolk Hoy, & Hoy, 1998), have better teacher performance (Klassen & Tze, 2014), display a greater commitment to teaching (Chestnut & Burley, 2015), and can even positively impact students' overall achievement (Caprara, Barbaranelli, Steca, & Malone, 2006). Although, this final claim requires further investigation (Klassen, Tze, Betts, & Gordon, 2011), establishing and nurturing teachers' confidence in their capabilities to teach is of clear importance.

Language teacher self-efficacy is now also an established domain of research. Unlike mainstream efficacy research, which has largely focused on North American contexts, language teacher self-efficacy research has predominantly focused on Asian and EFL contexts (Wyatt, 2018). The findings have shown interesting connections between teachers' self-efficacy and various pedagogical elements. For example, in Iran, more efficacious teachers showed a higher congruence between their stated pedagogical beliefs and teaching practices (Karimi, Abdullahi, & Haghghi, 2016). Other studies have noted the relationship between teachers' self-efficacy and elements such as teacher practical knowledge (Wyatt, 2010), teacher reflective practices (Moradkhani, Raygan, & Moein, 2017), and emotional intelligence (Rastegar & Memarpour, 2009). Higher levels of language teacher self-efficacy have also been linked to student achievement (Akbari & Karimi Allvar, 2010; Swanson, 2014). Although, similar to general education, such a significant claim must be made with caution because a causal relationship cannot be assumed, and further research is required (Wyatt, 2018). The burgeoning research in language teacher self-efficacy attests to its importance and impact.

Due to the nature of language teaching, a logical main focus of language teacher self-efficacy research has been on the relationship between self-efficacy and teachers' language proficiency. In the language classroom, language serves as both the content of the classroom but also, potentially, as the medium of instruction (Freeman, 2016). For many non-native speaking teachers, language proficiency has been a source of anxiety (e.g., Pasternak & Bailey, 2004).

Often (unfairly) compared to native speakers, non-native speaking teachers at times feel inadequate as language teachers because of (mis)perceptions of their own linguistic ability. Thus, the relationship between teacher self-efficacy and language proficiency has become a focal area of research. Some studies investigate this relationship broadly using overall measures of proficiency and self-efficacy (e.g. Sabokrouh & Barimani-Varandi, 2013), while others have investigated the relationship looking at specific language skills and different factors of self-efficacy (e.g. Chacon, 2005), with factors of self-efficacy (e.g. classroom management, student engagement, instructional strategies) most often drawn from the Teachers' Sense of Efficacy Scale (TSES) (Tschannen-Moran & Woolfolk Hoy, 2001), the most common self-efficacy scale used in language teacher self-efficacy research (Wyatt, 2018). Though numerous studies have looked at this relationship, quantitative results at times show a wide range of correlations (Faez & Karas, 2017). For example, Shim (2001) examined English language teachers in Korea and found little to no correlation between teachers' self-efficacy and their self-reported proficiency. Later studies found a moderate correlation in Venezuela (Chacon, 2005), Iran (Eslami & Fatahi, 2008), Turkey (Yilmaz, 2011), and other locations, while in the Middle East, Ghasemband (2014) found a very strong correlation. Generally speaking, most studies report a relationship, but the strength of the correlation varies, and results are less clear when examining specific pedagogical areas and language skills as this more nuanced analysis is less common. For example, looking at the relationship between speaking and classroom management, Eslami and Fatahi (2008) found a moderate significant correlation between the two, while Chacon (2005) found a weak non-significant correlation.

In this paper, we conduct a meta-analysis of these studies to gain further understandings about the relationship between teachers' self-efficacy beliefs (i.e., their confidence in their capabilities) and their language proficiency. In a meta-analysis, individual effect sizes from primary studies are pooled and averaged in order to yield a mean effect size of the relationship in question, which, in the current study, is the mean correlation between teacher self-efficacy beliefs and L2 proficiency. The next step is to conduct moderator analysis. Researchers often find that the effect size varies between studies. Thus, the value of conducting moderator analysis is to identify variables creating such between-study variation. In our meta-analysis, we investigated several variables as potential moderators, including characteristics of teachers (e.g., teaching experience) and different measures used in studies (e.g., teacher self-efficacy measures), which might impact the efficacy-proficiency correlation. By examining the impact of different moderator variables, we allowed for a more complete picture of the research literature that currently exists. The list of moderator variables and the rationale for inclusion of each of them is provided later in the paper.

## **Method**

The following research questions guided this meta-analysis:

- 1) What is the overall relationship between language proficiency and teaching self-efficacy?
- 2) To what extent is the relationship between language proficiency and teaching self-efficacy moderated by Type of Report, Teaching Degree, Teaching Experience, Measure of Self-Efficacy, Language of Self-Efficacy Scale, and Measure of Proficiency?
- 3) What is the correlation between different language skills (e.g., speaking, writing) and various self-efficacy factors (e.g., classroom management, instructional strategies, and

student engagement)?

### **Literature Search**

In line with previous meta-analytic studies (In'nami & Koizumi, 2010; Plonsky & Brown, 2015), initial literature searches were conducted on the following databases: Education Resources Information Center (ERIC), Linguistics and Language Behavior Abstracts (LLBA), ProQuest Dissertations and Theses, and PsycINFO. Following this, searches were also conducted on Google and Google Scholar, and finally the XXXXXX University library catalogue. Examples of search strings included, "Teacher efficacy and language proficiency," "Language teacher efficacy and language proficiency," "English teacher self-efficacy and English proficiency," and "Language teacher self-efficacy and proficiency." Finally, an ancestry search of materials was also conducted to find further materials.

### **Inclusion and Exclusion Criteria**

To be included in the meta-analysis, studies needed to address the relationship between language proficiency and teacher self-efficacy. While some studies examined the overall relationship between proficiency and self-efficacy (e.g., Crook, 2016), other studies focused on subskill analysis and the relationship between language skills (e.g., speaking) and self-efficacy factors (e.g., classroom management) (e.g., Chacon, 2005). We included both types of studies. For studies that focused on subskill analysis, we calculated the overall relationship by hand. Both published and unpublished materials (e.g., theses, proceedings, etc.) were included as this meta-analysis took an inclusive approach (Norris & Ortega, 2006). In such an approach, study quality is treated as an "empirical matter" (Norris & Ortega, 2006, p. 19), avoiding potential author bias and allowing the reader to make their own decisions. Furthermore, all studies needed to be written in English.

Studies that did not meet the above criteria were automatically excluded from analysis. For example, some studies (e.g., Liaw, 2004; Prayer, 2014) addressed the issue of proficiency by using nativeness (i.e., grouping teachers as native-speakers or non-native speakers) but did not include a measure of proficiency. Also, duplicate reports of the same data were omitted. For example, Chacon (2002) is a doctoral thesis, but Chacon (2005) draws on the same data and is in a peer-reviewed journal. Thus, Chacon (2002) was excluded to avoid misrepresenting the results. In total, we initially identified 20 studies.

### **Study Coding**

Studies were coded for (1) study context, including setting [e.g., foreign language (FL) or second language (SL)], institution type (e.g., elementary), L1, target language, and country; (2) participants (e.g., age, gender, teaching experience, academic degree, teacher L1, travel abroad experience); (3) study measures, including publication type, efficacy scale, proficiency measure, and type of proficiency (e.g., self-perceived or objective); and (4) methodological quality, related to measurement of teacher self-efficacy and language proficiency (i.e., number of scale points and items, reporting practice of reliability coefficients, and use of factor analysis). Coding was completed by the last two authors with an intercoder reliability of 99%. When there were coding inconsistencies, all three authors deliberated until a consensus was reached.

## Data Analysis

The current meta-analysis was conducted using the Comprehensive Meta-Analysis software (Version 3.3) using Fisher's  $z$ -transformation and inverse variance weighting (Borenstein, Hedges, Higgins, & Rothstein, 2011). The transformed value was converted back to the correlation coefficient in reporting the results. Following earlier meta-analytic studies focusing on correlations (e.g., Jeon & Yamashita, 2014; Li, 2016; Uchihara, Webb, & Yanagisawa, 2019), we used Fisher's  $z$  instead of  $r$  because it has better statistical properties such as normal distribution and stable variance (Lipsey & Wilson, 2001). Calculation of Fisher's  $z$  was based on the formula below, and the value for every study was automatically produced once we entered correlation coefficients in the spreadsheet of the Comprehensive Meta-Analysis software.

$$z = 0.5 \times \ln \left( \frac{1 + r}{1 - r} \right)$$

Another point we should consider when conducting a meta-analysis is a choice of model (Plonsky & Oswald, 2015). Two statistical models exist to conduct a meta-analysis: fixed-effect and random-effects model. A fixed-effect model assumes that there is only one true effect size across all studies, meaning that factors influencing the effect are the same in all studies. A random-effects model, in contrast, is based on the assumption that the true effect is different from study to study, influenced by sampling error and other factors such as age of participants, and study design. In the current study, since we assumed that there would be heterogeneity beyond that resulting from sampling error alone, the random-effects model was used to calculate the mean correlation between teachers' self-efficacy and their L2 proficiency, using a total of 20 effect sizes. After the effect-size aggregation, a series of moderator analyses was conducted using a random-effects model to identify the sources of variation in the size of correlations across studies. However, following Borenstein et al.'s (2011) recommendation, a fixed-effect model was used when the number of subsamples for a given moderator variable was smaller than five ( $k < 5$ ).

## General Features of the Data

All of the 20 studies are from the 2000s, with the earliest study being Shim's (2001) dissertation completed at Ohio State University, where a large body of teacher self-efficacy research has been completed (e.g., Tschannen-Moran & Woolfolk Hoy, 2001). The majority of studies were conducted in Asia ( $k = 15$ ), with four studies in Japan, three in South Korea, three in Iran, two in Thailand, and one each in Turkey, the Philippines, and the Middle East (country not specified in Ghasemboand, 2014). This matches Wyatt's (2018) findings that unlike general education self-efficacy research, which is mostly conducted in North America, language teacher self-efficacy research has been predominantly conducted in Asia. From the five remaining studies, one was conducted in Venezuela (Chacon, 2005) and four were conducted with language teachers in North America. Virtually all of the studies occurred in foreign language contexts, with the exception being Swanson (2012), which included French as a second language (FSL) teachers in Canada as well. With the exception of the work of Swanson (2010a; 2010b; 2012) and Swanson and Huff (2010), who included a mix of foreign/second language teachers (e.g., Spanish, French, German), all studies focused on English language teachers. See Table 1 for information on the 20 studies examined. Though initially included, Ghasemboand (2014) is eventually removed from analysis, as explained later, resulting in a final inclusion of 19 studies.

Table 1.  
*Studies included in Meta-Analysis (In chronological order)*

<b>Study</b>	<b>Study Context</b>	<b># of Participants</b>	<b>Type of Report</b>
Shim (2001)	South Korea	106	Thesis
Chacon (2005)	Venezuela	100	Journal Article
Eslami & Fatahi (2008)	Iran	40	Journal Article
Lee (2009)	South Korea	1327	Thesis
Swanson (2010a)	United States	463	Journal Article
Swanson (2010b)	United States	441	Conference Proceedings
Swanson & Huff (2010)	United States	167	Journal Article
Tayama (2011)	Japan	189	Journal Article
Yilmaz (2011)	Turkey	54	Journal Article
Nishino (2012)	Japan	139	Journal Article
Swanson (2012)	United States & Canada	1065	Journal Article
Sabokrouh & Barimani-Varandi (2013)	Iran	68	Journal Article
Best (2014)	Thailand	30	Thesis
Ghasemboland (2014)	Middle East	187	Thesis
Takayama (2015)	Japan	46	Thesis
Choi & Lee (2016) <sup>1</sup>	South Korea	167	Journal Article
Crook (2016)	Thailand	257	Thesis
Digap (2016)	Philippines	86	Journal Article
Thompson (2016)	Japan	113	Thesis

<sup>1</sup> In their manuscript, Choi and Lee (2016) use the gap between teachers' self-perceived proficiency and the level they believe is needed to teach in their context. To maintain consistency, we asked the authors to provide the simple correlation between self-efficacy and teachers' self-perceived proficiency, as found in the other studies included in this meta-analysis. The authors graciously provided this correlation ( $r = .443$ ) which is used in the analysis here.



Marashi & Azizi-Nassab (2018)	Iran	110	Journal Article
-------------------------------	------	-----	-----------------

### Moderators

**Type of Report.** As mentioned above, this meta-analysis adopted an inclusive approach in line with Norris and Ortega (2006). However, while this approach allows for a greater range and number of studies to be included, it creates the risk of including dubious low-quality research. To avoid doubt about our inclusion of both published and unpublished work and publications that were in lesser-known journals, we opted to use Type of Report as a moderator variable. To do this, we coded studies based on whether they were theses or published journal articles and if they were published in journals with or without impact factors. There were seven academic theses found, all of which were doctoral theses, except for Best (2014) which was a master's level thesis. All of the remaining studies were from published journals except for Swanson (2010b) which is a peer-reviewed published conference proceeding. However, because the publication for Swanson (2010b) is peer-reviewed, it is included as one of the seven publications without an impact factor, while the remaining six publications did have impact factors.

**Teaching Degree.** The level of teaching degree was also utilized as a moderator variable. Previous research has noted the importance of ELT specific degrees for teachers' confidence in their ability to engage with students (Akbari & Moradkhani, 2010), but research in this area appears limited. Outside of efficacy research, studies have noted the importance of level of degree for teachers (Akbari & Dadvand, 2011), and ELT specific degrees for English teachers (Akbari & Moradkhani, 2012). For this moderator, if participants had mixed educational backgrounds, we looked at the majority. If the majority of teachers in the study had a bachelor's degree, the study was considered majority BA. With this information, 10 studies were identified as having a majority of teachers with a bachelor's degree (BA  $k = 10$ ). Only four studies had the majority of teachers with an MA degree while six studies did not provide sufficient information to make a distinction.

**Teaching Experience.** Teaching experience was also coded as a moderator. This moderator was included because there has been some evidence that indicates more experienced teachers are more efficacious than novice teachers, from both general education (Tschannen-Moran & Woolfolk Hoy, 2001) and language teacher education (Akbari & Moradkhani, 2010). However, other studies have found no difference between experienced and novice teachers (Alemi & Pashmforoosh, 2013). While this study seeks to determine the relationship between self-efficacy and proficiency, these mixed results suggested it would be an interesting variable to include as part of this meta-analysis. Due to the great variation in the way researchers report information regarding teaching experience, it was not easy to code this variable in a consistent manner. Therefore, we followed criteria broadly, classifying teachers as experienced or less experienced. Some studies provided a mean of teaching experience (e.g., Swanson, 2010a) while others divided teachers' experience by a range of years (e.g., Crook, 2016), and some did not provide any information at all pertaining to teaching experience (e.g., Digap, 2016). Studies including more than half of the participants with at least 10 years of teaching experience, or with a mean teaching experience of over 10 years, were coded as experienced, with nine studies meeting this criterion. Other studies were coded as less experienced with five studies falling in this category. They were coded as 'less experienced' because the majority of teachers had less

than 10 years of experience. Finally, six studies did not provide sufficient information. For the most part, teachers in these studies were experienced.

**Measure of Self-Efficacy.** When coding for type of self-efficacy scale, three categories were used. The majority of studies use the Teacher Sense of Efficacy Scale (TSES), a general education measure created at Ohio State University (Tschannen-Moran & Woolfolk Hoy, 2001). Researchers have since modified the TSES to suit their own research purposes, but the range of modifications can vary. Thus, the first category includes studies that used the TSES without any major modifications, labeled ‘Original TSES’ ( $k = 13$ ). Studies in this category used some version of the TSES virtually ‘as is’ without significant modifications, either using the shortened 12-item version or the full 24-item scale. Certain studies added words to make the scale more relevant to their context (e.g., Chacon, 2005), or even translated the instrument into another language (e.g., Eslami & Fatahi, 2008), but for the most part, the TSES remained intact with its three-factor structure (classroom management, student engagement, instructional strategies). While some small modifications were made, these were considered minor. All of the studies by Swanson (2010a; 2010b; 2012) and Swanson and Huff (2010) are included in this category. These studies all used the TSES and Swanson’s Second/Foreign Language Teacher Efficacy Scale (S/FLTES). The S/FLTES was developed as a language teacher specific efficacy instrument, but researchers (e.g., Choi & Lee, 2016; Wyatt, 2018) noted an issue with one of its three factors, Content Knowledge. The subfactor of Content Knowledge appears to measure teachers’ general language proficiency rather than their beliefs about their pedagogical capabilities. Because we interpreted the subfactor of Content Knowledge as a measure of self-perceived proficiency, its correlation with the TSES scale was used for analysis.

Next, the second category included studies that used the TSES, but with significant alterations. This category, ‘Modified TSES’ ( $k = 3$ ), saw studies make more significant changes to the TSES by adding new items and/or subfactors. Because these changes were deemed more significant, they were placed in their own category. For example, Lee (2009) used the TSES, but added a new subfactor of Oral English Language Use, which measured teachers’ confidence in their abilities to use English in the classroom. This somewhat relates to the notion of English-for-teaching (Freeman, 2017) and thinking about language in relation to specific classroom actions. However, unlike the Content Knowledge subfactor on Swanson’s (2012) S/FLTES, it still relates to using English with classroom tasks, aligning it with self-efficacy doctrine (Bandura, 1997). Choi and Lee (2016) used parts of the TSES but also added items from Dellinger and colleagues (2008) and added an individual item of their own. Best (2014) also added items to make her scale more specific to the Thai context. Thus, to be considered in the modified category, studies needed to make more substantial changes than simple translation or adding a word to items as in the first category.

Finally, the third category consisted of studies that used a different scale altogether ( $k = 4$ ). Tayama (2011), Nishino (2012) and Thompson (2016) created their own study-specific scales while Shim (2001) used the Teacher Efficacy Scale (TES) created by Gibson and Dembo (1984). As can be seen from above, the vast majority of studies used the TSES in some form as a measure of self-efficacy.

**Language of Self-Efficacy Scale.** This moderator looked at the difference between

scales in the participants' L1 and scales that were in the L2. The rationale for including this moderator was to examine whether original items in English or their translated version in teachers' L1 had an impact on teachers' comprehension of the items. Some studies translated the self-efficacy measure into teachers' L1 ( $k = 8$ ). Crook (2016), for example, translated the TSES into Thai. Eslami and Fatahi (2008) translated the scale into Farsi for their Iranian participants. Other studies ( $k = 3$ ) did not translate the scale at all (e.g., Chacon, 2005), which potentially may have impacted teachers' understanding of items. The remaining 10 studies either did not provide information about scale language (e.g., Ghasemboand, 2014) or used a diverse participant pool that did not share a single L1 (e.g., Swanson 2010a). These remaining 10 studies were excluded from this analysis.

**Measure of Proficiency.** Finally, the type of proficiency measure was also coded. Most commonly, studies used proficiency measures from Butler (2004) and Chacon (2005), both of which are self-report measures. Butler's (2004) scale is drawn from the American Council on Teaching Foreign Languages (ACTFL) while Chacon (2005) created her own scale. The scales measure teachers' perceived language skills (e.g., speaking, writing), but the Chacon (2005) scale also includes culture as a measure of proficiency. Eslami and Fatahi (2008) and Yilmaz (2011) ( $k = 2$ ) used a combination of both scales. Best (2014), Choi and Lee (2016), and Lee (2009) ( $k = 3$ ) used only Butler (2004), while Chacon (2005) and Crook (2016) ( $k = 2$ ) used only Chacon's (2005) scale. The S/FLTES measure of Content Knowledge, which we interpreted as a measure of self-perceived general proficiency, was used by all of the Swanson studies ( $k = 4$ ). Finally, the 'Other' category had unique measures. Nishino (2012), Shim (2001), and Tayama (2011) all created their own proficiency measures. Thompson (2016) used multiple measures of proficiency, asking participants to estimate their current TOEIC scores and their current EIKEN score, which is a common Japanese English proficiency test. In order to not over represent the data sample, only one measure, the estimated current EIKEN score, was used. While the EIKEN is an objective measure of proficiency, because the participants were estimating their score, it was deemed a self-perceived measure. Objective proficiency measures were only used in  $k = 3$  studies. Sabokrouh and Barimani-Varandi (2013) used a TOEFL paper and pencil test but only measured grammar and reading ability. Marashi and Azizi-Nassab (2018) also used a sample TOEFL test but included listening, writing, reading, and speaking. Digap (2016) used the results of a local proficiency test administered to teachers in the Philippines.

### **Publication Bias**

Because studies reporting statistically significant findings or large effect sizes are more likely to be published, we were concerned about the overestimation of the aggregated effect sizes based on such biased samples of studies (Lipsey & Wilson, 2001). To assess publication bias in our data, we first employed a funnel plot (automatically produced using a function of the Comprehensive Meta-Analysis Software), which provides information regarding the relationship between measurement precision (i.e., standard error or function of sampling error) and the effect sizes in question. A well-balanced data set has a funnel shape, representing a symmetric shape of effect size distribution (i.e., the more sampling errors, the more likely effect sizes vary across studies). Conversely, an unbalanced data set represents an asymmetric funnel shape, an indication of publication bias. In the funnel plot (see Figure 1), precision values (i.e., the inverse of the standard error or  $1/\text{standard error}$ ) are plotted on the y-axis and effect sizes for each study on the x-axis. Large sample studies (or effect sizes) generally appear towards the top of the graph

and cluster around the mean effect size, whereas smaller studies expand across the bottom half of the graph. In the presence of publication bias, studies are normally missing on either side of the mean near the bottom.

The funnel plot (Figure 1) indicated that one study, Ghasemboland (2014), was a clear outlier ( $> 3SDs$ ) and was thus removed from subsequent analysis. Ghasemboland (2014) showed an overall correlation of  $r = .83$ , well above the correlations found in other studies. The plot also did not show a clustering of studies on the right side of the mean, indicating publication bias. This is not surprising given that the data set includes a number of unpublished (35% or 7 out of 20) and published works. Yet, it should be noted that studies appear to cluster on the left side of the mean, which suggests that studies with lower precision (or smaller sample sizes) tend to yield small effect sizes. Put simply, small-sample studies with large effect sizes were underrepresented, thereby underestimating the overall effect size. A trim-and-fill analysis was computed to search for the missing values that would change the mean effect size if these values were imputed. The result shows that under the random-effects model, five values were missing on the right side of the plot and imputing these values would have changed the mean effect size from .37 (95% CI = .33, .41) to .39 (95% CI = .35, .43). Acknowledging a slight underestimation of the mean effect size, we confirmed that the issue of publication bias was not serious in our data.

Figure 1

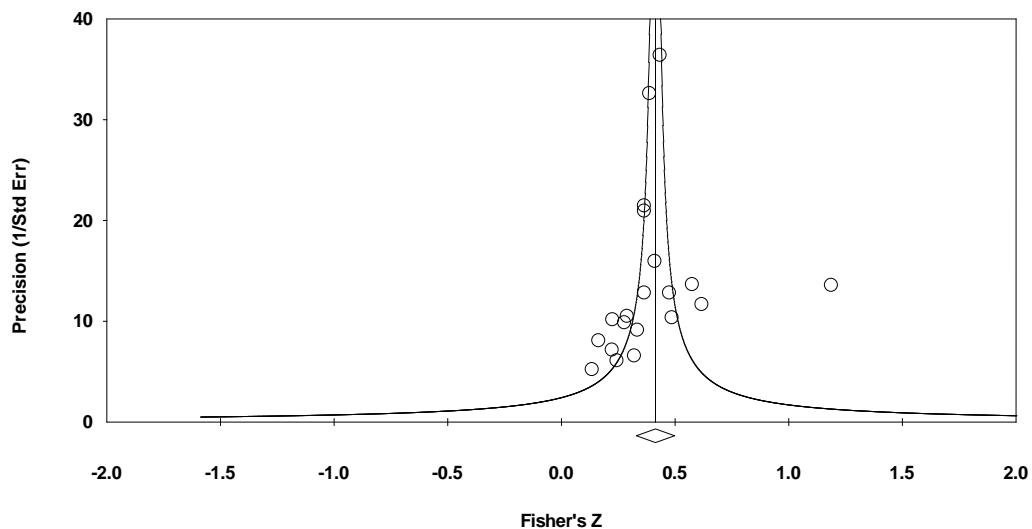


Figure 1. Funnel plot representing the relationship between measurement precision (the inverse of the standard error) and Fisher's z transformation of effect sizes.

## Results

### Methodological Quality

Before proceeding to our main research questions regarding the relationship between teacher self-efficacy and L2 proficiency, we examined the quality of methodological practices in our data set. Our primary concern related to data collection instruments (i.e., teacher self-efficacy scale and L2 proficiency measure) and reported reliability coefficients. For teacher self-efficacy measures, we coded four features: number of scale points, number of scale items, use of factor analysis, and reporting practice of reliability coefficients. For L2 proficiency measures, similarly, three methodological features were coded: number of scale points, number of scale items, and

reliability coefficient. It is possible that a greater number of scale points generate a higher correlation, whereas a restricted range of scale points might underestimate a true correlation (see Thorndike, 1949 for discussion of this issue in detail). Number of scale items are tied to issues of measurement errors because any data resulting from too few items are subject to construct-irrelevant errors and less likely to reflect a target construct. The rationale of inclusion of factor analysis also relates to quality of research in this field. Although the majority of primary studies in our data set used or adapted an established teacher self-efficacy scale (i.e., TSES), the factor structure obtained in earlier studies (e.g., Tschannen-Moran & Woolfolk Hoy, 2001) might not necessarily be generalized to another population. It is advisable to run factor analysis to inspect how scale items function in a given population in comparison to results of previous studies.

For teacher self-efficacy scale, the mean of number of scale points was 20.5 ( $SD = 40.9$ ) and noticeably much greater number of scale points (i.e., 101) was used in Swanson (2010a; 2010b; 2012) and Swanson and Huff (2010) compared to the remaining studies (range = 5 to 9). The mean of number of scale items was 13.6 ( $SD = 5.3$ , range = 4 to 24), indicating a fair amount of variation across studies. Ten out of 19 studies (53%) conducted factor analysis, and the majority of studies reported reliability coefficients (i.e., Cronbach alpha) (89%, 17 out of 19). The reported internal consistency was relatively high ( $M = .87$ ,  $SD = 0.08$ ) comparable or slightly higher than the median reliability coefficient found in SLA research (Plonsky & Derrick, 2016).

For L2 proficiency measure, the mean of number of scale points was 31.9 ( $SD = 41.7$ ) with, again, the greater number of points (i.e., 101) used by Swanson (2010a; 2010b; 2012) and Swanson and Huff (2010) compared to the other reports (range = 5 to 11). The mean number of scale items was 19.9 ( $SD = 35.0$ ) with a considerable variation across studies (range = 1 to 140). The majority of primary studies reported reliability coefficients (89%, 16 out of 18) with the average internal consistency ( $M = .89$ ,  $SD = 0.05$ ) higher than the median reliability coefficient found in SLA research (Plonsky & Derrick, 2016).

### **The Relationship between Teacher Self-Efficacy and L2 Proficiency**

After removing Ghasemboland (2014), 19 independent studies involving 4,968 participants were available for analyses for the overall relationship between teacher self-efficacy and L2 proficiency. The studies, including the relevant statistics (weighted correlation and their significance, 95% CI) and their graphic representations, are presented in Figure 2. The result shows an average correlation of .37, 95% CI [.33–.41], a medium effect according to (Plonsky & Oswald, 2014)<sup>2</sup> criteria for defining effect size values. As presented in Figure 2, although the CIs of some correlations crossed zero, every single correlation was positive (range = .135 to .550), indicating with great confidence and clarity the positive direction of the relationship between these two variables.

Figure 2

---

<sup>2</sup> Plonsky & Oswald (2014) considered small ( $r = .25$ ), medium ( $r = .40$ ), large ( $r = .60$ ) and suggested any value closer to  $r = .40$  is medium.

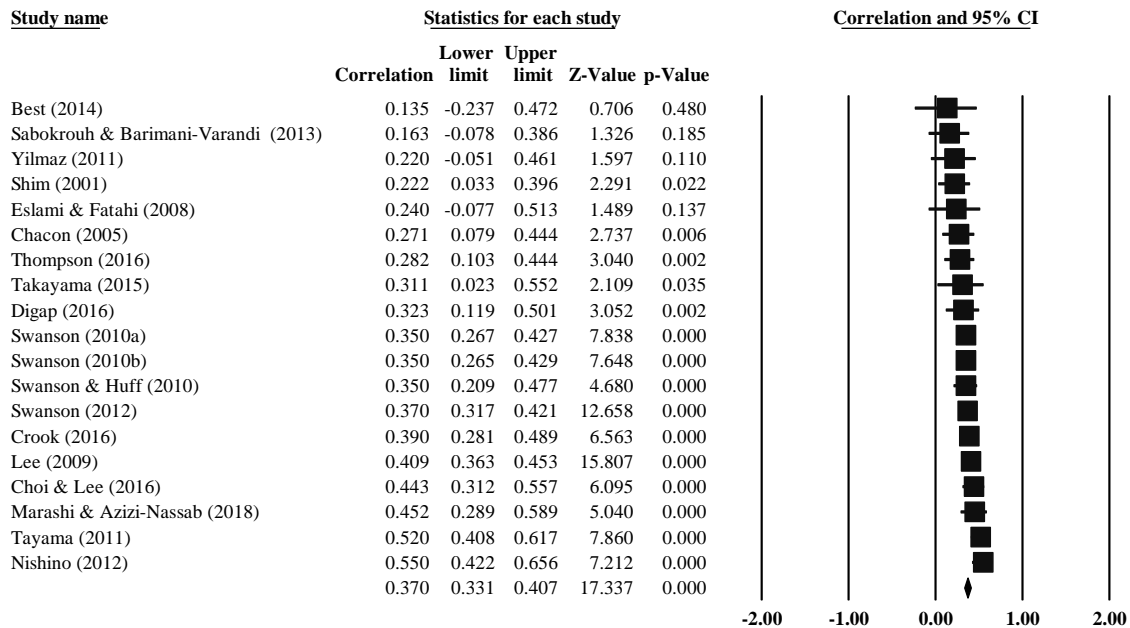


Figure 2. Overall average correlation (displayed by a diamond) and correlation with confidence interval for each study correlating teacher self-efficacy and L2 proficiency.

### Moderator Analysis

A moderator analysis was conducted to examine the extent to which the variance of the correlation coefficient between teacher self-efficacy and L2 proficiency across studies would be explained by six moderator variables: Type of Report, Teaching Degree, Teaching Experience, Measure of Self-Efficacy, Language of Self-Efficacy Scale, and Measure of Proficiency. Results show that none of these variables except for Measure of Self-Efficacy and Language of Self-Efficacy Scale were significantly related to the variability in the effect sizes across studies (see Table 2 for a summary of the results). However, it is important to still consider non-significant results as an overemphasis on null hypothesis testing can be overly simplistic and damaging to the advancement of theory (Plonsky, 2015). Furthermore, with small subgroups as found in this study, power may be a factor that contributes to the non-significant results. For Type of Report, all of the effect sizes were approaching medium (Plonsky & Oswald, 2015) with studies in journals with impact factors showing the largest correlation. In terms of the Teaching Degree moderator, studies with teachers who mostly held bachelor degrees showed a slightly higher correlation between self-efficacy and proficiency compared to studies with the majority of teachers with a master degree. The Teaching Experience moderator was also non-significant, but the studies categorized with teachers as ‘less experienced’ show a higher correlation compared to teachers who were deemed ‘experienced’. Finally, the difference between studies with self-reported proficiency and objective proficiency measures was not significantly different, but studies that used self-report measures showed a larger effect size than those with objective measures. Studies with objective measures showed a small effect while studies with self-reported proficiency were slightly larger and approaching a medium effect (Plonsky & Oswald, 2015).

The  $Q$  statistic showed that there was a significant difference in the size of correlation across the three types of SE scales. A further examination of the data revealed that L2

proficiency was more strongly correlated with teachers' self-efficacy when measured by non-TSES scales than by Original TSES (.43 vs. .35). For the most part, this difference is the result of the strong correlations found in Tayama (2011) and Nishino (2012), who created their own self-efficacy scales. The other two studies in this group, Shim (2001) and Thompson (2016), showed lower range correlations. The correlation for Modified TSES tends to be higher than that for Original TSES (.41 vs. .35). The magnitude of correlation for non-TSES scales and modified TSES was similar (.43 vs. .41). The result of language used in the SE questionnaire reveals that the correlation between teacher self-efficacy and L2 proficiency is significantly higher for L1-written SE scales ( $r = .41$ ) than the correlation for L2-written scales ( $r = .24$ ). This result indicates that the correlation between teachers' self-efficacy and L2 proficiency tends to be higher when participants are asked to answer SE questionnaires in their mother tongue than in an L2.

Table 2.

*Moderator Analysis of Six Variables*

Moderator Variable	<i>K</i>	<i>r</i>	95% CI	<i>Q<sub>between</sub></i>
<i>Type of Report</i>				0.711
Thesis	6	.347	[.273–.417]	
With impact factor	6	.367	[.292–.436]	
Without impact factor	7	.392	[.315–.463]	
<i>Teaching Degree</i>				0.982
Majority BA	9	.386	[.350–.421]	
Majority graduate	4	.360	[.322–.396]	
<i>Teaching Experience</i>				0.654
Experienced	9	.380	[.339–.419]	
Less experienced	4	.403	[.363–.442]	
<i>Measure of Self-Efficacy</i>				6.337*
Original TSES	12	.354	[.321–.385]	
Modified TSES	3	.408	[.366–.449]	
Non-TSES	4	.431	[.359–.497]	
<i>Language of Self-Efficacy Scale</i>				7.610**
L1	8	.405	[.369–.440]	
L2	3	.241	[.121–.353]	
<i>Measure of Proficiency</i>				0.516
Self-report measure	16	.381	[.356–.406]	
Objective measure	3	.341	[.229–.445]	

Note. \*\* $p < .01$ , \* $p < .05$

**TSES and L2 Proficiency**

To further examine the relationship between teacher self-efficacy and L2 proficiency, we used two analyses: We first examined the relationship between overall L2 proficiency and each of the efficacy subscales and next the relationship between overall teacher efficacy and each of the subskills of language proficiency. For the first analysis, we calculated average correlations ( $k = 8$ ) between L2 proficiency and each of the three TSES subscales: Student Engagement (SE), Classroom Management (CM), and Instructional Strategy (IS). As presented in Figure 3, IS-

proficiency associations ( $r = .37$ ) appear to be larger than either SE- or CM-associations ( $r = .28$  and  $.24$ , respectively). There was a statistically significant difference across the three associations ( $p = .03$ ), and post hoc comparison suggests that the difference between IS associations and SE or CM associations approached significance ( $p = .06$  and  $.07$ , respectively), a tendency indicating that the IS-proficiency relationship is stronger than either the SE- or CM-proficiency relationship.

Figure 3

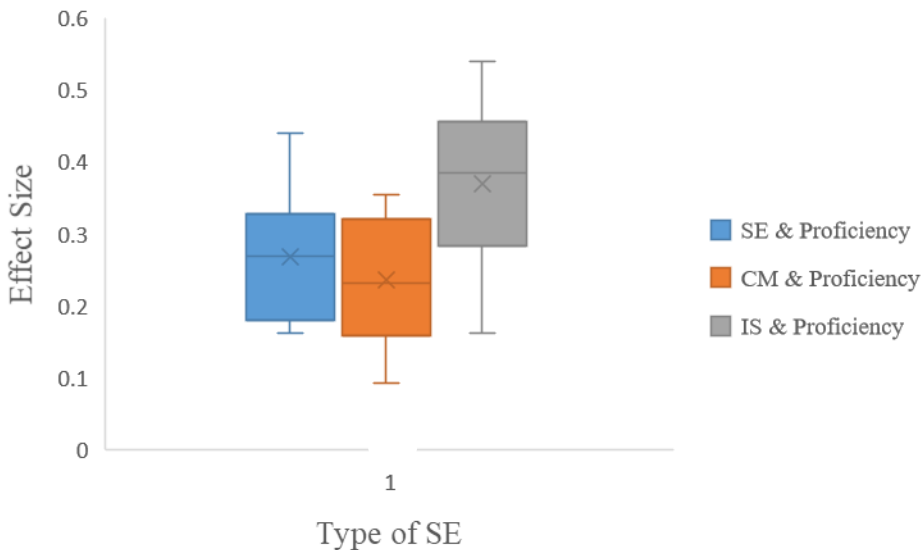


Figure 3. Comparison of the mean (denoted by x) and median (denoted by a line) effect sizes between teacher self-efficacy and L2 proficiency for each TSES subscale.

To probe into the relationship between teacher self-efficacy and four language skills (i.e., reading, listening, speaking, and writing), we also calculated average correlations ( $k = 4$ ) between overall TSES scores and each of the four language skills when data was available. Though tests of significance were not conducted due to small sample sizes, the size of average correlations appears to vary little across the four skills: speaking ( $r = .29$ ), listening ( $r = .30$ ), reading ( $r = .25$ ), and writing ( $r = .28$ ) (Figure 4).

Figure 4



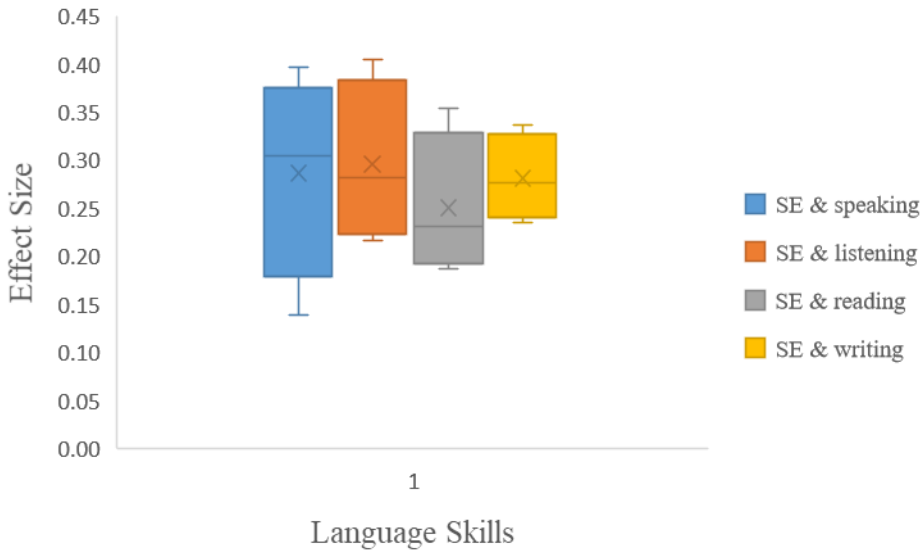


Figure 4. Comparison of the mean (denoted by x) and median (denoted by a line) effect sizes between teacher self-efficacy and L2 proficiency for each of the four language skills.

## Discussion

The findings of this meta-analysis showed that the overall relationship between language proficiency and teaching self-efficacy is  $r = 0.37$ , which is considered a moderate relationship. In quantitative terms, this finding means that only 13% of the variance in teaching self-efficacy is explained by language proficiency. Hence, the findings of this meta-analysis indicate that there is much more to language teachers' self-perceived teaching ability than just proficiency in the language. This finding aligns with arguments against using general proficiency as the only criterion for identifying effective teachers (e.g., Freeman, 2017; Richards, 2010, 2017). While proficiency is often noted as the key element for teachers' confidence (e.g., Kamhi-Stein, 2009; Murdoch, 1994), when analyzed in relation to the various required tasks of teachers, there is only a moderate relationship. The findings of this meta-analysis also support Tsang's (2017) claims, based on her qualitative study, that beyond a certain threshold of language proficiency, a teacher's pedagogical skills and personality become more important. While there are issues with the ways in which general language proficiency and teaching self-efficacy have been measured, a broader concern relates to the contribution of general language proficiency versus classroom-specific language proficiency or what Freeman et al. (2015) call English-for-teaching. A focus on teachers' capabilities to complete teaching tasks in English is a useful lens going forward.

The analysis of the moderator variables also yielded interesting findings. Only two of the six moderator variables, Measure of Self-Efficacy and Language of Self-Efficacy Scale, were related to the variability in the effect sizes. While the majority of studies used the TSES to measure self-efficacy, there was a significant difference between the four non-TSES studies and studies that used the TSES almost in its original form ( $k = 12$ ). As noted, the TSES is a validated tool to measure teacher efficacy in general education contexts and is not specific to language education. The highest correlations were found in Tayama (2011) and Nishino (2012), both of which utilized a study-specific measure of self-efficacy. Lee (2009) discussed the importance of language teaching specific measures of self-efficacy and noted this as a possible reason for the

low correlation found with Shim (2001) who used another general measure, the Teacher Efficacy Scale (TES). The difference between studies that used the modified TSES and original TSES was not big and was not statistically significant, although it was approaching significance ( $p = .08$ ). This seems to indicate that when studies use measures that specifically address the tasks of language teachers, rather than relying on general education measures alone, the relationship between self-efficacy and teacher language proficiency becomes somewhat stronger. Looking at the Language of Self-Efficacy Scale, it appears that when studies are written in the teachers' L1, correlations increase. Translation was deemed necessary for some studies as teachers' English proficiency was noted as very low (e.g., Crook, 2016), but others opted to leave the scales in English. Teacher's levels of understanding of the survey items may have impacted the results.

The remaining four moderator variables, Type of Resource, Teaching Degree, Teaching Experience, and Measure of Proficiency, all showed no significant difference between correlations. However, despite no statistical significance, the correlations were all positive and in the low to moderate range. For Type of Report, we were concerned about our broad inclusion criteria and the findings revealed that our concerns about published and unpublished materials, or inclusion of studies published in journals with and without impact factors, did not impact the findings. Measuring teachers' degrees and their experience in years was difficult as studies used different measures and reporting styles. We were forced to rely on broad categorizations, resulting in no significant differences. However, it should be noted that more precise measurements could impact the results. Finally, Measure of Proficiency showed no significant differences across correlations. Only three studies used objective measures of proficiency, while the remaining studies used self-report measures of general proficiency. Self-reported proficiency measures are notoriously inaccurate (e.g., Trofimovich, Isaacs, Kennedy, Saito, & Crowther, 2016); thus, it would be useful to see further studies use more objective measures of proficiency.

The examination of the relationship between teacher self-efficacy and L2 proficiency showed a stronger correlation between L2 proficiency and Instructional Strategies compared to Classroom Management and Student Engagement. This finding indicates that L2 proficiency is more closely related to Instructional Strategies than other classroom tasks such as Classroom Management and Student Engagement. Support for this finding is found in studies that included qualitative data (e.g., Lee, 2009), where participants stated that classroom management was mostly done in students' L1 in foreign language contexts. To further understand the contribution of different language skills required for effective teaching, our results did not show a particularly stronger relationship between any of the language skills (e.g., speaking and writing) with overall efficacy. This could be partially due to the small sample size used for this analysis (only 4 studies), but the findings indicated that average correlations are relatively similar for all language skills and hence all four language skills are equally valuable.

### **Directions for Future Research**

In spite of its significance, the area of teacher language proficiency and its relationship with teaching ability is a fairly undeveloped area of research. There are several reasons for this. Language proficiency is a construct that is contextually bound and hence difficult to define and assess. Different levels and types of proficiency are required for different contexts. Similarly, teaching efficacy is contextually and culturally bound (e.g., Tsui, 2003) and equally difficult to measure. Partly due to these reasons, most studies that have measured these constructs have relied on self-reports and self-reports can be unreliable (e.g., Trofimovich, et al., 2016). One

avenue of research in this area includes using more reliable measures of actual language proficiency and efficacy. While three studies in this meta-analysis were coded as using actual levels of language proficiency, these studies used a sample of a TOEFL paper and pencil test (e.g., Marashi & Azizi-Nassab, 2018; Sabokrouh and Barimani-Varandi, 2013) or used a local proficiency test (Digap, 2016), no actual standardized measure of proficiency has been used to date. The other issue pertains to examining the relationship between *general* language proficiency and teaching efficacy. With recent calls for acknowledging the significance of teacher *classroom* proficiency or *English-for-teaching* (Freeman, 2017; Richards, 2017) investigating which construct (general or classroom proficiency) is more strongly related to teaching self-efficacy is a future direction for research.

Another issue in examining teacher efficacy in language teacher education has been the pervasive use of a general education measure (e.g. TSES) for measuring language teacher self-efficacy. The TSES was developed to measure teacher efficacy in mainstream Kindergarten through grade 12 (K-12) classrooms. While its three-factor structure of Instructional Strategies, Classroom Management and Student Engagement all overlap with tasks of a language teacher, none of its items pertain specifically to language teaching. As noted in the moderator analysis section, fifteen studies used the TSES in either its original or modified form and only four studies used study-specific measures of teacher self-efficacy. By using study-specific measures, some of these studies reported higher correlations than the ones that used the original TSES. Therefore, another avenue of research includes using language teacher specific measures of teacher self-efficacy. However, regardless of what measure is used, there is also a gap in understanding how self-perceived teaching efficacy relates to teaching ability and teaching effectiveness. Studies that can juxtapose self-perceptions with those of learners, peers and administrator are needed but understandably more difficult to conduct. Also, studies that provide in-depth qualitative analysis of understanding what factors and why they contribute to teacher efficacy could add to our understanding of the issue. Finally, research is needed in second language contexts as the majority of studies appear to be conducted in EFL contexts (Wyatt, 2018).

As a final note for future research in light of methodological quality, researchers should be encouraged to examine the construct validity of the teacher self-efficacy measures in their specific population of participants. Our methodological synthesis showed that although the majority of studies (90%) followed recommended practices of reporting reliability coefficients, approximately half the studies failed to employ any factor analytic methods to examine factor structures underlying question items in use. Despite the frequent use of established scales like the TSES or its adapted version (79%, 15 out of 19), it is important not to assume that the self-efficacy measure works in the same way as in the original study (Tschannen-Moran & Woolfolk Hoy, 2001) but to inspect how factor structures in a given sample compare to results of earlier relevant studies. A possible account for this issue relates to sample size. The mean number of samples in the studies not conducting factor analysis was 103 (range = 40 to 257), whereas that of the studies conducting factor analysis was 411 (range = 106 to 1065). The former exceeded but the latter fell short of the mean of sample size found in SLA research ( $M = 381.8$ , range = 25 to 2278; Plonsky & Gonulal, 2015). We suggest that future studies of this kind should conduct factor analytic procedures to investigate the construct validity of teacher self-efficacy measures with a larger number of participants.

## Conclusion

The findings of this meta-analysis show a moderate relationship between language proficiency and self-perceived teaching ability, and only a small percentage (13%) of the variance in teaching can be accounted for by teachers' language proficiency. This finding suggests that while language proficiency is important, there is more to teaching self-efficacy than just language proficiency. The findings also revealed that only two variables, Measure of Self-Efficacy and Language of Self-Efficacy Scale were significantly related to the variability in the effect sizes across studies. This result indicates that the correlation between teachers' self-efficacy and L2 proficiency tends to be higher when scales specific to the language teaching task are used and when participants are asked to answer questionnaires in their mother tongue. This finding has significant implications for research on self-efficacy, as most research in the language teacher education context has used the TSES which is a self-efficacy scale developed for the general education context and not the language teaching context. The results also showed that L2 proficiency shows a stronger correlation with Instructional Strategies than Classroom Management and Student Engagement but that there is no significant difference among the different language skills for their contribution to efficacy factors. While teacher language proficiency will be an important variable for English language teachers, this study emphasizes that it is not the sole measure of teaching efficacy. This study has significant implications not only for language teachers but also for administrators and hiring managers. Hiring practices should not be biased towards favoring native speakers over non-native speakers. Instead, moving forward, it is important for all stakeholders to consider teacher language proficiency in relation to the specific tasks required of teachers to avoid general language proficiency comparisons couched in the native speakerist perspective.

## References

### **\* indicates one of the final 19 studies included in the meta-analysis**

- Akbari, R., & Dadvand, B. (2011). Does formal teacher education make a difference? A comparison of pedagogical thought units of B.A. versus M.A. teachers. *The Modern Language Journal*, 95(1), 44–60. <http://dx.doi.org/10.1111/j.1540-4781.2010.01142.x>
- Akbari, R., & Karimi Allvar, N. (2010). L2 teacher characteristics as predictors of students' academic achievement. *TESL-EJ*, 13(4), 1–22.
- Akbari, R., & Moradkhani, S. (2010). Iranian English teachers' self-efficacy: Do academic degree and experience make a difference? *Pazhuhesh-e Zabanha-ye Khareji*, 56(Spring), 25–47.
- Akbari, R., & Moradkhani, S. (2012). Is a degree relevant? A comparison pedagogical thought units of teachers with and without ELT-related academic credentials. *Australian Journal of Teacher Education*, 37(12), 77–93. <http://dx.doi.org/10.14221/ajte.2012v37n12.5>
- Alemi, M., & Pashmforoosh, R. (2013). EFL teachers' self-efficacy beliefs in teaching literature. *The International Journal of Literacies*, 19, 25–36. <https://doi.org/10.18848/2327-0136/CGP/v19i02/48770>
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: W.H. Freeman and Company.

- \*Best, B. (2014). A study of elementary school Thai English teachers' perceived English proficiency and self-reported English teaching efficacy (Master's dissertation). *Language in India*, 14(7), 75–195.
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2011). *Introduction to meta-analysis*. West Sussex, UK: Wiley.
- Butler, Y. G. (2004). What level of English proficiency do elementary school teachers need to attain to teach EFL? Case studies from Korea, Taiwan, and Japan. *TESOL Quarterly*, 38(2), 245–278. <http://dx.doi.org/10.2307/3588380>
- Canagarajah, A. S. (1999). Interrogating the Native Speaker Fallacy; non-linguistic roots, non-pedagogical results. In G. Braine (Ed.) *Non-native educators in English language teaching* (pp. 77–92). Mahwah NJ: Lawrence Erlbaum.
- Caprara, G. V., Barbaranelli, C., Steca, P., & Malone, P. S. (2006). Teachers' self-efficacy beliefs as determinants of job satisfaction and students' academic achievement: A study at the school level. *Journal of School Psychology*, 44, 473–490. <https://doi.org/10.1016/j.jsp.2006.09.001>
- Chacon, C. T. (2002). *Teachers' sense of efficacy and selected characteristics of selected English as a foreign language Venezuelan middle school teachers* (Unpublished doctoral dissertation). The Ohio State University, Columbus, OH.
- \*Chacon, C. T. (2005) Teachers' perceived efficacy among English as a foreign language teachers in middle schools in Venezuela. *Teaching and Teacher Education*, 21, 257–272. <https://doi.org/10.1016/j.tate.2005.01.001>
- Chestnut, S. R., & Burley, H. (2015). Self-efficacy as a predictor of commitment to the teaching profession: A meta-analysis. *Educational Research Review*, 15, 1–16. [doi:10.1016/j.edurev.2015.02.001](https://doi.org/10.1016/j.edurev.2015.02.001)
- \*Choi, E., & Lee, J. (2016). Investigating the relationship of target language proficiency and self-efficacy among nonnative EFL teachers. *System*, 58, 49–63. <https://doi.org/10.1016/j.system.2016.02.010>
- \*Crook, C. C. (2016). *The predictive relationship between specific teacher characteristics and the perceived sense of teacher self-efficacy of non-native English speaking teachers of English as a foreign language in rural Thailand* (Unpublished doctoral dissertation). Liberty University, Lynchburg, VA.
- Dellinger, A. B., Bobbett, J. J., Olivier, D. F., & Ellet, C. D. (2008). Measuring teachers' self-efficacy beliefs: Development and use of the TEBS-Self. *Teaching and Teacher Education*, 24, 751–766. <https://doi.org/10.1016/j.tate.2007.02.010>
- \*Digap, A. L. C. (2016). Self-efficacy, English proficiency and effectiveness of teachers of English in the secondary schools. *SMCC Higher Education Research Journal*, 2, 72–89. <http://dx.doi.org/10.18868/sher2j.02.00416.07>
- Elder, C. & Kim SHO (2014). Assessing teachers' language proficiency. In Junnan AJ (ed.) *The Companion to Language Assessment*, Chichester: John Wiley and Sons. Inc.

- \*Eslami, Z. R., & Fatahi, A. (2008). Teachers' sense of self-efficacy, English proficiency, and instructional strategies: A study of non-native EFL teachers in Iran. *TESL-EJ*, 11(4), 1–19.
- Eslami, Z., & Harper, K. (2018). Language proficiency and NNESTs. *The TESOL Encyclopedia of English Language Teaching*. <https://doi.org/10.1002/9781118784235.eelt0022>
- Faez, F., & Karas, M. (2017). Connecting language proficiency to (self-reported) teaching ability: A review and analysis of research. *RELC Journal*, 48(1), 135–151. <https://doi.org/10.1177/0033688217694755>
- Freeman, D. (2016). *Educating second language teachers: The same things done differently*. Oxford, UK: Oxford University Press.
- Freeman, D. (2017). The case for teachers' classroom English proficiency. *RELC Journal*, 48(1), 31–52. <https://doi.org/10.1177/0033688217691073>
- Freeman, D., Katz, A., Garcia Gomez, P. G., & Burns, A. (2015). English-for-teaching: Rethinking teacher proficiency in the classroom. *ELT Journal*, 69(2), 129–139. doi:10.1093/elt/ccu074
- Ghasemband, F. (2014). *Self-efficacy beliefs among English as a foreign language teachers in selected language centers*. (Unpublished doctoral thesis). University of Malaya, Kuala Lumpur, MA.
- Gibson, S., & Dembo, M. (1984). Teacher efficacy: A construct validation. *Journal of Educational Psychology*, 76(4), 569–582. <http://dx.doi.org/10.1037/0022-0663.76.4.569>
- Holliday, A. (2006). Native-speakerism. *ELT Journal*, 60(4), 385–387. doi:10.1093/elt/ccl030
- In'nami, Y., & Koizumi, R. (2010). Database selection guidelines for meta-analysis in applied linguistics. *TESOL Quarterly*, 44(1), 169–184. <https://doi.org/10.5054/tq.2010.215253>
- Jeon, E. H., & Yamashita, J. (2014). L2 reading comprehension and its correlates: A meta-analysis. *Language Learning*, 64(1), 160–212. <https://doi.org/10.1111/lang.12034>
- Kamhi-Stein, L. D. (2009). Teacher preparation and nonnative English-speaking educators. In A. Burns & J.C. Richards (Eds.) *The Cambridge guide to second language teacher education* (pp. 91–101). New York, NY: Cambridge University Press.
- Karimi, M. N., Abdullahi, K., & Haghghi, J. K. (2016). English as a foreign language teachers' self-efficacy as a determinant of correspondence between their professed orientations toward reading and their reading instructional practices. *Innovation in Language Learning and Teaching*, 10(3), 155–170. <https://doi.org/10.1080/17501229.2014.920847>
- Klassen, R. M., & Tze, V. M. C. (2014). Teachers' self-efficacy, personality, and teaching effectiveness: A meta-analysis. *Educational Research Review*, 12, 59–76. <https://doi.org/10.1016/j.edurev.2014.06.001>
- Klassen, R. M., Tze, V. M. C., Betts, S. M., & Gordon, K. A. (2011). Teacher efficacy research 1998 – 2009: Signs of progress or Unfulfilled promise? *Educational Psychological Review*, 23, 21–43. <https://doi.org/10.1007/s10648-010-9141-8>

- \*Lee, J. A. (2009). *Teachers' sense of efficacy in teaching English, perceived English language proficiency, and attitudes toward the English language: A case study of Korean public elementary school teachers* (Unpublished doctoral dissertation). The Ohio State University, Columbus, OH.
- Li, S. (2016). The construct validity of language aptitude. *Studies in Second Language Acquisition*, 38(4), 801–842. <https://doi.org/10.1017/S027226311500042X>
- Liaw, E. C. (2004). “How are they different?” *A comparative study of native and non-native foreign language teaching assistants regarding selected characteristics: Teacher efficacy, approach to language teaching/learning, teaching strategies and perception of nativenesship* (Unpublished doctoral dissertation). The Ohio State University, Columbus, OH.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.
- Mahboob, A., & Dutcher, L. (2014). Dynamic approach to language proficiency – A model. In A. Mahboob & L. Barratt (Eds.), *Englishes in multilingual contexts: Language variation and education* (pp. 117–136). Springer.
- Mahboob, A., & Golden, R. (2013). Looking for native speakers of English: Discrimination in English language teaching job advertisements. *Voices in Asia Journal*, 1(1), 72–81.
- \*Marashi, H., & Azizi-Nassab, F. (2018). EFL teachers' language proficiency, classroom management, and self-efficacy. *International Journal of Foreign Language Teaching & Research*, 6(22), 89–102.
- Medgyes, P. (2001). When the teacher is a non-native speaker. In M. Celce-Murcia (Ed.), *Teaching English as a Second or Foreign Language* (3<sup>rd</sup> Ed.) (pp. 415–427). Boston, MA: Heinle & Heinle.
- Moradkhani, S., Raygan, A., & Moein, M. S. (2017). Iranian EFL teachers' reflective practices and self-efficacy: Exploring possible relationships. *System*, 65, 1–14. <https://doi.org/10.1016/j.system.2016.12.011>
- Murdoch, G. (1994). Language development provision in teacher training curricula. *ELT Journal*, 48(3), 253–265. <https://doi.org/10.1093/elt/48.3.253>
- \*Nishino, T. (2012). Modeling teacher beliefs and practices in context: A multimethods approach. *The Modern Language Journal*, 96(3), 380–399. <https://doi.org/10.1111/j.1540-4781.2012.01364.x>
- Norris, J.M., & Ortega, L. (2006). The value and practice of research synthesis for language learning and teaching. In J. M. Norris & L. Ortega (Eds.), *Synthesizing research on language learning* (pp. 3–50). Amsterdam, NE: John Benjamins Publishing.
- Pasternak, M. & Bailey, K.M. (2004). Preparing nonnative and native English-speaking teachers: Issues of professionalism and proficiency. In L.D. Kamhi-Stein (Ed.) *Learning and teaching from experience: Perspectives on nonnative English-speaking professionals* (pp. 155–175). Ann Arbor, MI: The University of Michigan Press.

- Plonsky, L. (2015). Statistical power, p values, descriptive statistics, and effect sizes: A “back-to-basics” approach to advancing quantitative methods in L2 research. In L. Plonsky (Ed.) *Advancing quantitative methods in second language research* (pp. 23 – 45). New York: Routledge.
- Plonsky, L., & Brown, D. (2015). Domain definition and search techniques in meta-analyses of L2 research (Or why 18 meta-analyses of feedback have different results). *Second Language Research*, 31, 267-278.
- Plonsky, L., & Derrick, D. J. (2016). A meta-analysis of reliability coefficients in second language research. *Modern Language Journal*, 100, 538-553.
- Plonsky, L., & Gonulal, T. (2015). Methodological synthesis in quantitative L2 research: A review of reviews and a case study of exploratory factor analysis. *Language Learning*, 65, 9-36.
- Plonsky, L., & Oswald, F. L. (2014). How big is “big”? Interpreting effect sizes in L2 research. *Language Learning*, 64(4), 878–912. <https://doi.org/10.1111/lang.12079>
- Plonsky, L., & Oswald, F. L. (2015). Meta-analyzing second language research. In L. Plonsky (Ed.), *Advancing quantitative methods in second language research* (pp. 106–178). New York: Routledge.
- Praver, M. (2014). *Japanese university English language teachers’ self-efficacy beliefs: A mixed methods approach* (Unpublished doctoral dissertation). Temple University, Philadelphia, PA.
- Rastegar, M., & Memarpour, S. (2009). The relationship between emotional intelligence and self-efficacy among Iranian EFL teachers. *System*, 37, 700–707. <https://doi.org/10.1016/j.system.2009.09.013>
- Reves, V. & Medgyes, P. (1994). The non-native English speaking EFL/ESL teacher’s self-image: An international survey. *System*, 22(3), 353–67. [https://doi.org/10.1016/0346-251X\(94\)90021-3](https://doi.org/10.1016/0346-251X(94)90021-3)
- Richards, J. C. (2010). Competence and performance in language teaching. *RELC Journal*, 41(2), 101–122. doi:10.1177/0033688210372953
- Richards, J. C. (2017). Teaching English through English: Proficiency, pedagogy and performance. *RELC Journal*, 48(1), 7–30. <https://doi.org/10.1177/0033688217690059>
- \*Sabokrouh, F., & Barimani-Varandi, S. (2013). The effect of EFL teachers’ attitude toward English language and English language proficiency on their sense of efficacy. *Journal of Advances in English Language Teaching*, 1(4), 117–125.
- Seidlhofer, B. (1999). Double standards: Teacher education in the expanding circle. *World Englishes*, 18(2), 233–245. <https://doi.org/10.1111/1467-971X.00136>
- \*Shim, J. W. (2001). *The teacher efficacy beliefs of Korean teachers of English as a foreign language* (Unpublished doctoral dissertation). The Ohio State University, Columbus, OH.



- \*Swanson, P. (2010a). Teacher efficacy and attrition: Helping students at introductory levels of language instruction appears critical. *Hispania*, 93(2), 305–321.
- \*Swanson, P. (2010b). Efficacy and language teacher attrition: A case for mentorship beyond the classroom. *NECTFL Review*, 66, 48–72.
- \*Swanson, P. (2012). Second/foreign language teacher efficacy and its relationship to professional attrition. *The Canadian Modern Language Review*, 68(1), 78–101. doi:10.3138/cmlr.68.1.078
- Swanson, P. (2014). The power of belief: Spanish teachers' sense of efficacy and student performance on the National Spanish Examinations. *Hispania*, 97(1), 5–20. <http://dx.doi.org/10.1353/hpn.2014.0015>
- \*Swanson, P., & Huff, R. (2010). Georgia's rural foreign language teachers' sense of efficacy and how it relates to teacher attrition. *World Languages and Cultures Faculty Publications*, 54.
- \*Takayama, H. (2015). *What influences teacher efficacy among English teachers in Japanese secondary schools?* (Unpublished doctoral dissertation). The University of Iowa, Iowa City, IA.
- \*Tayama, Y. (2011). What influences teacher efficacy among English teachers in Japanese secondary schools. *Japan Society of English Language Education*, 217–232. [https://doi.org/10.20581/arele.22.0\\_217](https://doi.org/10.20581/arele.22.0_217)
- \*Thompson, G. (2016). *Japanese high school English teachers' self-efficacy beliefs about teaching English*. (Unpublished doctoral dissertation). Queensland University of Technology, Brisbane, AU.
- Thorndike, R. L. (1949). *Personnel selection: Test and measurement techniques*. New York, NY: Wiley.
- Trofimovich, P., Isaacs, T., Kennedy, S., Saito, K., & Crowther, D. (2016). Flawed self-assessment: Investigating self- and other perception of second language speech. *Bilingualism: Language and Cognition*, 19(1), 122–140. <https://doi.org/10.1017/S1366728914000832>
- Tsang, A. (2017). EFL/ESL teachers' general language proficiency and learners' engagement. *RELC Journal*, 48(1), 99–113. <https://doi.org/10.1177/0033688217690060>
- Tschannen-Moran, M., & Woolfolk Hoy, A. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education*, 17, 783–805. [https://doi.org/10.1016/S0742-051X\(01\)00036-1](https://doi.org/10.1016/S0742-051X(01)00036-1)
- Tschannen-Moran, M., Woolfolk Hoy, A., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research*, 68(2), 202–248. <https://doi.org/10.3102/00346543068002202>
- Tsui (2003). *Understanding expertise in teaching: Case studies of second language teachers*. New York: Cambridge University Press.

- Uchihara, T., Webb, S., & Yanagisawa, A. (in press). The effects of repetition on incidental vocabulary learning: A meta-analysis of correlational studies. *Language Learning*.  
<https://doi.org/10.1111/lang.12343>
- Wyatt, M. (2010). An English teacher's developing self-efficacy beliefs in using groupwork. *System*, 38, 603–613. <https://doi.org/10.1016/j.system.2010.09.012>
- Wyatt, M. (2018). Language teachers' self-efficacy beliefs: A review of the literature (2005 – 2016). *Australian Journal of Teacher Education*, 43(4), 92–120.  
[doi:10.14221/ajte.2018v43n4.6](https://doi.org/10.14221/ajte.2018v43n4.6)
- \*Yilmaz, C. (2011). Teachers' perceptions of self-efficacy, English proficiency, and instructional strategies. *Social Behavior and Personality*, 39(1), 91–100.  
<https://doi.org/10.2224/sbp.2011.39.1.91>