Is learning really just believing? A meta-analysis of self-efficacy and achievement in SLA

Julia Goetze

Pennsylvania State University, State College, USA
https://orcid.org/0000-0002-4856-760X
jzg5860@psu.edu

Meagan Driver

Michigan State University, Lansing, USA
https://orcid.org/0000-0002-6712-4521
driverme@msu.edu

Abstract

The positive psychology movement (Seligman, 1998) has contributed to the proclamation of a positive turn in second language acquisition (SLA) (MacIntyre et al., 2016). Within the context of individual differences, *self-efficacy* (Bandura, 1997), an individual's judgment of their capability to achieve goals, has gained particular interest in language learning (e.g., Lake, 2013). The present study meta-analyzes a body of research that has investigated the relationship between second language (L2) self-efficacy and L2 achievement by exploring 1) reporting practices in this domain, 2) the strength and direction of the relationship, and 3) the effects of moderator variables on the self-efficacy-achievement link. A comprehensive literature search uncovered 37 studies, which contributed to a total of 40 independent samples (*N* = 23,050). The average observed effect in the sample was *r* = .46. A moderator analysis showed systematic variations in the effect size for learners' first language, target language, proficiency level, and both self-efficacy and achievement type. We discuss our findings with respect to theoretical constructs and methodological practices and suggest implications for L2 pedagogy and future research into self-efficacy in SLA.

Keywords: positive psychology; individual differences; L2 achievement; self-efficacy

1. Introduction

For nearly three decades, self-efficacy has seen a steady line of interest in SLA. In fact, since the positive turn in the field (MacIntyre et al., 2016; MacIntyre & Mercer, 2014), studies addressing positive affect in the foreign language (FL) classroom, including enjoyment (e.g., Dewaele & MacIntyre, 2014; Jin & Zhang, 2018; Zhang & Tsung, 2021), motivation (e.g., Dörnyei, 2020; Le-Thi et al., 2020) and second or foreign language (L2) grit (e.g., Alamer, 2021; Teimouri et al., 2020), have flourished and garnered increased attention with respect to language learning outcomes. As one of the most influential variables in positive psychology, self-efficacy has been explored across a range of contexts (e.g., target languages and proficiency levels), age groups (e.g., children, adolescents, and adults), language skills (e.g., reading, writing, listening and speaking), and classroom participants (e.g., students and teachers). Encouragingly, findings have shown the benefits of self-efficacy with respect to learners' L2 achievement (e.g., Ghonsooly & Elahi, 2010; Hetthong & Teo, 2013), learning strategies (e.g., Balci, 2017; Gahunga, 2009; Jee, 2015; Ma et al., 2018; Mizumoto, 2013; Wang et al., 2012), and attitudes towards L2 learning (e.g., Bai et al., 2020; Murad Sani & Zain, 2011).

Bandura's (1997) and Pajares's (1997) seminal works on self-efficacy proved to be powerful driving forces for scholars in SLA and have sparked an abundance of research into the relationship between self-efficacy and achievement. Still, the only research synthesis to date investigates self-efficacy and English proficiency only (Wang & Sun, 2020), leaving the broader contributions of research across all languages unexamined. Given the theoretical implications of self-efficacy for individual tasks (Bandura, 1997), the various linguistic contexts and learner populations that have been addressed, the plethora of existing self-efficacy and achievement instruments, and the often inconsistent reporting standards of psychometric properties of instruments in SLA (Larson-Hall & Plonsky, 2015), a comprehensive meta-analysis of the overall effect of self-efficacy on L2 achievement would be beneficial for furthering our understanding of self-efficacy as well as for developing theory and practice within the framework of positive psychology in SLA. The present study addresses this gap by means of a systematic meta-analysis in accordance with now commonly accepted methodologies in SLA (e.g., Norris & Ortega, 2006; Oswald & Plonsky, 2010; Plonsky & Oswald, 2012), including a moderator analysis of variables, such as study contexts, learner characteristics, and instrumentation.

2. Literature review

2.1. Positive psychology in SLA

Positive psychology investigates the traits and processes that allow people to grow and flourish (Seligman & Csikszentmihalyi, 2000) and has encouraged scholars to explore positive emotions and experiences, such as flow (Csikszentmihalyi, 2000) and resilience (Pan & Chan, 2007), which support people in leading full, happy, and emotionally stable lives. Many scholars in education have adopted methods from positive psychology to investigate a range of positive emotions with respect to learning processes and achievement (e.g., Pekrun et al., 2002), a trend that has produced a number of meta-analyses (e.g., Lei & Cui, 2016; Marques et al., 2017; Möller et al., 2009; Möller et al., 2020; Petscher, 2010). For example, in a series of path meta-analyses of 118 studies (N = 213,121), Möller and colleagues (2020) found strong effects for the relationship between self-concept – "a self-description judgement that includes an evaluation of competence and the feelings of self-worth associated with the judgement in question" in a specific field (Pajares & Schunk, 2005, p. 105) – and K-12 achievement in mathematics (β = .57) and first language (L1) (β = .46), respectively. In another study, Marques et al. (2017) explored the relationship between hope and academic achievement through a meta-analysis of 45 studies (N = 9,250) carried out with K-12, undergraduate and graduate students. Their findings revealed a moderate, positive relationship (k = 24, mean ρ = .24, SD = .10) (95% CI [.20, .26]), with a stronger link for K-12 students (k = 8, mean ρ = .28, SD = .10) (95% CI [.24, .32]) than for undergraduate and graduate students (k = 16, mean ρ = .19, SD = .06) (95% CI [.14, .23]). These suggested links between positive affect and academic achievement lay the foundation for exploring the association in content-specific contexts, such as the world language classroom (i.e., including but not limited to English as second and foreign language classrooms).

The positive turn in second language acquisition (SLA) (MacIntyre et al., 2016; MacIntyre & Mercer, 2014) has acknowledged the relevance of positive psychology for improving the FL learning experience by supporting learners' L2 motivation, perseverance, and resilience as well as by fostering meaningful communication and interaction between learners and teachers. As a result, this area of inquiry has seen an influx of research on enjoyment (e.g., Dewaele & Alfawzan, 2018), L2 grit (e.g., Alamer, 2021; Teimouri et al., 2020), and self-efficacy, as detailed in the following section. Notably, L2 motivation has probably garnered the most attention (e.g., Lake, 2013; Papi et al., 2019). As an example, Al-Hoorie (2018) explored the relationship between the L2 motivational self system (L2MSS; Dörnyei, 2005, 2009) – specifically the ideal L2 self, the ought-to L2

self, and the L2 learning experience – and L2 achievement through a meta-analysis of 39 samples (N = 32,078). The findings revealed that all three components of the L2MSS were significant predictors of intended effort (rs = .61, .38, and .41, respectively), the measure for subjective learning outcomes, though weaker correlations were found between motivation and L2 achievement (rs = .20, -.05, and .17). This study highlighted the need for meta-analyses that examine the effects of other positive affective variables on L2 achievement and how certain moderators might explain variability of findings.

2.2. Self-efficacy and language achievement

In perhaps one of the most cited definitions of the term, Bandura (1986) describes self-efficacy as "people's judgements of their capabilities to organize and execute courses of action required to attain designated types of performances" (p. 391). Simply put, it refers to individuals' beliefs that they hold the necessary skills to complete a particular task (Chao et al., 2019). These ability-related beliefs regulate achievement by influencing the goals a learner sets out to reach and the amount of effort dedicated to their performance (Bandura, 1986, 1997; Pajares, 1997). Beyond the regulation of achievement-related cognitive processes, self-efficacy has also been found to regulate types of affective feeling states in SLA. For example, multiple studies have shown that a perceived lack of competence was related to both weak self-efficacy beliefs and the presence of negative emotions, such as anxiety (e.g., Hiver, 2013; Song, 2016). In contrast, Cheng and associates (Cheng et al., 1999) found that perceived competence was related to strong self-efficacy beliefs and the presence of positive emotions, such as self-confidence. Worth noting is Wyatt's (2018) recent argument that self-confidence should not be treated as an emotional variable. Instead, self-confidence is characterized as a "lay term" (p. 122) for self-efficacy beliefs, thereby calling into question the nature of self-confidence and its relation to self-efficacy. Indeed, the majority of existing studies in SLA have adopted an emotional understanding of self-confidence and clearly differentiate it from cognitive self-efficacy beliefs. Thus, the present work addresses self-efficacy, specifically, and does not treat self-confidence as a synonymous term.

Numerous studies have shown a positive correlation between L2 self-efficacy and L2 achievement (e.g., Bai & Wang, 2020; Golparvar & Khafi, 2021; Karbakhsh & Safa, 2020), suggesting that learners with high levels of self-efficacy are more likely to persevere even during the most challenging tasks, thus achieving higher levels of language learning success (Duckworth, 2013). Many scholars in SLA have also explored skill-specific self-efficacy, including reading (e.g., Al Khamisi et al., 2016; Balci, 2017; Chuang et al., 2018; Kitikanan & Sasimonton,

2017), writing (e.g., Raoofi & Maroofi, 2017; Sahril & Weda, 2018), speaking (e.g., Asakereh & Dehghannezhad, 2015; Kitikanan & Sasimonton, 2017) and listening (e.g., Kitikanan & Sasimonton, 2017; Phakiti et al., 2013), as well as content-specific self-efficacy for grammar (e.g., Mustapha et al., 2013) and vocabulary learning (e.g., Wu et al., 2012; Wu et al., 2013). For example, Al Khamisi et al. (2016) explored the relationship between reading self-efficacy and reading achievement for 636 primary and secondary English as a foreign language (EFL) students. Correlations showed that learners with higher self-efficacy had greater success on reading tests at both academic levels (elementary: $r = .541$, $p < .001$; secondary: $r = .518$, $p < .001$). Similarly, Kitikanan and Sasimonton (2017) investigated skill-specific self-efficacy (i.e., listening, speaking, reading, and writing) and achievement for 32 Thai EFL university students, finding strong positive correlations between self-efficacy for each skill and overall achievement ($r_{average} = .615$; 95% CI [.34; .79]). The authors thus proposed that developing self-efficacy in even just one skill might improve learners' overall L2 achievement.

Despite this general positive trend, some researchers of skill-specific self-efficacy and L2 achievement have observed merely weak relationships between the two variables. For example, Sahril and Weda (2018) examined writing self-efficacy and achievement with 50 Indonesian EFL university students and found a weak positive correlation ($r = .057$, $p < .001$). Similarly, Liem and associates (2008) found a small positive correlation ($r = .180$, $p < .001$) when investigating general L2 self-efficacy and achievement in a group of 1,475 high school EFL learners in Singapore. In a study of general self-efficacy and speaking achievement, Oliver et al. (2005) also reported a very small positive correlation ($r = .09$, $p > .05$) for 275 Australian elementary school students studying a range of foreign languages.

While all of these studies provide important insights, the variation of effect sizes across studies is noteworthy and motivates the current study, which aims to contribute to a more comprehensive and nuanced understanding of the relationship between L2 self-efficacy and L2 achievement in SLA.

3. The current study

The current study is motivated by three goals. First, we perform a thorough and systematic literature search and analysis to examine existing trends in reporting practices in the L2 self-efficacy and achievement domain. Second, we report on both the direction and size of the relationship between learner self-efficacy and achievement. Lastly, we examine our sample for systematic variation of effects and investigate a number of moderator variables. In order to meet these goals, the present meta-analysis addresses the following research questions:

1) What reporting practices are used in studies exploring L2 self-efficacy and L2 achievement? (RQ1)
2) What is the direction and magnitude of the relationship between L2 self-efficacy and L2 achievement? (RQ2)
3) What is the moderating effect of L1/L2 proficiency level, learner age, institutional context (e.g., primary school, university), self-efficacy type (e.g., general, skill-specific), and achievement measure on the relationship between L2 self-efficacy and L2 achievement? (RQ3)

4. Methods

4.1. Study identification and inclusion criteria

In order to identify relevant studies to answer our research questions, a set of inclusion criteria was applied. To be eligible, a study had to (1) present a measure of both L2 self-efficacy and L2 achievement, (2) report quantitative results, either as a Pearson correlation or a statistic that can be converted into an r index (e.g., t or F), and (3) be published in or after 1997, to include the studies published since both Bandura's (1997) and Pajares's (1997) influential works and the emergence of positive psychology.

Using the defined parameters and combinations of defined keywords in English: (1) *self-efficacy, beliefs, self-esteem, mindset, self-concept, talent*; (2) *performance, achievement, outcomes*; and (3) *language learning, second language learning, foreign language learning, heritage language learning*, we conducted a comprehensive search in two library-housed databases (ERIC, LLBA), and one public database (Google Scholar). We also conducted an ancestry search by reviewing the bibliographies of relevant studies and the publication lists of prominent researchers in self-efficacy and positive psychology in SLA, identifying a total of 640 publications.

Some studies were later excluded due to: (1) an absence of our target variables, (2) research designs that included neither L2 self-efficacy nor achievement instruments, (3) a qualitative methods or case study design, (4) missing data, such as correlations, means, and standard deviations, and (5) a self-efficacy construct unrelated to language learning, such as computer self-efficacy (Ale et al., 2017) or technological self-efficacy (Abdallah & Mansour, 2015), which measure beliefs regarding one's capability to use technology or digital tools in order to achieve one's goal. After applying all eligibility criteria, 37 studies published between 1999 and 2019 in 15 different countries were included in the final analysis, with a total N of 23,050 and sample sizes ranging from 32 to 11,036 ($M = 581.9$; $SD = 1770.8$).

4.2. Coding procedures

Each study was coded for a number of features according to a systematic coding scheme, which was adapted from an existing coding scheme of a recent meta-analysis of L2 anxiety (Teimouri et al., 2019). Specifically, we recorded study features according to five categories: (1) bibliographic details (e.g., authors), (2) study design (e.g., methods), (3) participant sample (e.g., target language), (4) instruments (e.g., self-efficacy scale), and (5) quantitative findings (e.g., effect sizes). To pilot the coding scheme, we each rated five studies from the sample and resolved any ambiguities in our results and coding scheme. The final coding scheme can be found in Appendix A. Next, each author coded half of the studies. Twenty percent of the sample was coded by both authors, establishing high inter-rater reliability (100%).

4.3. Analysis

To answer RQ1, we first identified all self-efficacy and achievement instruments. Next, we conducted frequency counts and examined the reporting tendencies for reliability and validity of instruments for each variable. By examining the psychometric properties, we followed the foundational principle of the methodological reform movement (Marsden & Plonsky, 2018), which aims for more rigorous reporting practices in the field of SLA through robust description and evaluation procedures.

To calculate the magnitude and direction of the relationship between self-efficacy and achievement (RQ2), we employed a fixed-effect model. That is, we assumed the existence of one common effect size across all studies in our sample and calculated both the weighted overall mean according to each study's sample size, and the 95% confidence intervals for each effect. The fixed-effect model was chosen a priori due to our small sample size. More specifically, after visually inspecting our coded studies, we assumed only a moderate amount of heterogeneity across our sample. That is, while the studies were not identical with regards to design features, such as their use of data collection instrument or participant recruitment procedure, we deemed them moderately homogenous in terms of participant characteristics, such as target language and institutional context. This assumption was later confirmed ($I^2 = 68.647$, $p < .001$), and we investigated the systematic variation of effect sizes within our model (Borenstein et al., 2010) in our moderator analysis (RQ3).

To address our last research question, we used the categorical variables in our coding scheme to form and analyze subsamples. Before running the analyses, we examined our data set for heterogeneity, calculating the goodness-of-fit statistic (Q) and checking for the existence of systematic variation among the observed effects in the sample (I^2). All analyses were performed using *Comprehensive Meta-Analysis* (CMA) Version 3 (Borenstein et al., 2013).

5. Results

The results concerning our first research question, which aimed to explore reporting practices, are summarized in Tables 1 and 2. Self-efficacy was measured by a total of 33 unique instruments, with two studies that did not provide details on their instrument. The most commonly referenced questionnaires were the *Motivated Strategies for Learning Questionnaire* (MSLQ; Pintrich et al., 1991), the *Questionnaire of English Self-Efficacy* (QESE; Wang, 2004; Wang et al., 2013), and the *Self-Efficacy Questionnaire* (SEQ; Sedighi et al., 2004). However, only a minority of our sample (k = 14; 36%) adapted any of these three instruments. Fifteen studies uniquely adapted instruments from other sources, and eight studies designed new instruments. Of the 39 times that studies reported on a measurement for L2 self-efficacy, reliability was reported 37 times (95%) in the form of Cronbach's alpha coefficients, with the exception of two studies that, respectively, conducted a test-retest correlation and split-half test of internal consistency.

With respect to language achievement measures, our analysis found 39 different measures representing three main methods for assessing L2 achievement: language tests (k = 22), course grades (k = 7), and Grade Point Average (GPA) (k = 3). Seven samples implemented a range of other methods, including student self-reported exam grades, teacher ratings of student achievement, and task-based assessment (Table 2). As shown in Table 2, 28 samples (72%) reported means, and 26 studies (66%) reported both the means and standard deviation for their instrument. Of the 22 samples implementing a language test, reliability coefficients were reported only eight times (36%): five reported Cronbach's alpha, one reported KR-20, one reported inter-rater reliability, and one reported test-retest correlation. The five samples reporting Cronbach's alpha yielded an average reliability of .93 (SD = .03).

Table 1 Reliability analyses (Cronbach's alpha) of L2 self-efficacy instruments

Instrument	K	# reporting reliability	# not reporting reliability	Min.	Max.	*Mean*	*SD*
MSLQ	6	6	0	.85	.97	.91	.05
QESE	5	4	1	.84	.96	.92	.06
SEQ	3	3	0	.74	.87	.82	.07
Novel instrument	8	8	0	.78	.95	.89	.06
Other	15	15	0	.63	.96	.86	.09
All SE instruments	39	37	2	.63	.97	.88	.08

Note. MSLQ – *Motivated Strategies for Learning Questionnaire* (Pintrich et al., 1991); QESE – *Questionnaire of English Self-Efficacy* (Wang, 2004; Wang et al., 2013); SEQ – *Self-Efficacy Questionnaire* (Sedighi et al., 2004)

Table 2 Reporting practices for L2 achievement assessments

Assessment type	K # reporting Mean		# reporting SD	# reporting reliability	Cronbach's alpha	
					Mean	*SD*
Language test	22	17	16	8	.93	.03
Course grade	7	7	7	0	-	-
GPA	3	3	3	0	-	-
Other	7	3	1	0	-	-
All achievement measures	39	28	26	8	.93	.03

The second research question targeted the magnitude and direction of the relationship between self-efficacy and achievement. We first computed the weighted average mean of all 40 effects in our sample: $r = .464$ (95% CI [.454, .474]; $p < .001$), and then examined their distribution for outliers, which we define as any result with a standardized residual larger than 3 in absolute value. This resulted in the exclusion of 13 values, which increased our result slightly to $r = .475$ (95% CI [.464, .486]; $p < .001$). Additionally, the Q statistic changed from 1303.27 to 82.93, and the I^2 statistic changed from 97.01 to 68.65. This indicated a reduction of sampling error and a sizeable amount of real and systematic variation in the data, thus justifying a moderator analysis. Table 3 provides a summary of our statistical model. An overview of all effects can be found in Appendix B.

Table 3 Overall correlation between L2 self-efficacy and L2 achievement

	K	N	M_r	95% CI		Q	I^2
				Lower	Upper		
L2 achievement	27	18,265	.475	.464	.486	82.927*	68.647*

Note. * = $p < .001$

Additionally, a funnel plot of the relationship between effect size and standard error (Figure 1) was created to examine the presence of publication bias in our sample. The plot includes both actual (i.e., hollow) and imputed (i.e., solid black) data points and suggests a slight publication bias in favor of studies that report strong, positive correlations between L2 self-efficacy and achievement.

The moderator analysis (Table 4) was conducted with 27 independent samples and calculated subgroup effects based on seven categories: age, learner L1, target language, institutional context, proficiency level, self-efficacy type, and achievement measure. The analysis for age showed weaker correlations for high-school age teenagers (16-18) and learners over 20 than for early high-school (15) and early college age students (19-20). However, the majority of primary studies did not report age as a learner characteristic, leading to small subsamples that warrant cautious interpretation.

Figure 1 Funnel plot of effect sizes and sampling errors
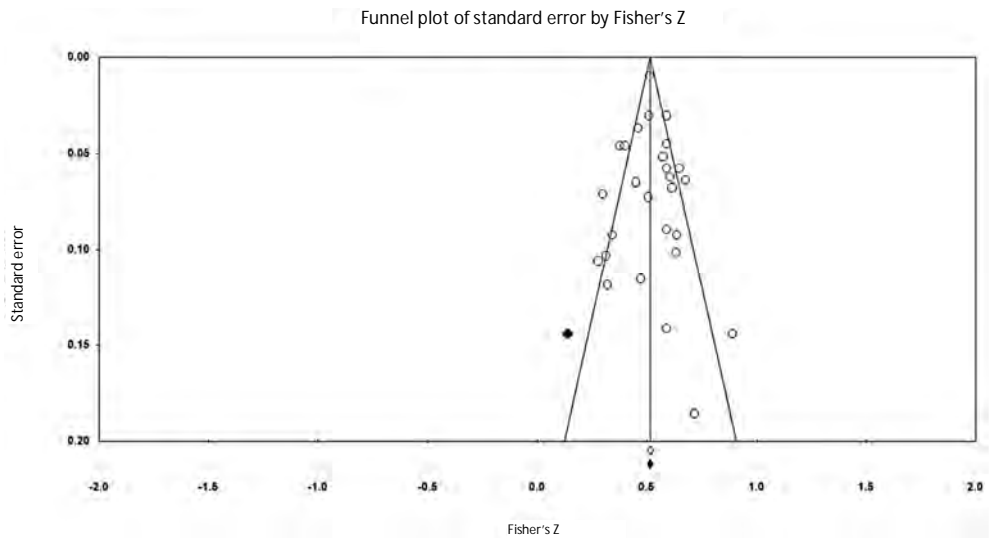
Note. Solid black point represents imputed data

Figure 1 Funnel plot of effect sizes and sampling errors

With respect to the learners' L1 and proficiency level, a similar pattern of underreporting was found. Regarding learners' L1, some languages, such as Thai ($r = .68$), showed stronger relationships between L2 self-efficacy and achievement than others, for example Japanese ($r = .31$). In terms of proficiency level, the analysis indicated a gradual increase in relationship strength when moving from beginner ($r = .44$) to advanced levels ($r = .62$).

The analysis of self-efficacy type revealed the strongest relationship for speaking self-efficacy ($r = .56$) and the weakest relationship for vocabulary self-efficacy ($r = .33$). Other skill-specific types (i.e., reading and writing self-efficacy), as well as general L2 self-efficacy (i.e., no skill or content-specific subtype), all showed findings similar to the overall mean effect of $r = .475$. Regarding achievement measure, learners' GPA yielded a stronger correlation ($r = .62$) than course grades or language tests, which both returned the same effect ($r = .47$).

The analyses for other moderators did not indicate such ranges in effect sizes. For example, findings for target language revealed only a slightly higher effect for English ($r = .53$) than for non-English languages ($r = .47$), and institutional contexts showed similar effects across all subgroups ($r = .47\text{-}.50$).

Table 4 Results of the moderator analysis

	K	N	M_r	95% CI	
				Lower	Upper
Age					
15	2	2,184	.501	.468	.531
16	1	200	.290	.158	.412
18	1	78	.440	.241	.603
19	1	53	.530	.303	.700
20	2	349	.582	.507	.647
21	6	1,594	.408	.366	.448
Not reported	14	13,807	.477	.465	.490
L1					
Arabic	3	857	.533	.483	.580
Chinese	2	1,212	.457	.411	.501
English	1	128	.530	.393	.644
Japanese	2	171	.307	.163	.438
Malaysian	2	320	.401	.304	.489
Persian	1	100	.560	.409	.681
Thai	2	83	.678	.539	.781
Multiple	1	304	.530	.444	.606
Not reported	13	15,090	.473	.460	.485
Target language					
English	23	17,296	.472	.460	.483
Non-English	4	969	.525	.478	.570
Institutional context					
Secondary	7	13,806	.474	.461	.487
Post-secondary	17	3,583	.471	.447	.495
Language institute	3	876	.500	.449	.548
Proficiency level					
Beginner	1	78	.440	.241	.603
Intermediate	4	696	.518	.461	.570
Advanced	1	32	.615	.339	.794
Multiple	5	842	.487	.433	.537
Not reported	16	16,617	.472	.461	.484
Self-efficacy type					
General	17	15,943	.474	.462	.486
Skill – vocabulary	1	120	.330	.160	.481
Skill – reading	4	889	.479	.427	.529
Skill – speaking	1	100	.560	.409	.681
Skill – writing	4	1,213	.483	.439	.525
Achievement measure					
Language test	17	16,677	.473	.461	.485
Course grade	5	783	.473	.417	.526
GPA	1	32	.615	.339	.794
Other	2	220	.442	.328	.543
Multiple	2	553	.530	.467	.588

6. Discussion

RQ1 explored reporting practices in L2 self-efficacy and achievement research. The majority of studies reported both basic descriptive statistics and reliability coefficients of their self-efficacy instruments, which all align with or exceed the average reliability estimates for instruments designed for SLA research (Plonsky & Derrick, 2016). Thus, we see rigorous reporting practices within this domain of research. With respect to language achievement measures, we found a weaker tradition of reporting practices. Only approximately one-third of the studies provided the reliability coefficient of their language achievement test. In terms of basic descriptives, only two-thirds reported both the mean and standard deviation for their sample. As noted in previous meta-analyses (e.g., Teimouri et al., 2019) and reviews of SLA research practices (e.g., Larson-Hall & Plonsky, 2015; Plonsky, 2013, 2017), this lack of reporting is a common limitation of SLA research. Furthermore, nearly half of the studies used instructor- or institution-developed achievement tools, which often lack the psychometric properties that are essential for robust quantitative analysis in SLA (Brown et al., 2018). Altogether, these results highlight the need for greater validation of instruments and more rigorous reporting practices on language achievement measures in SLA (Marsden & Plonsky, 2018).

  RQ2 investigated the direction and magnitude of the relationship between learners' self-efficacy and their L2 achievement. The mean correlation between these variables was $r = .475$, and self-efficacy accounted for approximately 22% of the variance in learners' achievement, signifying a medium effect in SLA (Plonsky & Oswald, 2014). Within the context of the positive turn in SLA, it is insightful to interpret our result in comparison to meta-analytical findings of other prominent psychological constructs, such as motivation, which has been paradigmatically categorized as playing a positive role in student learning. Somewhat surprisingly, Al-Hoorie's (2018) recent meta-analysis of the L2 motivational self system and L2 achievement found smaller mean correlations between learners' L2 achievement and their L2 ideal self ($r = .20$), ought-to self ($r = -.05$), and learning experience ($r = .17$). Considering the close relationship between learners' ideal selves and self-efficacy beliefs (Lake, 2013), one might have expected a similar effect size for both variables. Instead, finding a larger positive effect for learner beliefs raises questions about the complexity of the relationship between learner beliefs and motivations in the context of L2 achievement. While studies in the positive psychology paradigm in SLA have predominantly examined a host of emotional variables (Driver, 2021), our findings furthermore highlight the potential of examining the role of cognitive variables and their interaction in student thriving and learning success in future studies.

It is also helpful to interpret our result in comparison to meta-analytic findings in neighboring disciplines. In psychology, Multon and associates (1991) found a more moderate effect of $r = .38$ (95% CI [.36, .41]) through their meta-analysis of self-efficacy beliefs and general academic outcomes, while Huang's (2016) findings in education regarding the effect of self-efficacy on general achievement goals align more closely with our result ($r = .48$; 95% CI [.38, .46]). The similarity of Huang's finding raises questions regarding the underlying theoretical assumptions of the self-efficacy and achievement relationship in SLA (Wyatt, 2018), considering Huang's conceptualization of *achievement goals* as beliefs about achievement objectives rather than as a measurable performance outcome. Put differently, Huang's study examined the relationship between two separate beliefs, rather than between a belief (e.g., self-efficacy) and an outcome (e.g., achievement measure), arguably capturing more accurately Bandura's (1977) theorization of the relationship between self-efficacy and achievement (Figure 2). In contrast, most studies in our sample employed self-efficacy definitions by Bandura (1986, 1997) but used correlational designs to investigate the link to outcomes, thereby conflating multiple cognitive and behavioral constructs into one variable. Therefore, our estimate of the mean effect must be interpreted with caution, as neither the different types of beliefs nor the behaviors of participants were controlled for in the primary studies.
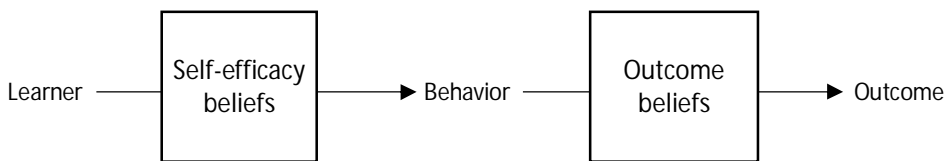

Figure 2 Representation of the self-efficacy and achievement relationship (Bandura, 1977)

RQ3 examined the effects of moderating variables. The results of this analysis showed a number of differences in effects sizes with respect to learner characteristics, language, and target skill (e.g., speaking vs. writing). However, considering the overall small sample size of this moderator analysis ($N = 27$) and the resulting limited number of studies in each analytical subcategory, caution should be taken when interpreting these findings.

Perhaps most notable were the findings for L2 proficiency, which showed strongest effects for advanced students and weakest for beginners. For self-efficacy to positively impact L2 achievement, learners must not only believe in their abilities, but these beliefs must also have a direct effect on how learners engage with a task, that is, through increased levels of attention and awareness, implementation of task-appropriate learning strategies, or stronger feelings of L2 grit,

all of which have been seen to result in better learning outcomes (e.g., Teimouri et al., 2020). Higher proficiency learners likely have more experience with how positive beliefs can motivate learning processes that lead to greater L2 achievement, as well as the linguistic resources to bridge the gap between simply believing in the ability to complete a task and actually successfully accomplishing the task.

Results also revealed differences across learners' L1, which may be related to questions around L2 accessibility and the relative "prestige" and utility of the L1 in global settings. Self-efficacy effects were strongest for L1 Thai speakers and weakest for L1 Japanese speakers. Given the value of Japanese in the worldwide business market and the socioeconomic developments in Japan (Terasawa, 2017) that have increased access to the L2 (i.e., English), L1 Japanese speakers may be motivated by factors other than self-efficacy and have less urgency to meet achievement objectives than other L1 speakers. In contrast, L1 Thai speakers encounter fewer opportunities and encouragement to use L2 English in Thailand (Anyadubalu, 2010) but, at the same time, may also be more dependent on an L2 in order to enter into global discussions, economics, and business partnerships. Thus, some L1 speakers in regions with less access to L2 resources are likely to rely more heavily on their self-efficacy for affecting achievement, which aligns with the relative effects and L1 profiles in our findings (e.g., Arabic; Aljaffery, 2015).

Similarly, given the "prestige" of English as a global language, L1 English speakers may also receive less support for learning an L2 and rely more heavily on their own self-beliefs for realizing L2 achievement goals. In addition, we found slightly weaker effects for L2 learners of English than for non-English languages, which strengthens our argument. L2 learners of English likely find primary motivations (e.g., ought-to L2 self; Ushioda & Dörnyei, 2011) from outside sources based on the global instrumentality of the English language. In contrast, learners of other languages likely see fewer motivations and obligations from their society to learn the L2, resulting in a stronger connection between self-efficacy and achievement.

Interestingly, results also indicated weaker effects for learners' beliefs about their L2 vocabulary knowledge than general or skill-specific self-efficacy (e.g., reading, writing, or speaking), suggesting that learners' beliefs in their ability to complete open-ended rather than narrowly focused tasks (i.e., on vocabulary or grammar) may have a greater impact on achievement. The strongest effect was found for speaking self-efficacy. Because speaking requires learners to produce language without the opportunity to review and revise before the final product (as is the case with both reading and writing), self-efficacy in this area suggests that learners have a strong belief in their ability to spontaneously and accurately organize and communicate their thoughts. Thus, speaking self-efficacy may encompass self-beliefs in a wider number of task-related abilities, leading to stronger effects for ultimate L2 achievement.

Finally, the positive effects of self-efficacy were uniform in relation to language tests and course grades, which suggests both as acceptable metrics of achievement in this domain. However, the much larger effect for GPA raises concerns about the weak ecological validity of GPA as a measure of achievement. Still, the subsample for this moderator was small and this finding should be taken with caution. Other moderators revealed either no differences (i.e., institutional context) or no clear variation patterns (i.e., age). Though a smaller effect size was seen for the 16-year-old age group, the self-efficacy-achievement connection for 15-year-olds was similar to the average effect size for college-aged learners (i.e., ages 18-21), and, due to underreporting of this learner characteristic, it is not clear if studies explored other age groups within the secondary or primary school context. The question whether or not the effects of L2 self-efficacy may be more beneficial for adults in comparison to younger learners, or perhaps for contextually younger students compared to older students, who tend to be closer to a graduation and perhaps reap fewer benefits of believing in their abilities, will require more research exploring a wider range of age groups, as well as robust reporting practices.

7. Limitations, directions for further research, and pedagogical implications

Some important limitations need to be considered when interpreting the findings of our study. First, small subsamples in our moderator analysis can only provide scant evidence for the role that our selected variables play in the systematic variation of the effect size. Second, we did not include unpublished studies or those published in languages other than English in our final sample, which may have introduced multiple biases into our analysis that possibly resulted in an overestimation of the strength of the observed relationship (Cooper, 2016). Lastly, we also excluded two studies that implemented advanced statistical methods, such as β-coefficients in structural equation modeling (SEM) to report on the relationship for self-efficacy and achievement. As a result, we are not able to provide evidence for existing trends or effect size variation for different analytical approaches.

We encourage future studies to carefully examine the theoretical underpinnings of the self-efficacy and achievement relationship. Bandura's (1986) frequently used definition distinguishes self-efficacy beliefs from beliefs about outcomes, which some scholars have defined as self-confidence (e.g., Karademas, 2006), thereby adding not only a behavioral, but also an emotional component to the theoretical conceptualization of the relationship. The use of conceptual frameworks that allow for the holistic examination of self-efficacy in relation to other psychological dimensions, such as emotions, behaviors, and motivations, simultaneously within cultural, societal, and institutional contexts (e.g., Ushioda,

2014) would move future studies away from cross-sectional bivariate analyses of variables toward system modeling that is able to capture variable change over time. The moderator analysis also revealed the need for research with populations learning other second or foreign languages than English and with K-12 learners. Additionally, our findings suggest that SLA would benefit from future studies that focus specifically on skill-specific self-efficacy beliefs. Lastly, we recommend more rigorous reporting practices in line with field-specific standards and reaffirm calls to improve the reporting of statistical and psychometric features of measurement tools, especially for L2 achievement.

In terms of FL pedagogy, the findings suggest that increased self-efficacy in the target language is beneficial for language learning, regardless of age or context. Given this positive relationship, educators might consider using methods that have been shown to promote self-efficacy, such as the flipped classroom model (Namaziandost & Çakmak, 2020), and project-based learning (Shin, 2018). Both of these methods have shown positive effects on learners' self-efficacy and motivation, and may have a positive influence on their future achievement objectives. Educators should consider preparing classroom activities, particularly relating to speaking, that feed into students' self-efficacy and nurture students' beliefs about their abilities to complete tasks in the L2, which will depend on learners' experiences with the language and their existing knowledge and skills. As our findings suggest, practitioners may find that materials designed to raise L2 self-efficacy are particularly beneficial for students at higher L2 levels. Importantly, the methods used in a majority of studies in our sample were not intended to replicate the classroom learning experience, and pedagogical implications should be considered with other well-founded pedagogical methods from SLA.

8. Conclusion

The goal of this meta-analysis was to systematically examine the relationship between L2 self-efficacy and achievement. The findings indicate that self-efficacy has a medium-sized, positive relationship with L2 achievement, and suggest that learning success may indeed be a question of believing in one's abilities during the learning process. Furthermore, the findings raise important questions about the theoretical nature of self-efficacy and the role of learner characteristics and contexts, and highlight the need for greater methodological rigor, all of which can serve as starting points for forthcoming empirical research into self-efficacy and achievement in the field of SLA, particularly within the framework of positive psychology.

References[1]

Abdallah, M., & Mansour, M. M. (2015). Virtual task-based situated language-learning with *Second Life*: Developing EFL pragmatic writing and technological self-efficacy. *Arab World English Journal (AWEJ)*, *Special Issue on CALL* (2), 150-182. https://doi.org/10.2139/ssrn.2843987

Al-Hoorie, A. H. (2018). The L2 motivational self system: A meta-analysis. *Studies in Second Language Learning and Teaching*, *8*(4), 721-754. https://doi.org/10.14746/ssllt.2018.8.4.2

Alamer, A. (2021). Grit and language learning: Construct validation of L2-Grit scale and its relation to later vocabulary knowledge. *Educational Psychology*, *41*(5), 544-562. https://doi.org/10.1080/01443410.2020.1867076

Ale, K., Loh, Y. A., & Chib, A. (2017). Contextualized-OLPC education project in rural India: Measuring learning impact and mediation of computer self-efficacy. *Educational Technology Research and Development*, *65*(3), 769-94. https://doi.org/10.1007/s11423-017-9517-2

Aljaffery, Y. (2015). *The influence of English language in the Arab world*. GRIN Verlag.

*Al Khamisi, H., Al Barwani, T., Al Mekhlafi, A., & Osman, M. E. T. (2016). EFL reading achievement: Impact of gender and self-efficacy beliefs. *International Journal of Learning, Teaching and Educational Research*, *15*(3), 54-73.

*Alrabai, F. (2018). The association between self-efficacy of Saudi learners and their EFL academic performance. *Theory and Practice in Language Studies, 8*(10), 1351-1360. https://doi.org/10.17507/tpls.0810.14

Anyadubalu, C. C. (2010). Self-efficacy, anxiety, and performance in the English language among middle-school students in English language program in Satri Si Suriyothai School, Bangkok. *International Journal of Human and Social Sciences*, *5*(3), 193-198.

*Asakereh, A., & Dehghannezhad, M. (2015). Student satisfaction with EFL speaking classes: Relating speaking self-efficacy and skills achievement. *Issues in Educational Research*, *25*(4), 345-363.

*Bai, B., Chao, G. C. N., & Wang, C. (2019). The relationship between social support, self-efficacy, and English language learning achievement in Hong Kong. *TESOL Quarterly, 53*(1), 208-221. https://doi.org/10.1002/tesq.439

Bai, B., Nie, Y., & Lee, A. N. (2020). Academic self-efficacy, task importance and interest: relations with English language learning in an Asian context. *Journal of Multilingual and Multicultural Development*. https://doi.org/10.1080/01434632.2020.1746317

---

[1] Starred references mark studies included in our meta-analysis.

Bai, B., & Wang, J. (2020). The role of growth mindset, self-efficacy and intrinsic value in self-regulated learning and English language learning achievements. *Language Teaching Research.* https://doi.org/10.1177/1362168820933190

*Balci, Ö. (2017). The effects of learning-style based activities on students' reading comprehension skills and self-efficacy perceptions in English foreign language classes. *Higher Education Studies, 7*(4), 35-54. https://doi.org/10.5539/hes.v7n4p35

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191-215. https://doi.org/10.1037/0033-295X.84.2.191

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Prentice Hall.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* Freeman.

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods, 1*(2), 97-111. https://doi.org/10.1002/jrsm.12

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2013). *Comprehensive meta-analysis version 3.* Biostat.

Brown, A. V., Plonsky, L., & Teimouri, Y. (2018). The use of course grades as metrics in L2 research: A systematic review. *Foreign Language Annals, 51*(4), 763-778. https://doi.org/10.1111/flan.12370

*Chao, C. N. G., McInerney, D. M., & Bai, B. (2019). Self-efficacy and self-concept as predictors of language learning achievements in an Asian bilingual context. *Asia-Pacific Education Researcher, 28*(2), 139-147. https://doi.org/10.1007/s40299-018-0420-3

Cheng, Y. S., Horwitz, E. K., & Schallert, D. L. (1999). Language anxiety: Differentiating writing and speaking components. *Language Learning, 49*(3), 417-446. https://doi.org/10.1111/0023-8333.00095

Chuang, H., Weng, C., & Chen, C. (2018). Which students benefit most from a flipped classroom approach to language learning? *British Journal of Educational Technology, 49*(1), 56-68. https://doi.org/10.1111/bjet.12530

Cooper, H. (2016). *Research synthesis and meta-analysis* (5th ed.). Sage.

Csikszentmihalyi, M. (2000). The contribution of flow to positive psychology. In J. E. Gillham (Ed.), *The science of optimism and hope: Research essays in honor of Martin E. P. Seligman* (pp. 387-395). Templeton Foundation Press.

Dewaele, J. M., & Alfawzan, M. (2018). Does the effect of enjoyment outweigh that of anxiety in foreign language performance? *Studies in Second Language Learning and Teaching, 8*(1), 21-45. https://doi.org/10.14746/ssllt.2018.8.1.2

Dewaele, J. M., & MacIntyre, P. D. (2014). The two faces of Janus? Anxiety and enjoyment in the foreign language classroom. *Studies in Second Language Learning and Teaching, 4*(2), 237-274. https://doi.org/10.14746/ ssllt.2014.4.2.5

*Doordinejad, F. G., & Afshar, H. (2014). On the relationship between self-efficacy and English achievement among Iranian third grade high school students. *International Journal of Language Learning and Applied Linguistics World*, *6*(4), 461-470.

Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Lawrence Erlbaum.

Dörnyei, Z. (2009). The L2 motivational self system. *Motivation, language identity and the L2 self*, *36*(3), 9-11. https://doi.org/10.21832/9781847691293-003

Dörnyei, Z. (2020). *Innovations and challenges in language learning motivation*. Routledge.

Driver, M. (2021). Emotion-laden texts and words: The influence of emotion on vocabulary learning for heritage and foreign language learners. *Studies in Second Language Acquisition*. https://doi.org/10.1017/S0272263121000851

Duckworth, A. L. (2013). *Grit: The power of passion and perseverance.* Scribner.

*Erkan, D. Y., & Saban, A. İ. (2011). Writing performance relative to writing apprehension, self-efficacy in writing, and attitudes towards writing: A correlational study in Turkish tertiary-level EFL. *The Asian EFL Journal Quarterly, 13*(1), 164-192.

*Fukuda, A. (2018). The Japanese EFL learners' self-regulated language learning and proficiency. *Journal of Pan-Pacific Association of Applied Linguistics*, *22*(1), 65-87. https://doi.org/10.25256/PAAL.22.1.4

*Gahunga, O. (2009). Are self-efficacy, language learning strategies, and foreign language ability interrelated? *The Buckingham Journal of Language and Linguistics, 2*, 47-60.

*Ghonsooly, B., & Elahi, M. (2010). Learners' self-efficacy in reading and its relation to foreign language reading anxiety and reading achievement. *Amirkabir, 53*(217), 45-67.

Golparvar, S. E., & Khafi, A. (2021). The role of L2 writing self-efficacy in integrated writing strategy use and performance. *Assessing Writing*, *47*, 100504. https://doi.org/10.1016/j.asw.2020.100504

*Hetthong, R., & Teo, A. (2013). Does writing self-efficacy correlate with and predict writing performance? *International Journal of Applied Linguistics and English Literature, 2*(1), 157-167. https://doi.org/10.7575/ijalel.v.2n.1p.157

Ghonsooly, B., & Elahi, M. (2010). Learners' self-efficacy in reading and its relation to foreign language reading anxiety and reading achievement. *Journal of English Language Teaching and Learning*, *53*(217), 45-67.

Hiver, P. (2013). The interplay of possible language teacher selves in professional development choices. *Language Teaching Research*, *17*(2), 210-227. https://doi.org/10.1177%2F1362168813475944

*Hsieh, P. (2008). "Why are college foreign language students' self-efficacy, attitude, and motivation so different?" *International Education*, *38*(1), 76-94.

*Hsieh, P. P., & Kang, H. (2010). Attribution and self-efficacy and their interrelationship in the Korean EFL context. *Language Learning*, *60*(3), 606-627. https://doi.org/10.1111/j.1467-9922.2010.00570.x

*Hsieh, P. P., & Schallert, D. L. (2008). Implications from self-efficacy and attribution theories for an understanding of undergraduates' motivation in a foreign language course. *Contemporary Educational Psychology*, *33*(4), 513-532. https://doi.org/10.1016/j.cedpsych.2008.01.003

Huang, C. (2016). Achievement goals and self-efficacy: A meta-analysis. *Educational Research Review*, *19*, 119-137. https://doi.org/10.1016/j.edurev.2016.07.002

*Jee, M. J. (2015). Language learners' strategy use and self-efficacy: Korean heritage learners versus non-heritage learners. *Language Research, 51*(1), 167-195.

Jin, Y., & Zhang, L. J. (2018). The dimensions of foreign language classroom enjoyment and their effect on foreign language achievement. *International Journal of Bilingual Education and Bilingualism.* https://doi.org/10.1080/13670050.2018.1526253

Karademas, E. C. (2006). Self-efficacy, social support and well-being: The mediating role of optimism. *Personality and Individual Differences*, *40*, 1281-1290.

Karbakhsh, R., & Safa, M. A. (2020). Basic psychological needs satisfaction, goal orientation, willingness to communicate, self-efficacy, and learning strategy use as predictors of second language achievement: A structural equation modeling approach. *Journal of Psycholinguistic Research*, *49*(5), 803-822. https://doi.org/10.1007/s10936-020-09714-7

*Khodadad, M., & Kaur, J. (2016). Causal relationships between integrative motivation, self-efficacy, strategy use and English language achievement. *3L: Language, Linguistics, Literature*, *22*(3), 111-125.

*Kitikanan, P., & Sasimonton, P. (2017). The relationship between English self-efficacy and English learning achievement of L2 Thai learners. *LEARN Journal: Language Education and Acquisition Research Network*, *10*(1), 149-164.

Lake, J. (2013). Positive L2 self: Linking positive psychology with L2 motivation. *Language Learning Motivation in Japan*, *71*, 225-244. https://doi.org/10.21832/9781783090518-015

Larson-Hall, J., & Plonsky, L. (2015). Reporting and interpreting quantitative research findings: What gets reported and recommendations for the field. *Language Learning*, *65*(S1), 127-159. https://doi.org/10.1111/lang.12115

Le-Thi, D., Dörnyei, Z., & Pellicer-Sánchez, A. (2020). Increasing the effectiveness of teaching L2 formulaic sequences through motivational strategies and mental imagery: A classroom experiment. *Language Teaching Research*. https://doi.org/10.1177/1362168820913125

Lei, H., & Cui, Y. (2016). Effects of academic emotions on achievement among mainland Chinese students: A meta-analysis. *Social Behavior and Personality: An International Journal*, *44*(9), 1541-1553. https://doi.org/10.2224/sbp.2016.44.9.1541

*Liem, A. D., Lau, S., & Nie, Y. (2008). The role of self-efficacy, task value, and achievement goals in predicting learning strategies, task disengagement, peer relationship, and achievement outcome. *Contemporary Educational Psychology*, *33*(4), 486-512. https://doi.org/10.1016/j.cedpsych.2007.08.001

*Ma, L., Du, X., Hau, K., & Liu, J. (2018). The association between teacher-student relationship and academic achievement in Chinese EFL context: A serial multiple mediation model. *Educational Psychology*, *38*(5), 687-707. https://doi.org/10.1080/01443410.2017.1412400

MacIntyre, P. D., Gregersen, T., & Mercer, S. (Eds.). (2016). *Positive psychology in SLA*. Multilingual Matters.

MacIntyre, P. D., & Mercer, S. (2014). Introducing positive psychology to SLA. *Studies in Second Language Learning and Teaching*, *4*(2), 153-172. https://doi.org/10.14746/ssllt.2014.4.2.2

*McCollum, D. L. (2003). Investigating non-cognitive components of foreign language achievement. *Applied Language Learning*, *13*(1), 19-32.

Marques, S. C., Gallagher, M. W., & Lopez, S. J. (2017). Hope-and academic-related outcomes: A meta-analysis. *School Mental Health*, *9*(3), 250-262. https://doi.org/10.1007/s12310-017-9212-9

Marsden, E., & Plonsky, L. (2018). Data, open science, and methodological reform in second language acquisition research. In A. Gudmestad & A. Edmonds (Eds.), *Critical reflections on data in second language acquisition* (pp. 219-228). John Benjamins.

*Meera, K. P., & Jumana, M. K. (2015). Self-efficacy and academic performance in English. *Research in Pedagogy*, *5*(2), 25-30. https://doi.org/10.17810/2015.13

*Mills, N., Pajares, F., & Herron, C. (2007). Self-efficacy of college intermediate French students: Relation to achievement and motivation. *Language Learning*, *57*(3), 417-442. https://doi.org/10.1111/j.1467-9922.2007.00421.x

*Mizumoto, A. (2013). Effects of self-regulated vocabulary learning process on self-efficacy. *Innovation in Language Learning and Teaching*, *7*(3), 253-265. https://doi.org/10.1080/17501229.2013.836206

Möller, J., Streblow, L., & Pohlmann, B. (2009). Achievement and self-concept of students with learning disabilities. *Social Psychology of Education*, *12*(1), 113-122. https://doi.org/10.1007/s11218-008-9065-z

Möller, J., Zitzmann, S., Helm, F., Machts, N., & Wolff, F. (2020). A meta-analysis of relations between achievement and self-concept. *Review of Educational Research*, *90*(3), 376-419. https://doi.org/10.3102/0034654320919354

Multon, K. D; Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology*, *38*(1), 30-38. https://doi.org/10.1037/0022-0167.38.1.30

*Murad Sani, A., & Zain, Z. (2011). Relating adolescents' second language reading attitudes, self efficacy for reading, and reading ability in a non-supportive ESL setting. *The Reading Matrix, 11*(3), 243-254.

*Mustapha, N. H., Mustapha, N. F., Daud, N., & Abd Wahab, M. (2013). Arabic language efficacy questionnaire (ALEQ): Assessing self-efficacy and achievement. *GEMA Online® Journal of Language Studies*, *13*(1), 155-167.

Namaziandost, E., & Çakmak, F. (2020). An account of EFL learners' self-efficacy and gender in the Flipped Classroom Model. *Education and Information Technologies*, 1-15. https://doi.org/10.1007/s10639-020-10167-7

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 and teaching* (pp. 3-50). John Benjamins.

*Oliver, R., Purdie, N., & Rochecouste, J. (2005). Affective aspects of language learning: Beliefs, attitudes, efficacy. *Babel*, *40*(2), 29-36.

Pajares, F. (1997). Current directions in self-efficacy research. *Advances in motivation and achievement*, *10*(149), 1-49.

Pajares, F., & Schunk, D. H. (2005). Self-efficacy and self-concept beliefs: Jointly contributing to the quality of human life. In H. Marsh, R. Craven, & D. McInerney (Eds.), *International advances in self-research* (Vol. 2, pp. 95-121). Information Age.

Pan, J. Y., & Chan, C. L. W. (2007). Resilience: A new research area in positive psychology. *Psychologia*, *50*(3), 164-176. https://doi.org/10.2117/psysoc.2007.164

Papi, M., Bondarenko, A. V., Mansouri, S., Feng, L., & Jiang, C. (2019). Rethinking L2 motivation research: The 2×2 model of L2 self-guides. *Studies in Second Language Acquisition*, *41*(2), 337-361. https://doi.org/10.1017/S0272263118000153

Phakiti, A., Hirsh, D., & Woodrow, L. (2013). It's not only English: Effects of other individual factors on English language learning and academic learning of ESL international students in Australia. *Journal of Research in International Education*, *12*(3), 239-258. https://doi.org/10.1177/1475240913513520

Pekrun, R., Götz, T., Titz, W., & Perry, R. (2002). Positive emotions in education. In E. Frydenberg (Ed.), *Beyond coping: Meeting goals, visions, and challenges* (pp. 149-173). Oxford University Press.

Petscher, Y. (2010). A meta-analysis of the relationship between student attitudes towards reading and achievement in reading. *Journal of Research in Reading*, *33*(4), 335-355. https://doi.org/10.1111/j.1467-9817.2009.01418.x

Pintrich, P. R., Smith, D. A. F., Garcia, T., & McKeachie, W. J. (1991). *A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ).*

Ann Arbor: University of Michigan, National Center for Research to Improve Postsecondary Teaching and Learning.

Plonsky, L. (2013). Study quality in SLA: An assessment of designs, analyses, and reporting practices in quantitative L2 research. *Studies in Second Language Acquisition*, *35*, 655-687. https://doi.org/10.1017/S0272263113000399

Plonsky, L. (2017). Quantitative research methods in instructed SLA. In S. Loewen & M. Sato (Eds.), *The Routledge handbook of instructed second language acquisition* (pp. 505-521). Routledge.

Plonsky, L., & Derrick, D. J. (2016). A meta-analysis of reliability coefficients in second language research. *Modern Language Journal*, *100*(2), 538-553. https://doi.org/10.1111/modl.12335

Plonsky, L., & Oswald, F. L. (2012). How to do a meta-analysis. In A. Mackey & S. M. Gass (Eds.), *Research methods in second language acquisition: A practical guide* (pp. 275-295). Blackwell.

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

Oswald, F. L., & Plonsky, L. (2010). Meta-analysis in second language research: Choices and challenges. *Annual Review of Applied Linguistics*, *30*, 85-110. https://doi.org/10.1017/S0267190510000115

*Rahemi, J. (2007). Self efficacy in English and Iranian senior high school students majoring in humanities. *Novitas-Royal*, *1*(2), 98-111.

*Raoofi, S., & Maroofi, Y. (2017). Relationships among motivation (self-efficacy and task value), strategy use and performance in L2 writing. *Southern African Linguistics and Applied Language Studies*, *35*(3), 299-310. https://doi.org/10.2989/16073614.2017.1391706

*Sahril, & Weda, S. (2018). The relationship of self-efficacy beliefs, motivation, and writing performance of Indonesian EFL students in higher education. *Journal of English as an International Language*, *13*, 47-63.

Sedighi, F., Alavi, S., & Samani, S. (2004). Developing a foreign language learning self-efficacy scale for Iranian students. *Journal of Social Sciences and Humanities of Shiraz University*, *21*(1 (40)), 94-101.

*Shah, P. M., Mahmud, W. H., Din, R., Yusof, A., & Pardi, K. M. (2011). Self-efficacy in the writing of Malaysian ESL learners. *World Applied Sciences Journal*, *15*(11), 8-11.

*Shang, H. F. (2010). Reading strategy use, self-efficacy and EFL reading comprehension. *Asian EFL Journal*, *12*(2), 18-42.

Seligman, M. P. (1998). *Positive psychology network concept paper*. Retrieved from https://tinyurl.com/yzmwm7vz

Seligman, M. E. P., & Csikszentmihalyi, M. (2000). Positive psychology: An introduction. *American Psychologist, 55*(1), 5-14. https://doi.org/10.1037/0003-066X.55.1.5

Shin, M. H. (2018). Effects of project-based learning on students' motivation and self-efficacy. *English Teaching*, *73*(1), 95-114.

Song, J. (2016). Emotions and language teacher identity: Conflicts, vulnerability, and transformation. *TESOL Quarterly*, *50*(3), 631-654. https://doi.org/10.1002/tesq.312

Teimouri, Y., Goetze, J., & Plonsky, L. (2019). Second language anxiety and achievement: A meta-analysis. *Studies in Second Language Acquisition*, *41*(2), 363-387. https://doi.org/10.1017/S0272263118000311

Teimouri, Y., Plonsky, L., & Tabandeh, F. (2020). L2 grit: Passion and perseverance for second-language learning. *Language Teaching Research*. https://doi.org/10.1177/1362168820921895

*Templin, S. A. (1999). The relationship between self-efficacy and language learners' grades. *JALT Journal*, *21*(1), 112-124.

Terasawa, T. (2017). Has socioeconomic development reduced the English divide? A statistical analysis of access to English skills in Japan. *Journal of Multilingual and Multicultural Development*, *38*(8), 671-685. https://doi.org/10.1080/01434632.2016.1221412

Ushioda, E. (2014). Context and complex dynamic systems theory. In Z. Dörnyei, P. MacIntyre, & A. Henry (Eds.), *Motivational dynamics in language learning* (pp. 47-54). Multilingual Matters.

Ushioda, E., & Dörnyei, Z. (2011). *Teaching and researching: Motivation*. Pearson Education.

Wang, C. (2004). *Self-regulated learning strategies and self-efficacy beliefs of children learning English as a second language* [Unpublished doctoral dissertation, The Ohio State University]. No. osu1091546670. https://etd.ohiolink.edu/

*Wang, C., Hu, J., Zhang, G., Chang, Y., & Xu, Y. (2012). Chinese college students' self-regulated learning strategies and self-efficacy beliefs in learning English as a foreign language. *Journal of Research in Education*, *22*(2), 103-135.

Wang, C., Kim, D.-H., Bong, M., & Ahn, H. S. (2013). Examining measurement properties of an English self-efficacy scale for English language learners in Korea. *International Journal of Educational Research*, *59*, 24-34. https://doi.org/10.1016/j.ijer.2013.02.004

Wang, C., & Sun, T. (2020). Relationship between self-efficacy and language proficiency: A meta-analysis. *System*, *95*. https://doi.org/10.1016/j.system.2020.102366

*Weda, S., Samad, I. A., Patak, A. A., & Fitriani, S. S. (2018). The effects of self-efficacy belief, motivation, and learning strategies on students' academic performance in English in higher education. *Asian EFL Journal*, *20*(9), 140-168.

*Woodrow, L. (2011). College English writing affect: Self-efficacy and anxiety. *System*, *39*(4), 510-522. https://doi.org/10.1016/j.system.2011.10.017

Wu, X., Lowyck, J., Sercu, L., & Elen, J. (2012). Self-efficacy, task complexity and task performance: Exploring interactions in two versions of vocabulary learning tasks. *Learning Environments Research*, *15*(1), 17-35.

*Wu, X., Lowyck, J., Sercu, L., & Elen, J. (2013). Task complexity, student perceptions of vocabulary learning in EFL, and task performance. *British Journal of Educational Psychology*, *83*(1), 160-181. https://doi.org/10.1111/j.2044-8279.2011.02061.x

Wyatt, M. (2018). Language teachers' self-efficacy beliefs: An introduction. In S. Mercer & A. Kostoulas (Eds.), *Language teacher psychology* (pp. 122-140). Multilingual Matters.

Zhang, L., & Tsung, L. (2021). Learning Chinese as a second language in China: Positive emotions and enjoyment. *System*, *96*, 102410. https://doi.org/10.1016/j.system.2020.102410

APPENDIX A

Descriptions of variables used in the coding scheme (adapted from Teimouri et al., 2019)

Variable	Description
Bibliographic information	
Author(s)	Name(s) of researchers who conducted and published the study
Publication year	Year in which the study was published
Publication title	Title of the article or chapter
Publication venue	Title of the journal or volume in which the study was published
Participant sample	
Sample size	Number of participants
Age	Average age or age range of participants
Nationality	Nationality of participants
L1	L1 of participants
Target language	Language of interest for self-efficacy and achievement measures
Academic level	Educational level (elementary, middle, or high school, university, language institute, tutoring/private lessons) of participants
Proficiency level	Proficiency level (beginner, intermediate, advanced, multiple) of participants
Language learning experience	Average number of months/years of experience that participants had with learning the language
Study design	
Design	Experimental or observational
Methodology	Quantitative, qualitative, or mixed methods
Context	Target language context (L2, FL, heritage, combined L2/FL, study abroad)
Setting	Setting in which the research took place (classroom, lab, combined classroom/lab, computer- or mobile-assisted learning, distance learning)
Country	Country in which the research took place
Self-efficacy instrument	
Terminology	Terminology used to refer to self-efficacy (e.g., L2 self-efficacy)
Type	General or skill-specific
Definition	Definition used for self-efficacy and whether the definition was adapted from previous work or newly-formulated
Variable role	Dependent or independent variable, or not applicable
Instrument type	Whether the instrument was adapted from previous work or newly-developed; if adapted, original publication from which the instrument was adapted
N of items	Number of items in the instrument
Items type	Whether or not the instrument was a Likert-scale questionnaire and, if so, the range of the Likert scale
Mean	Reported mean for the self-efficacy measure
Standard deviation	Standard deviation for the self-efficacy measure
Reliability	Whether or not a reliability measure was reported
Reliability index	Reported reliability index (e.g., Cronbach's alpha)
Reliability estimate	Observed reliability coefficient
L2 achievement instrument	
Type	Type of achievement measure (e.g., language test, course grade, self-perceived competence, GPA)
Mean	Reported mean for the achievement measure
Standard deviation	Standard deviation for the achievement measure
Reliability	Whether or not a reliability measure was reported
Reliability index	Reported reliability index (e.g., Cronbach's alpha)
Reliability estimate	Observed reliability coefficient
Results	
Type of analysis	Type of analysis run between self-efficacy and L2 achievement (e.g., Pearson correlation, regression)
Effect size (*r*)	Observed correlation between self-efficacy and L2 achievement

APPENDIX B

Effect sizes in the studies under review

Primary study	N	M	95% CI	
			Lower	Upper
Alrabai (2018)	221	.549*	.450	.635
Asakereh & Dehghannezhad (2015)	100	.560*	.409	.681
Bai et al. (2019)	1092	.530*	.486	.571
Balci (2017)	78	.440*	.241	.603
Chao et al. (2019)	1092	.470*	.422	.515
Fukuda (2018)	97	.304**	.111	.475
Hetthong & Teo (2013)	51	.712*	.543	.826
Hsieh (2008)	249	.590*	.503	.666
Hsieh & Kang (2010)	192	.468*	.350	.572
Hsieh & Schallert (2008)	500	.530*	.464	.590
Jee (2015)	92	.272**	.071	.452
Khodadad & Kaur (2016)	240	.423*	.313	.522
Kitikanan & Sasimonton (2017)	32	.615*	.339	.794
Ma et al. (2018)	11,036	.472*	.457	.486
McCollum (2003)	128	.530*	.393	.644
Mizumoto (2013)	303	.570*	.489	.641
Al Khamisi et al. (2016) – Study 1	260	.541*	.449	.622
Al Khamisi et al. (2016) – Study 2	376	.518*	.440	.588
Raoofi & Maroofi (2017)	304	.530*	.444	.606
Shah et al. (2011)	120	.563*	.427	.674
Shang (2010)	53	.530*	.303	.700
Templin (1999)	74	.310**	.088	.503
Wang et al. (2012) – Study 1	474	.380*	.300	.455
Wang et al. – Study 3	476	.360*	.279	.436
Woodrow (2011)	738	.430*	.369	.487
Wu et al. (2013)	120	.330*	.160	.481
Weighted average***	18,265	.475*	.464	.486

Note. * $p < .001$; ** $p < .01$; *** Weighted means with 95% confidence intervals of final sample ($k = 27$)