Studies in Second Language Acquisition (2024), **46**, 1515–1536
doi:10.1017/S0272263124000512

REPLICATION STUDY

Reexamining the role of regulatory focus in second language achievement

An approximate replication of Papi and Khajavy (2021)

Li Wang and Xinyi Sun

School of Foreign Languages, Shanghai Normal University, Shanghai, China
Corresponding author: Li Wang; Email: Wanglily22@shnu.edu.cn

(Received 06 May 2023; Revised 21 June 2024; Accepted 03 July 2024)

Abstract

In this study, we approximately replicated Papi and Khajavy's (2021) investigation into the regulatory focus in second language acquisition, applying their approach to a significantly larger cohort of 855 Chinese second language (L2) learners. In contrast with the original study, our research employed the College English Test Band 4 (CET-4), a standardized English test, to better align with the Chinese educational context. This methodological shift allowed for a nuanced exploration of the regulatory focus's role in language learning among Chinese students. Our results predominantly reinforce the crucial role of regulatory focus in language learning, echoing Papi and Khajavy's findings. We discovered notable parallels in how promotion positively influences ideal own and ideal other, and how prevention negatively impacts ought own and ought other. However, a unique finding of our study was the stronger impact of ought selves on L2 anxiety and enjoyment in the Chinese context. This highlights the nuanced influence of socioeducational environments on regulatory focus strategies, suggesting contextual variability in language learning strategies.

Keywords: replication study; second language learners; second language acquisition; regulatory focus; motivation

Introduction

In an endeavor to understand the complexities of motivation in language acquisition, Dörnyei (2005, 2009) amalgamated existing second language (L2) motivation theories with constructs from motivational psychology, leading to the conceptualization of the Second Language Motivational Self System (L2MSS). This framework, which includes the ideal L2 self, the ought–to L2 self, and the L2 learning experience, has influenced research on language learners' motivations. However, recent studies have raised critical questions about the L2MSS model, highlighting issues such as the "fantasy problem," the "ought–to L2 self problem," and the "context problem" (Henry & Liu, 2023; Al-Hoorie, 2018). This study aims to explore the applicability of the 2X2 model among Chinese EFL learners.

The ideal L2 self in this system is envisioned as a learner's future self, adept in the target language, typically associated with positive emotional states and aspirations. In contrast, the ought–to L2 self mirrors the learner's perceived duties and responsibilities in the realm of language learning, often spurred by external factors. The ideal L2 self and the ought–to L2 self, as key components of the L2MSS, influence the L2 learning experience. This experience includes situational motivations linked to the learner's immediate context, shaping the overall motivational landscape. However, recent studies have raised critical questions about the L2MSS model, highlighting issues such as the "fantasy problem," the "ought–to L2 self problem," and the "context problem" (Henry & Liu, 2023; Al-Hoorie, 2018). In response to these drawbacks, Papi, Bondarenko, Mansouri, Feng, and Jiang (2019) proposed the 2X2 model to provide a more nuanced understanding of motivational dynamics by integrating regulatory focus theory with L2 self-guides. This study aims to explore the applicability of the 2X2 model among Chinese EFL learners.

Scholarly inquiry into L2MSS spans various domains: from theoretical explorations (Dörnyei, 2005; Dörnyei & Ushioda, 2009) to investigations into the constructs of L2MSS (Saito, Dewaele, Abe, & In'nami, 2018; Mahdavy, 2020), studies probing influencing factors (Dewaele & MacIntyre, 2016; Dörnyei & Chan, 2013), and research dissecting the nexus between motivation and pedagogical approaches (Moskovsky, Alrabai, Paolini, & Ratcheva, 2013; Henry, Korp, Sundqvist, & Thorsen, 2018).

Building on this foundational work, Papi et al. (2019) introduced the 2X2 model, which was designed to further refine the conceptualization of future self-guides by integrating promotion and prevention regulatory foci. The model categorizes learners' motivational orientations into four types based on their regulatory focus (promotion vs. prevention) and their self-guides (ideal self vs. ought–to self). This integration offers a more detailed understanding of how different motivational dimensions interact to affect L2 learning outcomes, providing a nuanced perspective on these processes.

Several studies have explored the applications of the 2X2 model in various contexts. Research has demonstrated the model's effectiveness in predicting language learning behaviors and outcomes across different cultural and educational settings. Papi et al. (2019) found that the 2X2 model could explain variations in learners' engagement and persistence in L2 learning. Teimouri (2017) showed that learners with a promotion focus were more likely to adopt approach–oriented strategies, while those with a prevention focus tended to avoid language learning tasks perceived as threatening. Tahmouresi and Papi (2021) found that future selves, enjoyment, and anxiety predicted L2 writing achievement, while Zhou and Papi (2023) demonstrated the role of future L2 selves in L2 speech development in an instructional setting. Kermad (2018) explored speaker and listener variability in the perception of non-native speech, contributing further to the understanding of motivational and perceptual dynamics in L2 learning. Feng and Papi (2020) explored the relationship between the 2X2 model and language learning anxiety, revealing that a prevention focus was positively correlated with higher levels of anxiety, while a promotion focus was associated with greater language enjoyment.

Papi and Khajavy (2021) applied the 2X2 framework in their investigation into the role of regulatory focus in L2 achievement. Conducted with a smaller sample in a different cultural setting, their research highlighted significant interactions between promotion and prevention foci and L2 achievement, revealing how these regulatory focuses differently impact language learning outcomes. Their findings suggest that promotion focus is positively associated with ideal self-guides and language learning success, while prevention focus is linked to ought self-guides and is associated with higher levels of anxiety and

https://doi.org/10.1017/S0272263124000512 Published online by Cambridge University Press

vigilant L2 use strategies. By examining these processes in a novel context, their study provided valuable insights into the applicability of the 2X2 model.

Thus, our study, with its extensive sample size of 855 Chinese L2 learners and robust methodological approach, is uniquely positioned to enrich the understanding of L2 motivational factors among Chinese EFL learners. By aiming to test the generalizability of Papi and Khajavy's (2021) theoretical model in the specific context of Chinese university students, this research contributes to the understanding of second language motivation in light of these comparative insights and seeks to deepen our understanding of the applicability of this theoretical model across diverse educational and cultural backgrounds. This approach, therefore, offers potential implications for second language learning strategies, particularly in the unique socioeducational landscape of China, thereby making a meaningful contribution to the field.

Motivation for replication

This study is primarily motivated by the critical need to explore the motivational factors in Chinese EFL learners' second language acquisition. While a significant body of research, including Papi and Khajavy's (2021) study, has examined L2 motivation, further research is needed to fully understand how these motivational constructs function within the L2MSS framework, particularly in the unique cultural and educational context of China. Our research aims to bridge this gap by delving into how the "ought–to L2 self" interacts with other motivational constructs in a Chinese educational setting, aiming for a more holistic understanding of its role in L2 acquisition.

Papi et al.'s (2019) 2X2 model of future self-guides, intertwining promotion, and prevention regulatory foci, presents a novel lens through which to examine these motivational factors. Our decision to replicate and extend this model in the Chinese context is driven by its potential to uncover unique insights into the motivational processes among Chinese EFL learners—an area previously underexplored. This replication study, leveraging a substantial sample size and a robust methodological framework, is positioned not only to validate and adapt Papi and Khajavy's findings but also to deepen our comprehension of the interplay between regulatory focus, L2 self-guides, emotional reactions, strategic behaviors, and L2 achievement in a distinct cultural setting. Through this approach, we aim to contribute significantly to the broader discourse on L2 motivation, underscoring how diverse cultural and educational contexts can influence the efficacy and applicability of established motivational models.

Context comparison

In comparing English language education in Iran and China, we observe distinct approaches to learning, teaching, assessment, and policy prioritization. While both countries recognize English's global importance, their educational strategies and implementation significantly differ. This comparative backdrop is essential for our study, particularly in understanding the unique socioeducational factors influencing EFL learners' motivation in China.

For instance, in Iran, English is introduced as a foreign language in secondary school, while in China, English education begins earlier, typically in Grade 3. This early exposure in China forms a critical part of the context for our investigation into motivational strategies among Chinese English as a Foreign Language (EFL) learners (Yang, 2014). Further, the assessment methodologies, with Iran favoring international assessments and

https://doi.org/10.1017/S0272263124000512 Published online by Cambridge University Press

China relying on the Gaokao (the National College Entrance Examination), may distinctively shape Chinese students' learning strategies and motivational orientations, aligning with our study's focus on regulatory focus and self-guides (Fang, 2018).

Moreover, policy perspectives in both nations shape their approaches to English education. In China, the aggressive promotion of English learning aligns with a broader strategy for global competitiveness (Gil & Adamson, 2011). In contrast, in Iran, the importance of English is primarily in the role that it plays in educational advancement (Moharami & Daneshfar, 2021). The practical use of English also diverges between the two countries, limited by political factors in Iran and increasingly valued for academic and occupational opportunities in China. As we explore the impact of these educational policies on language learning motivation, it becomes imperative to adopt a systematic approach to replicate and extend existing research. To this end, we have embraced the replication study framework as defined by Porte and McManus (2019). According to their categorization, "approximate replication" allows for more modifications than "close replication," which is limited to altering only one major variable. However, it is crucial in an approximate replication not to change all methodological aspects of the original study; otherwise, it transitions into a conceptual replication. This framework is essential in guiding our approach to examining Papi and Khajavy's (2021) model in the Chinese educational context while maintaining methodological coherence.

Research questions

In the original study by Papi et al. (2021), a comprehensive model was explored that integrated various elements such as regulatory focus, future L2 self-guides, emotions, and strategic inclinations in the context of second language achievement. Although Papi et al. did not explicitly frame their investigation around distinct research questions, their study implicitly addressed key aspects of how these elements interact and influence language learning outcomes. In our replication study, we aim to explicitly articulate these underlying inquiries, adapting them to the context of Chinese EFL learners. Our research questions are formulated to closely mirror the implicit inquiries of the original study, focusing on the unique processes of motivational and emotional factors among Chinese learners of English as a foreign language.

RQ1: How does the regulatory focus (promotion and prevention) relate to the ideal and ought–to selves of Chinese EFL learners?

RQ2: What is the impact of these self-guides on the emotional experiences (anxiety and enjoyment) of Chinese EFL learners?

RQ3: How do these emotional patterns influence the strategic behaviors (eager vs. vigilant L2 use) of Chinese EFL learners?

RQ4: How do these strategic behaviors contribute to the overall L2 achievement among Chinese EFL learners?

The predicted outcomes of this replication are expected to align with Papi & Khajavy's (2021) findings, aiming to replicate the original results and potentially extend them by identifying new predictive pathways specific to the context of Chinese EFL learners. Specifically, (a) the promotion focus will positively predict the ideal L2 selves, and the prevention focus will positively predict the ought selves; (b) the ideal L2 selves will predict L2 enjoyment positively and L2 anxiety negatively, whereas the ought selves

https://doi.org/10.1017/S0272263124000512 Published online by Cambridge University Press

will predict L2 anxiety positively and L2 enjoyment negatively; (c) L2 enjoyment will predict eager strategies positively and vigilant strategies negatively, whereas L2 anxiety will predict vigilant strategies positively and eager strategies negatively; and (d) eager strategies will positively, and vigilant strategies will negatively, predict L2 achievement. By replicating and extending Papi and Khajavy's (2021) study, we aim to contribute to the understanding of motivational mechanisms in second language learning, particularly for Chinese EFL learners.

Method

In alignment with the framework proposed by Porte & McManus (2019), this study adopts an "approximate replication" approach, which permits a broader scope of modifications compared to "close replication." Close replication is characterized by its restriction to altering a single major variable. In contrast, approximate replication allows for multiple changes, provided these alterations do not encompass all methodological aspects of the original study. Exceeding this boundary would categorize the study as a "conceptual replication" rather than an approximate one. In our research, we have judiciously adjusted three specific elements: the participants' language background, the language used in the questionnaire, and the method of measuring L2 achievement. These changes, detailed in Table 1, are within the permissible scope of approximate replication and serve to contextualize our study within the framework of the original research by Papi and Khajavy (2021).

Participants

This investigation involved 855 university students from diverse academic backgrounds in China, with a gender distribution of 20.58% male and 79.42% female,

Table 1. Comparative overview of variables in the initial study and Its replication

Variable	Initial study	Replication study
L1 background	Farsi	Chinese
Number of participants	324	855
Age range	18 to 38	18 to 36
Education level	Undergraduate (most in first year)	Undergraduates: 66.43%; Graduates: 33.57%
Language learning context	General English course at a university in Iran	General English course at universities in China
Major Distribution	Humanities, Engineering, Arts, Sciences, Unknown Literature, Education Science, Management Science, Engineering, Natural Science, Economics, Others and Unknown	Humanities, Engineering, Arts, Sciences, Unknown Literature, Education Science, Management Science, Engineering, Natural Science, Economics, Others and Unknown
Questionnaire language	From English to Farsi	From English to Chinese
L2 achievement	Final exam: Reading comprehension, Vocabulary knowledge, Grammar, Pronunciation: Listening, Reading Comprehension, Translation and Writing	CET–4: Listening, Reading comprehension, Translation and Writing

https://doi.org/10.1017/S0272263124000512 Published online by Cambridge University Press

spanning an age range of 18 to 36 years (M = 21.31, SD = 2.30). This broad demographic spectrum was deliberately selected to reflect the evolving landscape of English language learning within the Chinese higher education system, particularly emphasizing China's growing prominence in the global context of language acquisition studies.

The sample size, substantially exceeding the minimum requirement of 434 as recommended by MacCallum, Browne, and Sugawara (1996) and validated through the Preacher and Coffman (2006) calculator, provides robust statistical power for the study. This ensures a higher degree of reliability and validity in the empirical testing of the Papi and Khajavy (2021) model on Chinese EFL learners' motivation, a critical aspect of the study's methodological framework.

The participants, enrolled in a mandatory general English course, represented a wide array of disciplines, including literature (27.13%), education science (23.04%), management science (12.28%), engineering (11.35%), natural science (8.07%), and economics (7.25%), with additional representation from history, philosophy, agronomy, art, law, and medicine. This interdisciplinary composition of the sample not only enhances the generalizability of the findings but also provides a comprehensive insight into the multifaceted nature of English language acquisition across different academic domains.

Comprising 568 undergraduates (66.43%) and 287 graduate students (33.57%), the participants were chosen to encapsulate a range of learning stages, from a test–centric approach to a more practical application of English in academic and professional settings. This strategic selection underscores the relevance of the study in examining the nuanced processes of language motivation among adult learners in a rapidly globalizing educational context.

Procedures and instruments

The present study adopted the same questionnaire (available from IRIS, https://www.iris-database.org) as Papi and Khajavy (2021), only the language being translated into the participants' native language for better understanding. The questionnaire has two sections. The first section collected the personal information of the participants while the second section measured the participants' regulatory focus, L2 self-guides, L2 anxiety and enjoyment, and L2 strategic inclinations. The Likert scale, ranging from 1 (strong disagreement) to 6 (strong agreement), was employed in this study. The survey was distributed through Wenjuanxing, a professional survey platform widely used in China (http://www.wjx.cn/).

The questionnaire was translated into Chinese and the Chinese version was revised by a professor proficient in linguistics and translation over three rounds. As the participants had varying levels of English proficiency, a pilot study with 10 students was conducted to test the comprehensibility of the questionnaire. Based on the feedback received from the pilot study, the items that were found to be confusing or difficult to understand were adjusted. The finalized version of the questionnaire, along with the raw data from our study, has been archived and made publicly accessible for further scholarly review and analysis on the Open Science Framework (OSF) platform. These resources can be found at the following link: https://osf.io/mhj25/.

In this study, we utilized the College English Test Band 4 (CET-4) to measure L2 achievement, diverging from the approach taken by Papi and Khajavy (2021). CET-4, administered by the National College English Testing Committee on behalf of the Higher Education Department, Ministry of Education, the People's Republic of China,

https://doi.org/10.1017/S0272263124000512 Published online by Cambridge University Press

is a comprehensive assessment tool specifically designed for the Chinese educational context. It evaluates students' listening, reading comprehension, translation, and writing skills across a total of 710 points, with specific subscores for listening, reading comprehension, translation, and writing. The choice of CET-4 was driven by its relevance within the Chinese context and the availability of data. While we recognize that international assessments like the International English Language Testing System (IELTS) might offer broader cross-cultural applicability, the limited availability of such data guided our decision. It is important to note that the findings from CET-4, while robust within the Chinese language learning environment, may not be directly transferable to other educational contexts. Therefore, this study's focus is on the specific nuances of language learning within China, a context where CET-4 provides a detailed and relevant assessment.

Yang and Weir (1998) have provided clear guidelines for grading CET-4 and have established the reliability and validity of the tests. They reported instrument reliability and inter-rater reliability by considering intramarker consistency, intermarker consistency, and intercenter consistency. In each test, the internal reliability was over 0.9. They also considered construct validity, content validity, concurrent validity, predictive validity, and face validity, and up to 92% of the teachers thought CET-4 reflected students' actual English competence. The validity and reliability of CET-4, an English language proficiency test, have been well-established in the research, notably by Jin and Wu (1998). Since the reliability of the tests is already well-established, no further test was conducted to determine the reliability of the tests in our study.

Data analysis

Data collection was conducted between March 30 and May 20, 2022, adhering to the same procedures as outlined in Papi and Khajavy's (2021) study.

As a preliminary step for conducting SEM, we performed a thorough examination of missing data, outliers, and assumptions relevant to SEM. Among 855 data samples, 25 lacked values of CET-4 grades. The reason for the 2.9% missing data rate was that some participants had forgotten their CET-4 grades. In contrast to Papi and Khajavy's (2021) study, no univariate outliers were detected in our analysis. However, we identified 103 multivariate outliers. They were subsequently removed to avoid affecting the statistical inference and to maintain the stability of the model fit, leaving a final sample of 752 participants. Among the 752 participants, 257 are graduates and 495 are undergraduates.

As with the original study, we assessed the skewness and kurtosis of all variables and found them to be within acceptable ranges, as shown in Table 2, indicating univariate normality. Despite this, Mardia's test for multivariate normality (Korkmaz, Goksuluk, & Zararsiz, 2014) revealed that the data did not exhibit multivariate normality. This result was further confirmed by a quantile–quantile plot (Figure 1), which clearly illustrates the absence of multivariate normality.

Following the feedback received and the revised confirmatory factor analysis (CFA) for the measurement models of all constructs, we continued to investigate the constructs using CFA as per Papi and Khajavy (2021). The goodness–of–fit indices for the revised model were utilized to confirm model fit in CFA and SEM. The standards for the comparative fit index (CFI), Tucker–Lewis index (TLI), root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR) were reassessed and found to be within the optimal range postrevision. Specifically, CFI and

https://doi.org/10.1017/S0272263124000512 Published online by Cambridge University Press

Table 2. Descriptive statistics and reliability of the final scales

	n	Mean	SD	Skewness	Kurtosis	α	ω
Promotion	752	4.26	0.89	0.01	–0.15	0.81	0.89
Prevention	752	4.00	1.12	–0.11	–0.39	0.67	0.82
Ideal own	752	4.21	1.14	–0.24	–0.41	0.95	0.90
Ideal other	752	4.65	1.04	–0.62	0.19	0.83	0.96
Ought own	752	3.94	1.02	0.07	–0.38	0.81	0.88
Ought other	752	3.12	1.20	0.44	–0.24	0.88	0.92
Enjoyment	752	3.81	1.12	–0.12	–0.19	0.91	0.94
Anxiety	752	4.11	0.95	–0.15	–0.12	0.83	0.88
Eager	752	3.75	1.02	0.39	0.21	0.90	0.93
Vigilant	752	3.85	0.97	–0.17	–0.06	0.85	0.89
L2 achievement	752	487.89	159.38	–8.22	74.63	——	——

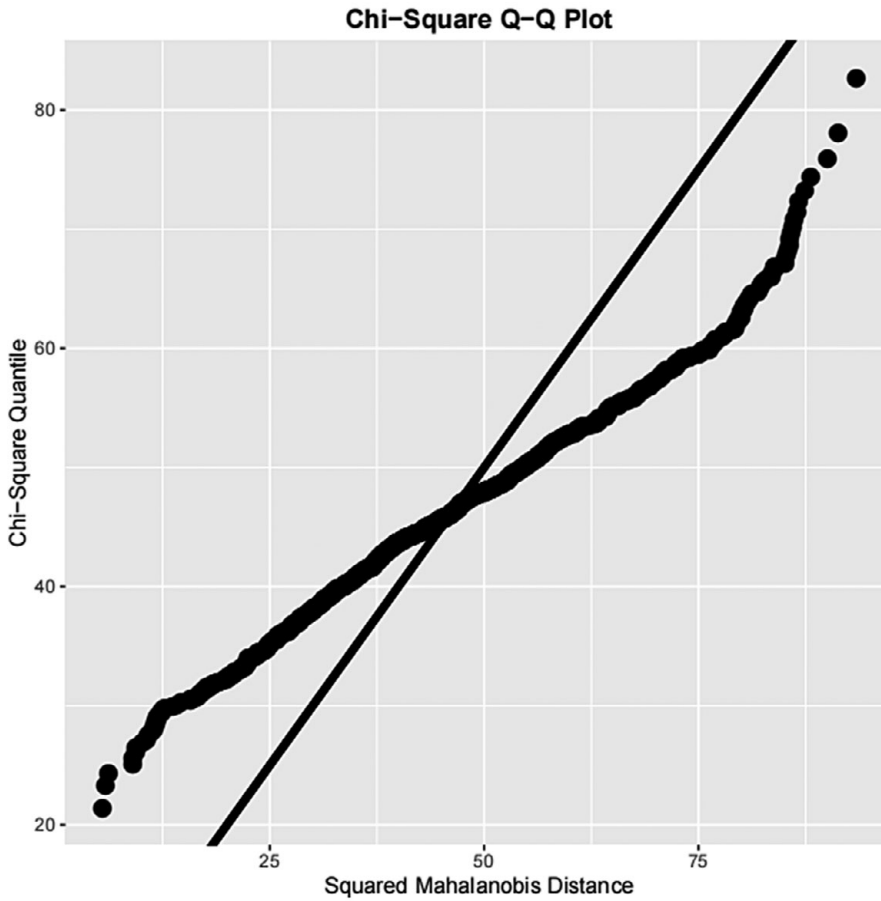

Figure 1. Quantile–quantile plot.

https://doi.org/10.1017/S0272263124000512 Published online by Cambridge University Press

TLI values are now higher than .90 and .95, indicating an adequate and excellent fit to the data respectively; RMSEA and SRMR values are less than .08 and .06, respectively, showing adequate and excellent fit (Hu & Bentler, 1999).

When reporting effect sizes for the paired–sample t-tests, Cohen's d was calculated as per the original study. We continued to reference the framework proposed by Plonsky and Oswald (2014) for within–subject comparisons: $0.60 \leq d < 1.00$ as generally small; $1.00 \leq d < 1.40$ as medium; and $d \geq 1.40$ as large. For SEM analyses, Cohen's f^2 (Cohen, 1992) was used for effect sizes with the following interpretation: $0.02 \leq f^2 < 0.15$, a small effect; $0.15 \leq f^2 < 0.35$, a medium effect; and $f^2 \geq 0.35$, a large effect.

The data were analyzed following agreed–upon steps for latent variable models (SEM) with the assistance of Mplus 8 (Muthén & Muthén, 1998-2017). Drawing from the results of the revised CFA, we examined a series of measurement models to evaluate the congruence of the data with the theoretical constructs. The structural model was subsequently estimated, examining the relationships between the variables within the hypothesized model.

Following Hiver and Al-Hoorie (2020), we assessed the convergent and discriminant validity of the constructs. Convergent validity was evaluated by examining whether the average variance extracted (AVE) for each construct exceeded .50. Discriminant validity was assessed by comparing the square root of the AVE for each construct with the correlations between that construct and other constructs. Specifically, discriminant validity is considered satisfactory if the square root of the AVE for a construct is greater than the correlations between that construct and any other constructs. The AVE is calculated as the mean of the squared loadings of the indicators related to the construct, and an AVE value above .50 indicates that the construct explains more than half of the variance of its indicators. Our results were not satisfactory, which implies a weak convergent and discriminant validity.

Factor analysis

Due to the low construct validity of the Regulatory Focus measure, we conducted an exploratory factor analysis (EFA) to reassess its construct validity. The Kaiser–Meyer–Olkin measure and Bartlett's test of sphericity were performed on the data of the four question items of Promotion. The data in question showed that the KMO value was 0.706 (greater than 0.7) and Bartlett's spherical test reached a significance level of 0.001, which indicated that the raw data were suitable for factor analysis. Factor analysis was performed using principal component analysis and varimax orthogonal rotation. In this study, the factors were extracted according to the criterion of eigenvalue greater than 1, and the items with factor loading less than 0.6 were gradually eliminated from the smallest to the largest. Therefore "Promotion 2[Compared to most people, I am typically unable to get what I want out of life]" was eliminated, and the scale items were streamlined to three, and one common factor was extracted. Further exploratory factor analysis was conducted on the remaining three items (shown in Table 1), and the KMO value of the sample data was 0.687, with a significant Bartlett's spherical test result ($p<0.001$), and one common factor was still extracted, with a cumulative variance contribution rate of 69.322%, and the overall coefficient α of the scale was 0.776, indicating that the scale has good reliability.

The same analytical approach was applied to the five question items of Prevention and finally, the scale items were streamlined to three. According to Bollen (1989), there should be at least three items for each potential construct, with five to seven items being

https://doi.org/10.1017/S0272263124000512 Published online by Cambridge University Press

preferred. Therefore, we retained three items for promotion (prom1, prom3, prom4; α= 0.85) and three for prevention (prev1, prev2, prev4; α= 0.80).

Results of the measurement models

Following Papi and Khajavy (2021), we ran CFA to assess the construct validity. As shown in Table 3, goodness–of–fit indices of Regulatory Focus, 2X2 model, and Strategic Inclinations all fit the data. The 'Emotions' factor achieved marginal fit but did not fully meet the criteria. As can be seen from Figure 2, factor loadings of Emotions ranged from .523 to .859, which indicates the strength of the relationship between the observed variables and the latent variable (joy and anxiety) ranging from middle to strong.

We also performed CFA for other constructs. Results for all the scales supported their construct validity, as demonstrated by Figures 2, 3, 4, and 5, and Table 3. The convergent validity and discriminant validity of each construct are further confirmed in Table 4.

As can be seen from Table 5, the results of measurement invariance testing confirmed the configural, metric, and scalar invariance. This implies that the items and latent factors were interpreted similarly by men and women in the current study.

To check the reliability of the model, we measured the internal consistency and composite reliability with coefficientαand coefficientω, respectively. As shown in Table 2, all measures had an acceptable level of coefficientαand coefficientω. This implies that all the scales are internally consistent.

In response to concerns about the model fit, we have conducted further analyses to ensure the robustness of our findings. After an initial review indicating borderline or poor model fit, we meticulously revised our measurement models. This involved removing items with poor loadings from the Regulatory Focus scale and reassessing the scale's structure. Subsequently, we explored and tested an alternative structural model. These steps were crucial in enhancing the model's fit to our data and addressing potential confirmation bias. The revised model, now demonstrating improved fit indicators, better represents the data in our current study context.

Results of t-tests

Following Papi and Khajavy (2021), we carried out paired–sample t-tests to examine the mean differences among the subscales. As shown in Table 6, among the four self–guide subscales, our participants showed higher means for three subscales when compared with Ought Other: Ideal Own (large effect), Ideal Other (large effect), and Ought Own (medium effect). Additionally, Ideal Other had a slightly higher mean than Ideal Own (small effect) and Ought Own (medium effect). Ideal Own also had a slightly higher mean than Ought Own (small effect). Participants used slightly more Vigilant

Table 3. Goodness–of–fit indices of CFA and SEM

	x^2	*df*	CFI	TLI	RMSEA (90%CI)	SRMR
Regulatory Focus	13.885	8	0.993	0.987	0.000 (0.000, 0.072)	0.022
2X2 model	449.226	98	0.938	0.924	0.077 (0.070, 0.093)	0.055
Emotions	351.597	26	0.872	0.822	0.145 (0.122, 0.174)	0.104
Strategic Inclinations	198.606	34	0.937	0.917	0.086 (0.078, 0.112)	0.061
Final SEM model	2517.337	783	0.882	0.871	0.064 (0.059, 0.070)	0.076

https://doi.org/10.1017/S0272263124000512 Published online by Cambridge University Press

Table 4. Convergent validity and discriminant validity of each construct

	Convergence Validity	Discriminant Validity									
Dimension	AVE	Ideal other oth Other Other	Oughtown Own	Ideal own Own	Ought other Other	Prom	Prev	Enjoy	Anxiety	Eager	Vigilant
Ideal Other	0.381	0.617									
Ought Own	0.298	0.528	0.546								
Ideal Own	0.607	0.616	0.507	0.779							
Ought Other	0.424	0.459	0.696	0.364	0.651						
Prom	0.452	0.463	0.362	0.654	0.357	0.672					
Prev	0.242	0.172	0.05	0.259	0.024	0.322	0.492				
Enjoy	0.527	0.575	0.470	0.833	0.413	0.660	0.266	0.726			
Anxiety	0.273	0.214	0.373	0.034	0.352	0.01	–0.025	–0.003	0.522		
Eager	0.434	0.516	0.448	0.745	0.482	0.676	0.280	0.832	–0.059	0.659	
Vigilant	0.281	–0.008	0.213	–0.245	0.271	–0.127	–0.064	–0.276	0.715	–0.266	0.531

Note: Ideal Other refers to Ideal L2 Self/Other; Ought Own refers to Ought L2 Self/Own; Ideal Own refers to Ideal L2 Self/Own; Ought Other refers to Ought L2 Self/Other; Prom refers to promotion focus; Prev refers to prevention focus.

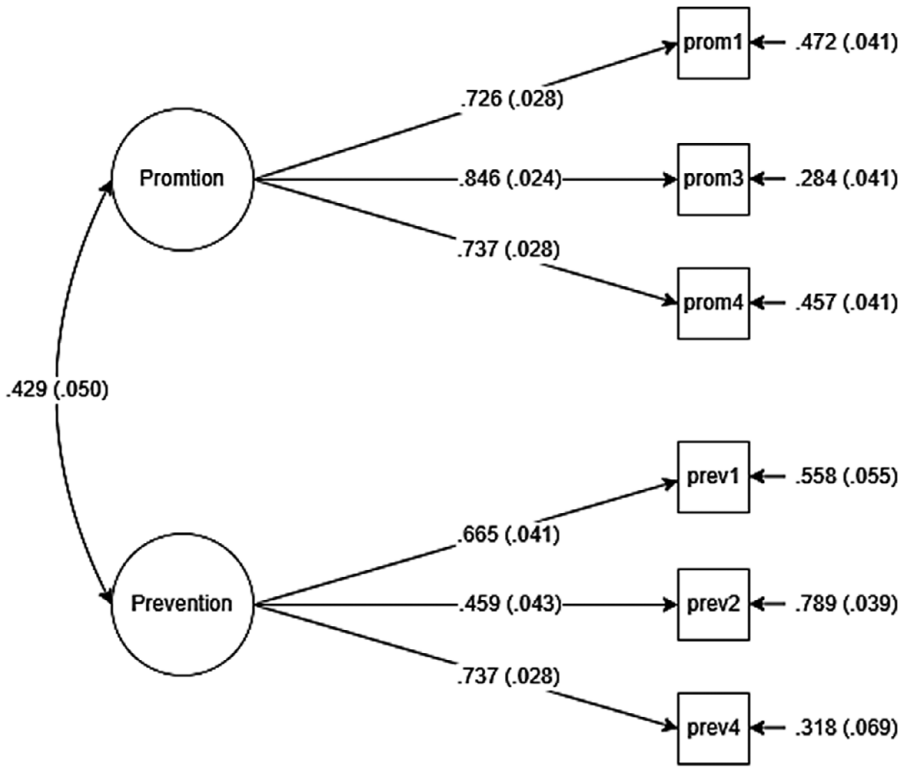

Figure 2. The results of CFA for the regulatory focus with standardized estimates.

than Eager strategies (medium effect), experienced more Anxiety than Enjoyment (small effect), and were slightly more Promotion-focused than Prevention-focused.

Results of the structural model

We used SEM to test the structural model with five layers of variables (see Figure 6). As can be seen from Table 3, goodness–of–fit indices displayed that the model fitted the data adequately.

In terms of paths in the first layer, Promotion was a significant positive predictor of Ideal Own (f^2 = 1.37, large effect), Ideal Other (f^2 = .53, large effect), Ought Own (f^2 = .39, large effect), and Ought Other (f^2 = .46, large effect). Prevention was a significant negative predictor of Ought Own (f^2 = .03, small effect) and Ought Other (f^2 = .08, small effect).

In the second layer, Ideal Own was a significant positive predictor of Enjoyment (f^2 = 4.82, large effect) and is a significant negative predictor of Anxiety (f^2= .020, medium effect), as anticipated. However, Ought Own (f^2 = .35, large effect) was a significant positive predictor of Anxiety and Ought Other was a significant positive predictor of Enjoyment (f^2 = .13, small effect).

In the third layer, Enjoyment was a significant positive predictor of Eager L2 Use (f^2 = 6.51, large effect) and a significant negative predictor of Vigilant L2 Use (f^2 = .16, medium effect), whereas Anxiety was a significant positive predictor of Vigilant L2 Use

https://doi.org/10.1017/S0272263124000512 Published online by Cambridge University Press

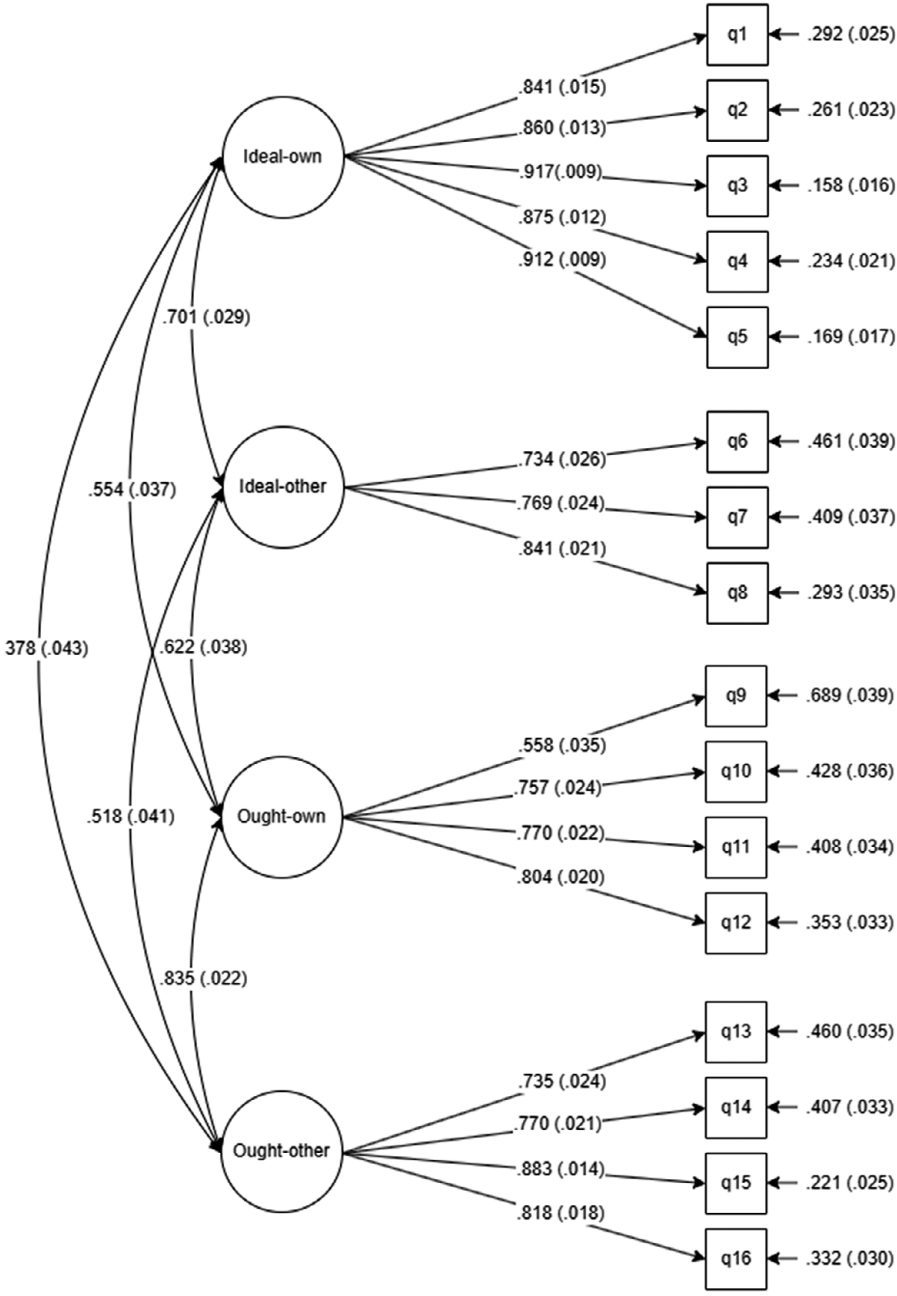

Figure 3. The results of CFA for the L2 self-guides with standardized estimates.

https://doi.org/10.1017/S0272263124000512 Published online by Cambridge University Press

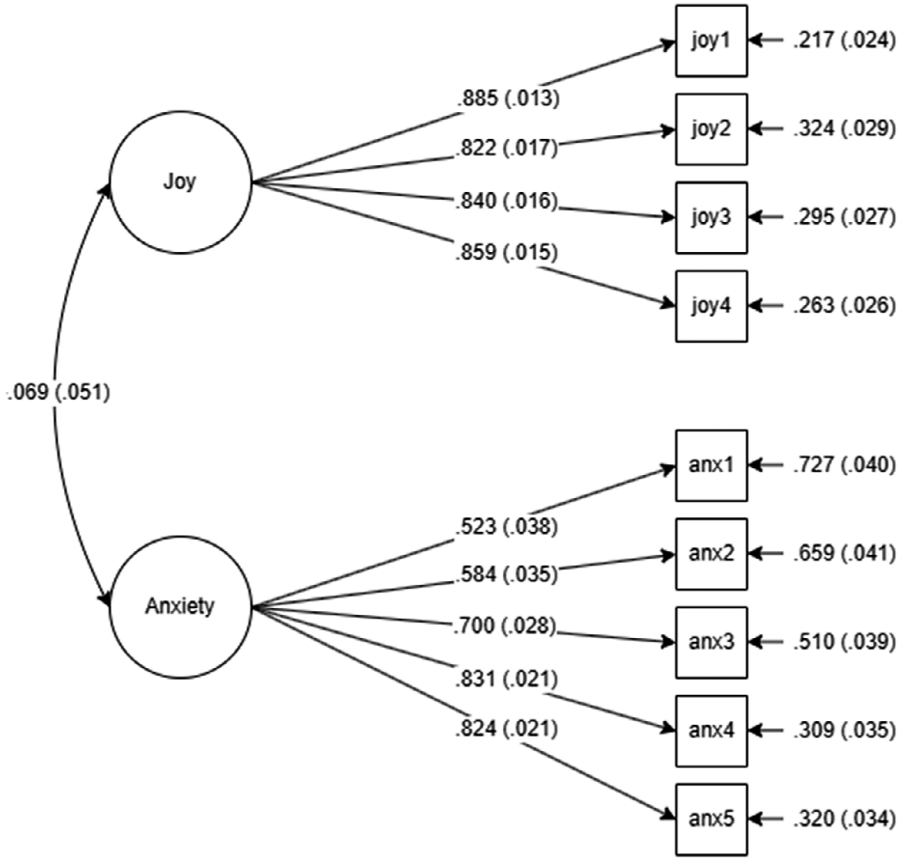

Figure 4. The results of CFA for the emotions with standardized estimates.

(f^2 = 1.58, large effect) and a negative predictor of Eager L2 Use (f^2 = .01, small effect), precisely as anticipated. In addition, Promotion (f^2 = .09, small effect) was a significant positive predictor but Ideal Own (f^2 = .08, small effect) was a significant negative predictor of Eager L2 Use. Ought Own turned out to be a significant negative predictor of Vigilant L2 Use (f^2 = .01, small effect) whereas Ought Other turned out to be a significant positive predictor of Vigilant L2 Use (f^2 = .10, small effect).

In the fourth layer, Eager L2 Use was a significant positive predictor of L2 Achievement, although the effect size was small (f^2 = .01). Similarly, Vigilant L2 Use was a significant negative predictor of L2 Achievement but again with a small effect size (f^2 = .02). These predictors together accounted for approximately 10% of the variance in L2 Achievement, indicating that while they are statistically significant, their practical impact may be limited.

Discussion

In the forthcoming discussion, we will compare the outcomes of our proposed model with those of Papi and Khajavy (2021), which can be observed in Figure 7. The

https://doi.org/10.1017/S0272263124000512 Published online by Cambridge University Press

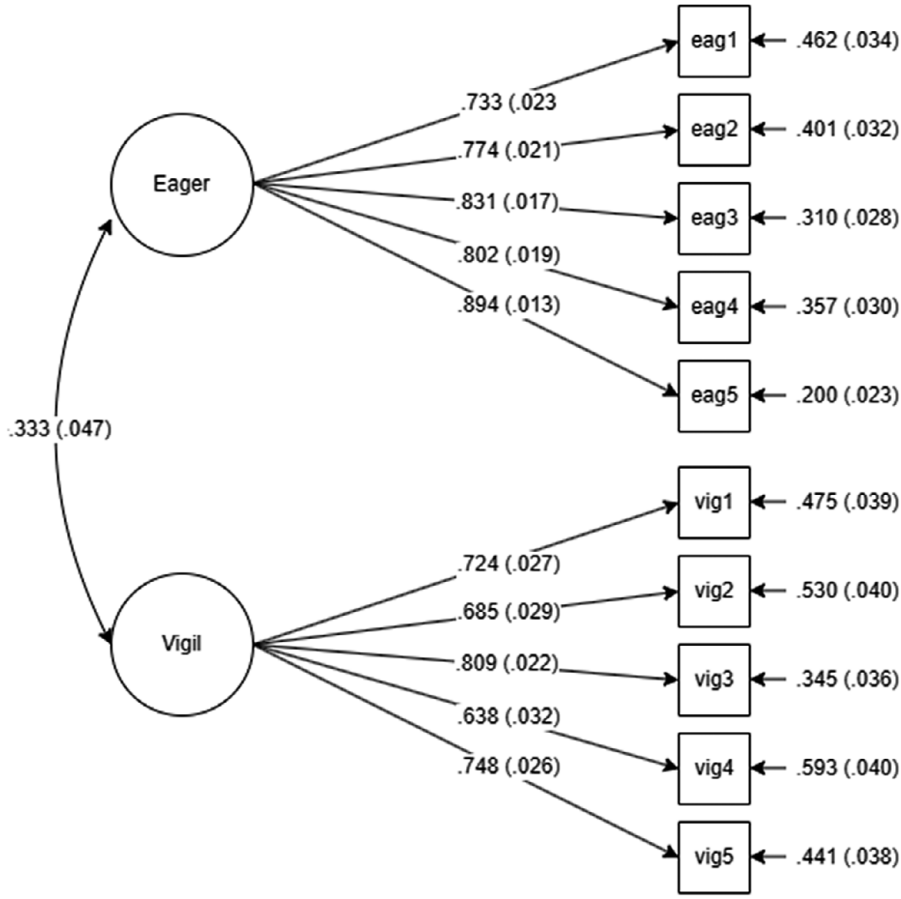

Figure 5. The results of CFA for the strategic inclinations with standardized estimates.

Table 5. Measurement invariance test of the 2X2 model for males and females

	χ2	*df*	CFI	TLI	RMSEA	SRMR	ΔCFI	ΔRMSEA	ΔSRMR
Configural	630.045	228	0.930	0.916	0.084	0.058			
Metric	635.950	240	0.931	0.922	0.082	0.060	0.001	–0.002	0.002
Scalar	652.386	252	0.930	0.924	0.080	0.060	–0.001	–0.002	0.000

similarities and dissimilarities between the two studies will be thoroughly analyzed and discussed.

First layer of paths: regulatory focus—L2 self-guides

Based on the analysis of Figure 6 and Figure 7, our study partially aligns with Papi and Khajavy's (2021) findings in the first layer of paths. The first similarity is that: promotion is a significant positive predictor of Ideal L2 Self/Own and Ideal L2 Self/Other. The results support the hypothesis that a higher chronic promotion focus

https://doi.org/10.1017/S0272263124000512 Published online by Cambridge University Press

Table 6. Paired–sample t-tests between subscales

	Mean	SDD	*t*	*df*	*p*	Cohen's *d*	95% CI
Ideal Other - Ought Own	15.985	12.87	15.985	494	0	0.718	[0.630, 0.806]
Ideal Other - Ideal Own	10.307	12.87	10.307	494	0	0.444	[0.360, 0.529]
Ideal Other - Ought Other	29.147	12.87	29.147	494	0	1.534	[1.431, 1.637]
Ought Own - Ideal Own	–5.658	12.87	–5.658	494	0	–0.274	[–0.369, –0.178]
Ought Own - Ought Other	20.663	12.87	20.663	494	0	0.816	[0.739, 0.894]
Ideal Own - Ought Other	18.388	12.87	18.388	494	0	1.090	[0.973, 1.206]
Promotion - Prevention	N/A	N/A	4.839	494	0	0.258	[0.153, 0.363]
Enjoy - Anxiety	N/A	N/A	–4.641	494	0	–0.306	[–0.435, –0.176]
Eager - Vigilant	N/A	N/A	–5.012	494	0	–0.358	[–0.498, –0.218]

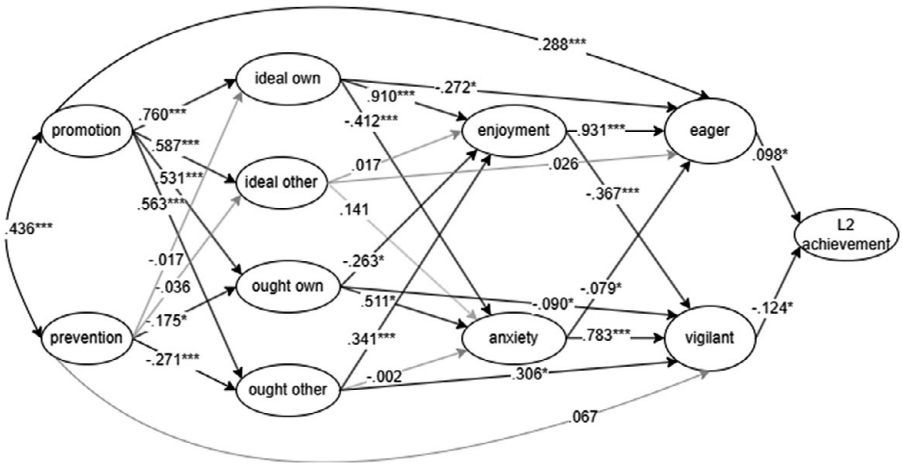

Figure 6. The model of Chinese EFL learners' motivation and achievement with standardized estimates.

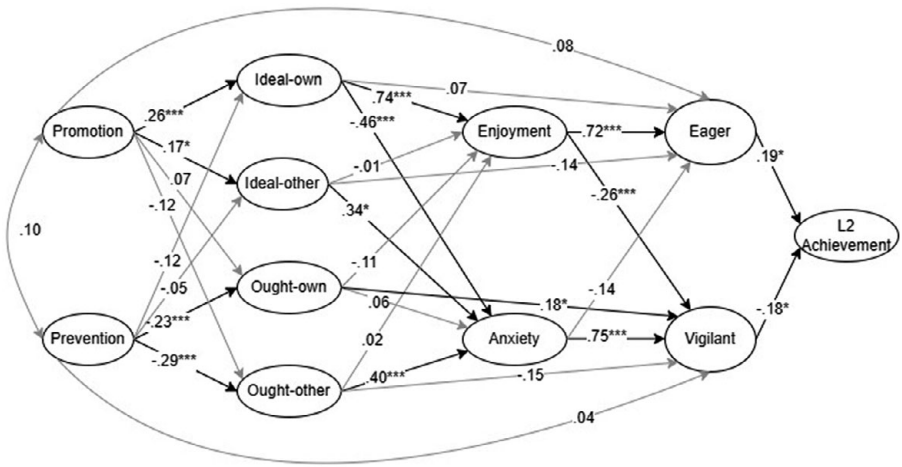

Figure 7. Papi and Khajavy's (2021) integrated model of motivation and achievement in language learning with standardized estimates.

https://doi.org/10.1017/S0272263124000512 Published online by Cambridge University Press

increases the likelihood of reaching one's ideal self. This is due to the fact that the promotion focus is regulated in a way that matches the ideal selves. As per the regulatory focus theory (Higgins, 1997), the promotion focus emphasizes accomplishments, advancement, and growth. To achieve a more desirable end state, such as an ideal state, people with promotion focuses tend to move from their current state to their ideal selves. Previous research has suggested that language learners who are motivated by promotion are more likely to perceive their current language abilities as inadequate and strive to reach their ideal levels of L2 proficiency (Papi et al., 2019; Teimouri, 2017).

The second similarity is that prevention is a significant negative predictor of Ought L2 Self/Own and Ought L2 Self/Other. Similar to Papi and Khajavy (2021) who found a negative effect of prevention on ought–own, our study reinforced such effect, which is attributable to the successful allocation of learners' time and effort toward meeting their obligations and oughts, leading to a decrease in these obligations and oughts.

It shows that prevention focus plays an important role in meeting the ought selves and fulfilling obligations, supporting the findings of Papi and Khajavy (2021) that ought selves are influenced by prevention focus. However, it is the promotion focus that enhances both the ideal selves and the ought selves. Therefore, these results suggest that while the prevention focus helps learners meet their obligations, it is the promotion focus that serves as the principal chronic regulatory focus motivating language learners to fulfill their foreign language learning ambitions.

Second layer of paths: L2 self-guides—emotions

Upon examination of Figure 6 and Figure 7, our study shares similar paths with those of Papi and Khajavy (2021). That is, Ideal L2 Self/Own in the two studies both significantly positively predicted L2 enjoyment and negatively predicted L2 anxiety. This supported the results in previous studies. For instance, Papi (2010) found that the ideal L2 self negatively predicted L2 anxiety and Teimouri (2017) found that the ideal L2 self positively predicted joy. The ideal L2 self represents a more desirable end-state than the learners' current proficiency status and is therefore a motivation with a promotion focus. Learners experience the elation–related emotion of joy when they envision reaching their promotion–focused end-state (Higgins, 1987). As a consequence, there is no surprise that Ideal L2 Self/Own has arisen as a strong predictor of L2 enjoyment. Regarding the negative association between Ideal L2 Self/Own and L2 anxiety, this finding can be attributed to the mutually inhibitory nature of regulatory focus (Higgins, 1998; Klenk, Strauman, & Higgins, 2011). This inhibition effect hinders the cooccurrence of promotion–focused and prevention–focused emotions. To further illustrate, when learners are motivated by promotion–focused motives, they tend to experience positive emotions such as enjoyment, whereas when they are motivated by prevention–focused motives, they tend to experience negative emotions such as anxiety. This may explain why the Ideal L2 Self/Own did not only positively predict L2 enjoyment but also negatively predict L2 anxiety.

The biggest difference between Papi and Khajavy's (2021) results and ours exists in the strength of the predictive power of ought L2 selves. In the current study, Ought L2 Self/Own strongly positively predicts L2 anxiety but strongly negatively predicts L2 enjoyment; Ought L2 Self/Other strongly positively predicts L2 enjoyment. However, in Papi and Khajavy's (2021) study, the predictive power of Ought L2 Self/Own on L2 anxiety was moderate and only Ought L2 Self/Other positively predicts L2 anxiety. Possible reasons could be as follows. The participants in Papi and Khajavy's (2021)

https://doi.org/10.1017/S0272263124000512 Published online by Cambridge University Press

study were college students from a public university, who were studying Persian (Farsi) as their language of instruction and had the opportunity to retake the exam if they failed. In contrast, the present study was conducted with participants taking the CET-4, which not only determines their eligibility for graduation but also carries social weight. Although participants have many chances for CET-4, failure on the CET-4 may lead to shame and peer pressure, which could explain the stronger predictive power of ought selves in this study. Furthermore, the results suggested that in Asian collectivist cultures, ought–related constructs have been found to be more prevalent than ideal–related motives (Apple, Da Silva, & Fellner, 2016; Lee, Aaker, & Gardner, 2000).

To conclude, the results of Ideal L2 Self/Own in the second layer agree with those of Papi and Khajavy (2021); however, the predictive power of ought selves did not show the same consistency. Therefore, teachers are encouraged to motivate students and cultivate their interest in L2 learning to help foster their aspirations toward achieving a higher level of L2 proficiency. This insight underscores the need for sensitivity to the context–specific impacts of social anxiety on motivational processes.

Third layer of paths: emotions—strategic inclinations

In the third layer of paths, L2 enjoyment strongly positively predicted eager L2 use and strongly negatively predicted vigilant L2 use. L2 anxiety strongly positively predicted vigilant L2 use and strongly negatively predicted eager L2 use. The four paths generally agreed with Papi and Khajavy's (2021) results. This powerfully indicated that L2 enjoyment and anxiety are indeed the connections between motives and motivated behaviors (Papi & Khajavy, 2021). In addition, it echoes Higgins' (1998) viewpoint that there are many ways in which people experience the world. Some people experience a world filled with objects and events that bring them joy or pain, while others experience objects and events that bring them relaxation or nervousness. There may also be different strategic responses resulting from this difference in how we experience the world.

Besides the similarity in the paths from emotions to strategies, the strength of the promotion focus and Ideal L2 Self/Own on eager L2 use has been reinforced. While Papi and Khajavya's (2021) study found that promotion focus moderately predicted eager L2 use, our study observed a similar trend, though the increase in predictive strength did not reach statistical significance. Our results support Higgins, Roney, Crowe, and Hymes (1994) that a promotion–focused individual is constantly searching for opportunities and making sure they do not miss any potential successes. This indirectly echoes Higgins' (1998) argument that chronic goal attainment and emotional experiences are moderated by the strength of regulatory focus.

Ideal L2 Self/Own moderately predicted eager L2 use in Papi and Khajavy's (2021). **In this study, this relationship became stronger, indicating that their Ideal Self/Own leads to more eagerness in L2 use**. One potential explanation is that Chinese students are prone to be more introverted and less outgoing. The negative relationship observed in this study might be due to the full mediation effect of enjoyment. Therefore, they might not act as the questions described that "I take advantage of every chance I get to use and practice my English in my classes." Additionally, in a university setting, students may be hesitant to speak up in front of the entire class, even if they aspire to improve their English proficiency. Instead, they may prefer to practice on their own to build their skills more privately.

https://doi.org/10.1017/S0272263124000512 Published online by Cambridge University Press

The predictive power of L2 selves on eager and vigilant L2 use differed in terms of direction and strength. The predictive power of L2 selves on eager and vigilant L2 use differed in terms of direction and strength. For instance, in this study, Ought L2 Self/Own strongly positively predicted vigilant L2 use, whereas Ought L2 Self/Other did not significantly predict vigilant L2 use. However, in Papi and Khajavy's (2021) study, Ought L2 Self/Own was found to strongly positively predict vigilant L2 use, and Ought L2 Self/Other did not emerge as a significant predictor of vigilant L2 use. This variation suggests that, depending on cultural or educational context, the influence of these self-guides on vigilant L2 use can vary significantly, highlighting that Chinese students may employ vigilant L2 use strategies to meet external expectations and avoid negative outcomes, reflecting a heightened sensitivity to the demands of others.

To conclude, the findings of the third layer are largely consistent with those of Papi and Khajavy (2021). However, the paths from regulatory focus and future L2 selves to strategic inclinations in this study do not completely align with Papi and Khajavy's (2021) results.

Fourth layer of paths: strategic inclinations—L2 achievement

In the fourth layer of the hypothesized paths, eager L2 use was found to positively predict L2 achievement, although the effect size was small. Similarly, vigilant L2 use negatively predicted L2 achievement, also with a small effect size. These findings suggest that while these predictors are significant, their overall contribution to explaining L2 achievement is relatively modest. This fully agrees with Papi and Khajavy's (2021) results, which indicate that taking eager strategies is more likely to yield higher L2 achievement, while taking vigilant strategies may have an adverse effect on L2 achievement. One potential reason for the modest effect sizes could be that both eager and vigilant L2 use primarily focus on oral L2 use, whereas our achievement test did not include an oral production measure. It is possible that including such a measure in future studies might result in stronger predictive relationships.

Future replication research

While this replication study has contributed to our understanding of the motivational mechanisms of Chinese EFL learners, several limitations should be acknowledged. Firstly, some specific items in the questionnaires, particularly those measuring aspects of motivation and language anxiety, were found to be less effective. Secondly, the study was limited to university students, omitting younger learners from middle and high schools.

With reference to the findings and limitations of the study, some recommendations for future replication studies can be made. Firstly, researchers could undertake a detailed review and modification of the questionnaires used by Papi and Khajavy (2021), especially focusing on enhancing the enjoyment and anxiety scales to better capture nuanced aspects of learner motivation. Therefore, future research could consider modifying these questionnaires to enhance their reliability and validity.

Secondly, future replication research may be conducted with students from primary, middle, or high school because these groups are pivotal in understanding the full spectrum of foreign language learning in China. Insights from these diverse educational stages could significantly enrich our understanding of EFL learners' motivational

https://doi.org/10.1017/S0272263124000512 Published online by Cambridge University Press

processes. As such, future research can collect questionnaires from EFL learners in other periods.

Thirdly, future replication research might adopt a dynamic perspective to investigate the mechanism of motivation. Employing longitudinal studies or mixed methods could provide a more comprehensive understanding of how motivation evolves and is influenced by various personal and social factors over time. Studying the dynamic changes in students' motivation over an extended period could help identify appropriate strategies to enhance L2 learning motivation.

Conclusion

This replication study investigates whether Papi and Khajavy's (2021) theoretical model, which integrates regulatory focus, 2X2 future L2 selves, emotions, strategic inclinations, and L2 achievement, applies to Chinese EFL learners. In accordance with Papi and Khajavy's (2021) research method, this study collected questionnaires and analyzed the theoretical model with CFA and SEM. The study aimed to identify similarities and differences between its findings and those of Papi and Khajavy (2021). The first research question yielded a bifurcated answer: The questionnaire items related to regulatory focus, L2 selves, and strategies were found to be applicable to Chinese university students while those related to emotions required to be improved.

Regarding the second research question, our findings support those of Papi and Khajavy (2021) in the following layers: (1) the promotion focus significantly positively predicts the ideal selves and the prevention focus significantly negatively predicts the ought selves; (2) the Ideal L2 Self/Own was found to predict L2 enjoyment positively and L2 anxiety negatively, whereas the Ought L2 Self/Own was found to predict L2 anxiety positively and L2 enjoyment negatively; (3) L2 enjoyment was found to predict eager strategies positively and vigilant strategies negatively, whereas L2 anxiety was found to predict vigilant strategies positively; (4) eager strategies was found to positively, and vigilant strategies negatively, predict L2 achievement.

Based on these findings, the present research has implications for language teachers and learners. Firstly, teachers are advised to promote a promotion–focused mindset among students. For instance, teachers can set simple tasks first so that students can experience the thrill of success and build up their self-confidence. During the process, teachers can encourage students to set achievable goals, provide regular feedback, and create a supportive learning environment that fosters a growth mindset. By implementing goal–setting activities and providing constructive feedback, teachers can help students develop the skills and mindset needed to pursue success (Dweck, 2006). Secondly, they should facilitate the development of the ideal selves in language learning. To achieve this, teachers can initially encourage students to envision their future L2 selves. Specifically, they can prompt students to articulate their ideal English selves in a clear, concrete, and attainable manner. Additionally, teachers can guide students in assessing their current language proficiency and creating an ideal self that is both realistic and achievable through diligent efforts. This approach can reduce anxiety and fear associated with L2 learning and serve as a motivating factor to help students strive toward their ideal L2 self (Liu, Yao, & Hu, 2012).

Acknowledgments. We would like to express our deepest appreciation to the handling editor, Dr. Kevin McManus, and the two anonymous reviewers from SSLA for their insightful and invaluable feedback. Their contributions have been instrumental in enhancing the quality of our work.

This research is supported by the General Project Funding of Shanghai Educational Science Research 2024 (No. C2024197), the Major Project of the National Social Science Fund of China for 2023 (No. 23&ZD320), and the Innovation Team for Comparative Linguistics and International Communication of Chinese Shanghai Normal University (No. 202106140059).

References

Al-Hoorie, A. H. (2018). The L2 Motivational Self System: A meta-analysis. *Studies in Second Language Learning and Teaching*, *8*(4), 721–754

Apple, M. T., Da Silva, D., & Fellner, T. (Eds.). (2016). L2 selves and motivations in Asian contexts (Vol. 106). Multilingual Matters.

Bollen, K. A. (1989). *Structural equations with latent variables* (Vol. 210). John Wiley & Sons.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159.

Dewaele, J. M., & MacIntyre, P. D. (2016). Foreign language enjoyment and foreign language classroom anxiety: The right and left feet of the language learner. In MacIntyre, P. D., Gregersen, T., & Mercer, S. (Eds.), *Positive psychology in SLA* (pp. 215–236). Multilingual Matters.

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. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 9–42). Multilingual Matters.

Dörnyei, Z., & Chan, L. (2013). Motivation and vision: An analysis of future L2 self images, sensory styles, and imagery capacity across two target languages. *Language Learning*, *63*, 437–462.

Dörnyei, Z., & Ushioda, E. (Eds.). (2009). *Motivation, language identity and the L2 self* (Vol. 36). Multilingual Matters.

Dweck, C. S. (2006). *Mindset: The new psychology of success*. Random House.

Fang, F. (2018). Ideology and identity debate of English in China: Past, present and future. *Asian Englishes*, *20* (1), 15–26.

Feng, L., & Papi, M. (2020). Persistence in language learning: The role of grit and future self-guides. *Learning and Individual Differences*, *81*, 101904.

Gil, J., & Adamson, B. (2011). The English language in Mainland China: A sociolinguistic profile. In A. Feng (Ed.), *English language education across Greater China* (pp. 23–45). Bristol: Multilingual Matters.

Henry, A., Korp, H., Sundqvist, P., & Thorsen, C. (2018). Motivational strategies and the reframing of English: Activity design and challenges for teachers in contexts of extensive extramural encounters. *Tesol Quarterly*, *52*(2), 247–273.

Henry, A., & Liu, M. (2023). Can L2 motivation be modelled as a self-system? A critical assessment. *System*, *119*, 103158.

Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review*, *94*, 319–340.

Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist*, *52*, 1280–1300.

Higgins, E. T. (1998). Promotion and prevention: Regulatory focus as a motivational principle. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 30, pp. 1–46). Academic Press.

Higgins, E. T., Roney, C. J. R., Crowe, E., & Hymes, C. (1994). Ideal versus ought predilections for approach and avoidance: Distinct self-regulatory systems. *Journal of Personality and Social Psychology*, *66*(2), 276–286.

Hiver, P., & Al-Hoorie, A. H. (2020). Reexamining the role of vision in second language motivation: A preregistered conceptual replication of You, Dörnyei, and Csizér (2016). *Language Learning*, *70*(1), 48–102.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55.

Klenk, M. M., Strauman, T. J., & Higgins, E. T. (2011). Regulatory focus and anxiety: A self-regulatory model of GAD-depression comorbidity. *Personality and Individual Differences*, *50*, 935–943.

Kermad, A. A. D., (2018). *Speaker and listener variability in the perception of non-native speech (Doctoral dissertation)*. Northern Arizona University

Korkmaz, S., Goksuluk, D., & Zararsiz, G. (2014). MVN: An R package for assessing multivariate normality. *The R Journal*, *6*, 151–162.

Lee, A. Y., Aaker, J. L., & Gardner, W. L. (2000). The pleasures and pains of distinct self-construals: the role of interdependence in regulatory focus. *Journal of personality and social psychology*, *78*(6), 1122.

Liu, Z., Yao, X. J., & Hu, S. F. (2012). A structural analysis of college students' L2 selves, anxiety, and motivated learning behavior. *Foreign Language World*, *6*, 28–37.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological methods*, *1*(2), 130.

Mahdavy, B. (2020). Ideal L2 self in the Expanding Circle: The case of English language learners in Iran. *International Journal of Applied Linguistics*, *30*(2), 280–292.

Moharami, M., & Daneshfar, S. (2021). The impacts of learning English on Iranians' everyday life: An ethnographic example from Piranshahr. *Issues in Educational Research*, *31*(4), 1156–1174.

Moskovsky, C., Alrabai, F., Paolini, S., & Ratcheva, S. (2013). The effects of teachers' motivational strategies on learners' motivation: A controlled investigation of second language acquisition. *Language learning*, *63*(1), 34–62.

Muthén, L.K. and Muthén, B.O. (1998-2017). *Mplus user's guide*. Eighth Edition. Muthén & Muthén.

Papi, M. (2010). The L2 motivational self system, L2 anxiety, and motivated behavior: A structural equation modeling approach. *System*, *38*, 467–479.

Papi, M., & Khajavy, G. H. (2021). Motivational mechanisms underlying second language achievement: A regulatory focus perspective. *Language Learning*, *71*(2), 537–572.

Papi, M., Bondarenko, A. V., Mansouri, S., Feng, L., & Jiang, C. (2019). Rethinking L2 motivation research: The 2X2 model of L2 self-guides. *Studies in Second Language Acquisition*, *41*, 337–361.

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

Porte, G., & McManus, K. (2019). *Doing replication research in applied linguistics*. Routledge.

Preacher, K. J., & Coffman, D. L. (2006). *Computing power and minimum sample size for RMSEA [Computer software]*. Available from http://quantpsy.org/

Saito, K., Dewaele, J. M., Abe, M., & In'nami, Y. (2018). Motivation, emotion, learning experience, and second language comprehensibility development in classroom settings: A cross-sectional and longitudinal study. *Language Learning*, *68*(3), 709–743.

Teimouri, Y. (2017). L2 selves, emotions, and motivated behaviors. *Studies in Second Language Acquisition*, *39*, 691–709.

Yang, H.Z., & Weir C. J. (1998). *Validation study of the National College English Test*. Shanghai Foreign Language Education Press.

Yang, R. (2014). China's removal of English from gaokao. *I*nternational Higher Education, *75*, 12–13.

Zhou, M., & Papi, M., (2023). *The role of future L2 selves in L2 speech development: A longitudinal study in an instructional setting*. System, 103156.

Cite this article: Wang, L., & Sun, X. (2024). Reexamining the role of regulatory focus in second language achievement: An approximate replication of Papi and Khajavy (2021). *Studies in Second Language Acquisition,* 46: 1515–1536. https://doi.org/10.1017/S0272263124000512