Abjad Journal of Humanities & Education, 2026
Vol. 4. No. 2. pp. 49 – 91, https://doi.org/10.62079/abjad.v4i2.146

CENTRE FOR LITERARY AND CULTURAL STUDIES
CLCS

RESEARCH ARTICLE

Intercultural Sensitivity, L2 Motivation, and Willingness to Communicate in EFL: A Systematic Review and Meta-Analysis

Ahmad Sugianto*, Sherina Gema Millennium, Rojab Siti Rodliyah, Eri Kurniawan

Received: June 23, 2026 / Revised: July 13, 2026 / Accepted: July 27, 2026

Abstract

This study aims to synthesise evidence on intercultural sensitivity, L2 motivation, and willingness to communicate (WTC) in English as a foreign language (EFL) contexts to examine their contributions to the social psychology of communication. A systematic literature review (SLR) combined with meta-analysis was conducted in accordance with the PRISMA guidelines. Twenty-nine studies between 2005 and 2026 were identified in the Scopus database based on predefined inclusion criteria, which encompassed English-language and open-access publications. The selected studies were analysed using thematic analysis to scrutinise the conceptual trends and meta-analysis to estimate the magnitude and direction of the relationships and intervention effects. The qualitative synthesis suggests that the dynamic interplay among psychological, social, and intercultural factors shapes communication in EFL contexts. The meta-analysis of cross-sectional studies showed a significant moderate positive relationship between L2 motivation and WTC (random-effects r=0.449, 95% CI: 0.361-0.530, p< 0.001) and a significant positive relationship between intercultural sensitivity/intercultural competence and L2 motivation (random-effects r=0.544, 95% CI:0.009-0.837, p=0.047). Meta-analyses of intervention studies report moderate-to-large positive effects of intercultural sensitivity (g=0.779, 95% CI: 0.551-1.008) and L2 motivation (g=0.495, 95% CI: 0.242-0.747). A large pooled effect was also reported for WTC (g= 2.637, 95% CI: 0.628-4.646); however, this estimate should be interpreted with caution, as it is based on only very limited studies with substantial between-study heterogeneity. Taken together, the findings suggest that communication in EFL extends beyond linguistic competence, encompassing interconnected psychological and intercultural processes. Nevertheless, the findings should be interpreted in light of the review's methodological constraints, encompassing the exclusive use of the Scopus database and the

* Ahmad Sugianto, Rojab Siti Rodliyah, and Eri Kurniawan, Universitas Pendidikan Indonesia, Indonesia. Sherina Gema Millennium, IKIP Siliwangi, Indonesia. Corresponding author: ahmadsugianto@upi.edu

restriction to open-access publications, which may have limited the coverage of relevant studies. Future studies are necessarily required to scrutinise these constructs within integrated theoretical frameworks and across more diverse educational and sociocultural contexts.

Keywords social psychology of communication; EFL learners; intercultural sensitivity; L2 motivation; meta-analysis; systematic literature review (SLR); willingness to communicate (WTC)

Abstrak

Penelitian ini bertujuan untuk menyintesis bukti mengenai sensitivitas antarbudaya, motivasi bahasa kedua (L2), dan kesediaan untuk berkomunikasi (WTC) dalam konteks bahasa Inggris sebagai bahasa asing (EFL) guna mengkaji kontribusinya terhadap psikologi sosial komunikasi. Tinjauan pustaka sistematis (SLR) yang dikombinasikan dengan meta-analisis dilakukan sesuai dengan pedoman PRISMA. Sebanyak dua puluh sembilan penelitian yang diterbitkan antara tahun 2005 dan 2026 diidentifikasi melalui basis data Scopus berdasarkan kriteria inklusi yang telah ditetapkan, yang mencakup publikasi berbahasa Inggris dan akses terbuka. Penelitian-penelitian terpilih dianalisis menggunakan analisis tematik untuk mencermati tren konseptual serta meta-analisis untuk mengestimasi besaran dan arah hubungan serta efek intervensi. Sintesis kualitatif menunjukkan bahwa interaksi dinamis antara faktor psikologis, sosial, dan antarbudaya membentuk komunikasi dalam konteks EFL. Meta-analisis terhadap studi potong-lintang (cross-sectional) menunjukkan adanya hubungan positif yang signifikan dengan kekuatan sedang antara motivasi L2 dan WTC (efek acak r=0,449; IK 95%: 0,361–0,530; p<0,001) serta hubungan positif yang signifikan antara sensitivitas antarbudaya/kompetensi antarbudaya dan motivasi L2 (efek acak r=0,544; IK 95%: 0,009–0,837; p=0,047). Meta-analisis terhadap studi intervensi melaporkan adanya efek positif berskala sedang hingga besar terkait sensitivitas antarbudaya (g=0,779; IK 95%: 0,551–1,008) dan motivasi L2 (g=0,495; IK 95%: 0,242–0,747). Efek gabungan yang besar juga dilaporkan untuk WTC (g=2,637; IK 95%: 0,628–4,646); namun, estimasi ini perlu ditafsirkan secara hati-hati karena didasarkan pada jumlah penelitian yang sangat terbatas dengan heterogenitas yang substansial antarpenelitian. Secara keseluruhan, temuan ini mengindikasikan bahwa komunikasi dalam konteks EFL melampaui sekadar kompetensi linguistik, melainkan mencakup proses psikologis dan antarbudaya yang saling terkait. Kendati demikian, temuan ini perlu ditafsirkan dengan mempertimbangkan keterbatasan metodologis tinjauan tersebut—termasuk penggunaan eksklusif basis data Scopus dan pembatasan pada publikasi akses terbuka—yang mungkin telah membatasi cakupan studi-studi yang relevan. Studi lanjutan sangat diperlukan untuk mengkaji secara mendalam konstruk-konstruk ini dalam kerangka teoretis yang terintegrasi serta di berbagai konteks pendidikan dan sosiokultural yang lebih beragam.

Kata kunci *psikologi sosial komunikasi; pembelajar EFL; kepekaan antarbudaya; motivasi bahasa kedua (L2); meta-analisis; tinjauan pustaka sistematis (SLR); kesediaan untuk berkomunikasi (WTC)*

Introduction

Communication in EFL extends beyond learners' mastery of vocabulary or grammar. A growing body of research suggests that successful communication also involves psychological, social, and intercultural processes that contribute to learners' readiness, motivation, and confidence to use English in diverse communicative situations. Accordingly, communication in EFL should be construed not merely as linguistic performance but as a socially and psychologically mediated process through which learners perceive themselves, understand their interlocutors, and respond to the broader communicative context before engaging in interactions. Previous studies have consistently suggested that learners' WTC, L2 motivation, and intercultural orientation play important roles in promoting communicative engagement across various contexts encompassing traditional face-to-face classrooms and online or technology-enhanced learning settings (Aladini, et al., 2024; Balouchi & Samad, 2021; Cao, 2022) Notwithstanding these advances, the three constructs have predominantly been scrutinised in isolation or in pairs, limiting a comprehensive understanding of how they jointly contribute to communication in EFL contexts. This gap suggests a need for a systematic synthesis of the evidence on intercultural sensitivity, L2 motivation, and WTC.

Moreover, social-psychological constructs, apart from linguistic competence, are essential to communication. As MacIntyre et al. (1998) reveal and advocate, communication in L2 is mainly concerned with various variables encompassing communication behaviour, situated intention, situated antecedents, motivational propensities, affective-cognitive context, and social and individual context. The notion above shows that there are interplays among social psychological variables in L2 communication. In a similar vein, Gardner (1985), through his Socio-Educational Model, explains that both students' linguistic outcomes (such as L2 proficiency) and non-linguistic ones (such as attitudes or values) are shaped by language training or experiences, involving psychological variables within cultural contexts. This supporting notion strongly asserts that psychological variables, along with cultural and cognitive variables, are interdependent in L2 communication.

Notwithstanding the theories advocated above, conducting the present study in the recent EFL contexts is deemed crucial. A substantial body of literature and evidence has shown the importance of the constructs comprising intercultural sensitivity, motivation, and willingness to communicate in contributing to EFL learning. However, they are commonly scrutinised in partial or fragmented fashions. For instance, Sugianto et al. (2022) scrutinised students' intercultural sensitivity with respect to extroversion, Feng et

al. (2023) investigated the use of a particular pedagogical strategies, such as mind-mapping, in relation to students' motivation and willingness to communicate, and Joe et al. (2017) conducted a study on motivation and willingness to communicate along with other variables comprising classroom social climate and achievement.

Given the aforementioned problems, further research is required to develop an in-depth understanding of the three constructs in an integrated manner. To that end, a systematic review might be beneficial in mapping the theories and findings related to the constructs under scrutiny. In detail, the use of systematic reviews provides a means for interpretation, synthesis, judgment, and the identification of research gaps and novelty in previous related studies (Patole, 2021). To that end, the present study is guided by the following research questions (RQs): 1) How is social psychology of communication conceptualised in EFL literature through the constructs of intercultural sensitivity, motivation, and willingness to communicate?2) What unresolved conceptual issues and research gaps emerge from the literature, and how can they inform the development of an integrated systematic review framework on social psychology of communication in EFL? 3) What is the magnitude of the effects and relationships among intercultural sensitivity, L2 motivation, and WTC in EFL contexts?

Methods

The present study employed a systematic literature review (SLR) combined with meta-analysis to synthesise qualitative and quantitative evidence on intercultural sensitivity, L2 motivation, and WTC in EFL contexts. The thematic synthesis was conducted to identify conceptual patterns, research trends, and gaps in the literature. At the same time, the meta-analysis was used to quantitatively estimate the magnitude and direction of the relationship and intervention effects reported across eligible studies. The review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines. Although the review protocol was not prospectively registered, a predefined protocol specifying the research questions, eligibility criteria, search strategy, data extraction, and synthesis procedures was developed and followed throughout the study to ensure methodological transparency and consistency (Cheung & Vijayakumar, 2016).

To conduct the SLR study, rigorous, systematic protocols and guidelines were utilised. In this regard, the PRISMA guidelines (Page et al., 2021), which guided the identification, screening, eligibility assessment, and inclusion processes, were employed. To facilitate study management throughout the review, the Watase Uake platform (Wahyudi, 2018) was used to organise the review workflow, while Napkin

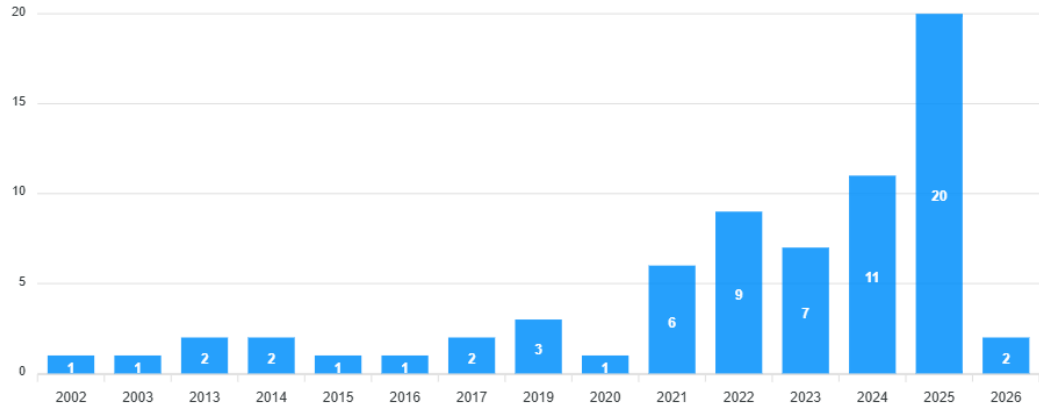
(Sharma & Scholler, 2026) and Tableau (Chabot et al., 2026) were employed to generate visual summaries of the findings. Methodological quality was independently assessed by two reviewers using the Joanna Briggs Institute (JBI) Critical Appraisal Checklists (Joanna Briggs Institute, 2026), with the appropriate checklist selected for each study design. Studies that met the predefined quality thresholds of 6/8 for cross-sectional designs and 7/9 for quasi-experimental designs were retained for synthesis. Any discrepancies between the reviewers were resolved through discussion until consensus was achieved. A standardised data extraction form was used to record data systematically. A standardised data extraction form was used to systematically record study characteristics, participant information, research design, instruments, key findings, and statistical data required for the thematic synthesis and meta-analysis. These procedures were adopted to promote the transparency, methodological rigour, and reproducibility of the review.

Before the literature search was conducted, explicit inclusion criteria were established to ensure a transparent and systematic study selection process. Eligible studies should meet the following criteria: 1) be indexed in the Scopus database; 2) be published between 2005 and 2026; 3) be written in English; 4) investigate one or more the focal constructs, namely intercultural sensitivity/intercultural communicative competence, L2 motivation, and WTC in EFL context; 5) involve learners in EFL settings; 6) be available as open-access publications to ensure full-text accessibility for quality appraisal and data extraction. Both cross-sectional and quasi-experimental studies were included because they provided evidence relevant to the research questions. For the meta-analysis, studies were additionally required to report sufficient quantitative information (e.g., correlation coefficients or statistics convertible to effect sizes, sample sizes, and measures of variance) to enable effect size estimation. The restriction to Scopus-indexed and open-access publications was adopted to ensure the quality, accessibility, and consistency of the evidence-based. However, this criterion may have excluded relevant studies indexed by other databases or published under subscription access, and this limitation is acknowledged when interpreting the findings.

As the inclusion criteria had already been set, the process moved on to the identification stage. The Boolean Operators involving “OR/AND” and a combination of words, e.g. (TITLE-ABS-KEY (“intercultural sensitivity” OR “intercultural competence” OR “intercultural awareness”) AND TITLE-ABS-KEY (“motivation” OR “L2 motivation” OR “second language motivation” OR “English learning motivation”) AND TITLE-ABS-KEY (“willingness to communicate” OR “WTC”) AND TITLE-ABS-KEY (“EFL” OR “English as a Foreign Language”) AND TITLE-ABS-KEY (“Young Learngers” OR “Primary School Students”) AND TITLE-ABS-KEY (“Indonesia”)) were input to Watase Uake platform. The search on the Scopus database

was conducted on 6-7 March 2026. This initial stage generated 69 articles (see Pictures 1 and 2). Then, as the articles were inspected, 6 duplicate articles were removed: 2 were outside the proposed publication years, and 5 did not meet the Scopus Q1, Q2, Q3, or Q4 tiers. In the final identification stage, 56 articles remained for review (see Picture 2).

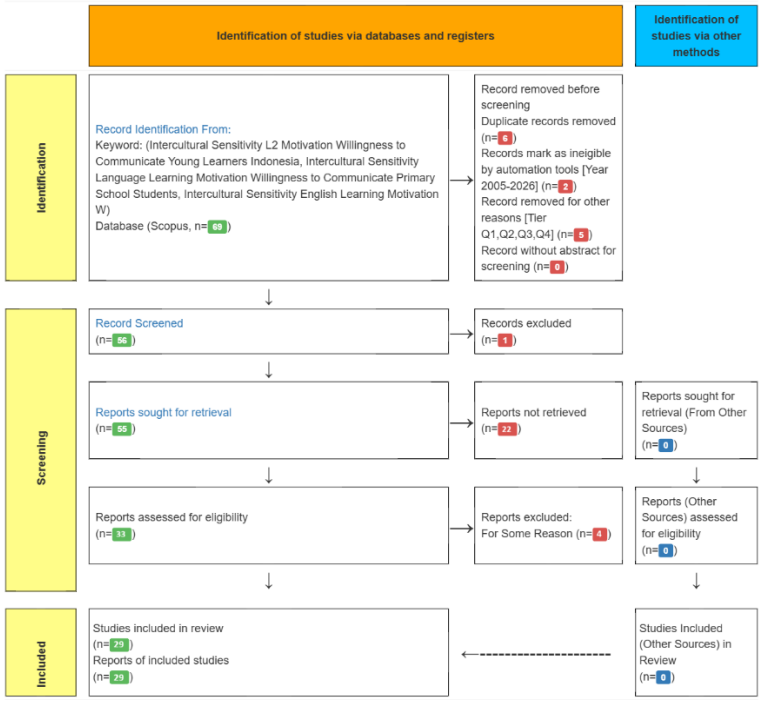
It is noteworthy that, to maximise the available evidence while maintaining conceptual consistency, the present study includes articles that scrutinise intercultural sensitivity and closely related constructs, such as intercultural competence and global/intercultural competence. The rationale for including them was their shared theoretical foundation in intercultural communication research. In this respect, as Chen and Starosta (2000) point out, intercultural sensitivity constitutes the affective dimension of broader frameworks of intercultural competence. Moreover, during the screening stage, articles were included if the constructs explicitly reflected or were related to the constructs or dimensions of intercultural sensitivity, such as learners' intercultural awareness, openness, respect for cultural differences, empathy, or intercultural communication abilities.



Picture 1. Keyword Search Result in the Initial Stage of Identification by Year.

After the identification stage was completed, the documents proceeded to the subsequent stage, namely the screening stage. In the initial stage of screening, the bibliographical information, including titles, abstracts, keywords, source details, and topics, was evaluated to determine whether the articles aligned with the constructs under scrutiny. In this first screening stage, 1 article was removed because its topic was in management education. In the second screening stage, the articles were assessed for availability. At this stage, 22 articles could not be retrieved because they were not published in open-access journals. In the second screening stage, 33 articles remained. The 33 remaining articles continued to the last screening stage. In the last screening stage, the articles were

evaluated using JBI Critical Appraisal checklists (Joanna Briggs Institute, 2026). The checklists aligned with the methodological design of each article encompassing qualitative studies, conceptually oriented works, analytical cross-sectional (ACS) studies, and quasi-experimental studies. Using the checklists, the writers evaluated the articles item by item to determine whether their designs, sampling, data collection, analyses, and interpretations adequately met the criteria for inclusion in the final synthesis or review process. The writers agreed that 4 articles were removed. Therefore, only 29 articles remained eligible to undergo the final process of rigorous review (see Pictures 2 and 3).

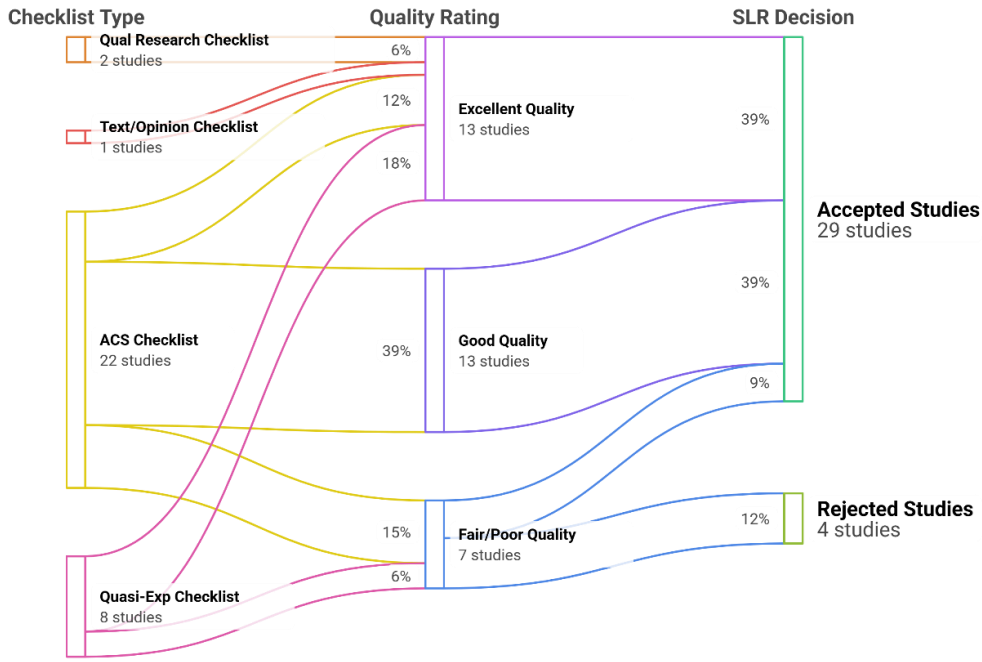


Picture 2. Report of the PRISMA Protocol

The articles were reviewed and analysed using both qualitative and quantitative approaches. The qualitative analysis involved thematic analysis of all included studies, while the quantitative analysis involved meta-analysis of empirical studies employing cross-sectional and quasi-experimental designs. Pearson's correlation coefficient (r) was used as the effect size for correlational studies, whereas Hedge's g was applied for quasi-experimental studies. The effect sizes were analysed using Comprehensive Meta-Analysis (CMA) Version 3.0, which applies Fisher's z-transformation to correlation coefficients during pooling. Effect sizes were interpreted based on (Cohen, 1988) benchmarks, where r values of 0.10, 0.30, and 0.50 and Hedge's g values of 0.20, 0.50,

and 0.80 indicate small, medium, and large effects, respectively (Borenstein et al., 2009, 2021; Lipsey & Wilson, 2001).

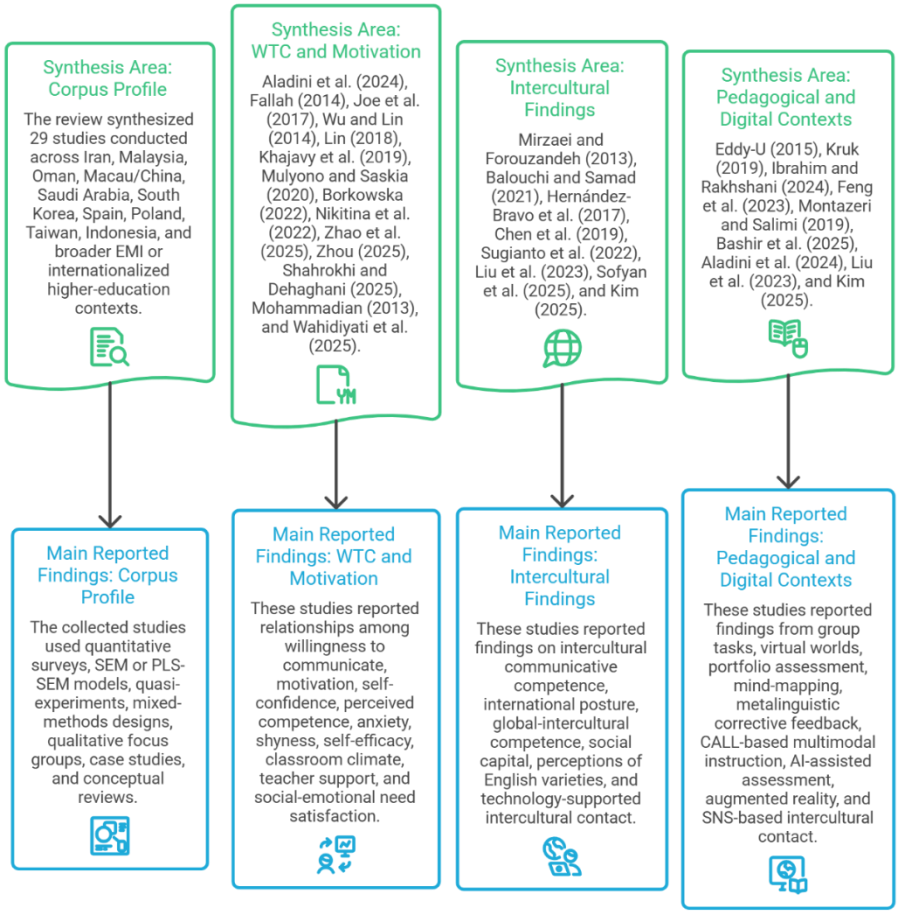
A random-effects model was employed due to variation among studies in participants, contexts, and measurement instruments. Heterogeneity was assessed using Cochran's Q, I2, and τ2 statistics (Borenstein et al., 2021; Higgins, Thomson, Deeks, & Altman, 2003). When multiple effect sizes were reported in a single study, the effect size most closely aligned with the targeted relationship was selected to maintain independence among observations. Publication bias was examined using funnel plots and Egger's test when at least 10 studies were available, while trim-and-fill and fail-safe N analyses were considered as supplementary assessments for a smaller number of studies (Duval & Tweedie, 2000a, 2000b; Higgins et al., 2019; Orwin, 1983; Rosenthal, 1979). CMA version 3.3.070 was utilised due to its comprehensive functions for effect-size calculation, heterogeneity assessment, and publication bias analysis (Borenstein et al., 2014). In addition to the use of CMA version 3.3.070, R+ metafor software was employed as a complementary tool to conduct statistical analyses, including prediction intervals, leave-one-out, influence analysis, subgroup analysis, and publication bias (Viechtbauer, 2010).



Picture 3. JBI Critical Appraisal Result.

Results

After conducting the rigorous systematic protocols, some results were obtained. The qualitative results, obtained through thematic analysis, were divided into four areas. They comprise a corpus profile, willingness to communicate with respect to motivation, an area focusing on intercultural research, and an area that synthesises various contexts, from pedagogical to digital ones. They are illustrated and summarised in detail in Picture 4 below.



Picture 4. Synthesis Area Results

 DOI: 10.62079/abjad.v4i2.146

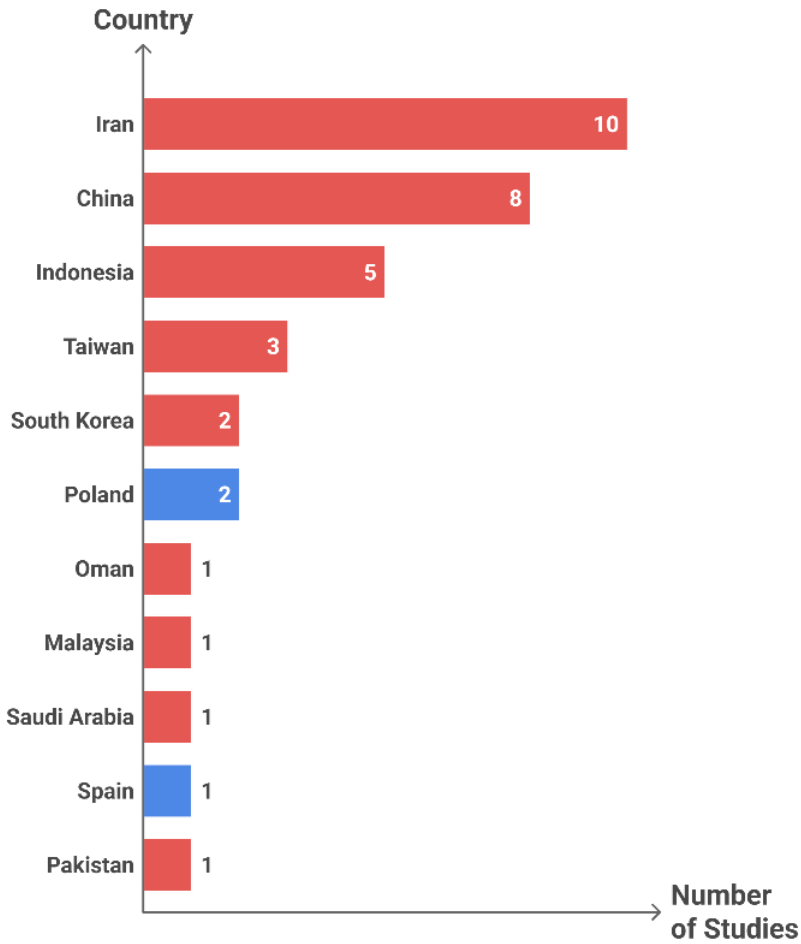
Corpus Profile

The present study identified several typical features among the 29 articles reviewed. First, Pictures 5 and 6 show that the geographical settings were mainly distributed into two continents, namely Asia (coloured in red) and Europe (coloured in blue). The former surpassed the latter in total number of studies. In this regard, 26 studies were conducted in Asia, whilst 3 were conducted in European countries. Moreover, the contributor countries from Asia are Iran, China, Indonesia, Taiwan, South Korea, Oman, Malaysia, Saudi Arabia, and Pakistan, ranked from the most to the fewest contributors. Additionally, two European countries were represented: Poland (2 studies) and Spain (1 study).

In addition, the present study found diversity in methodological profiles. In this regard, the methodological profile comprise quantitative correlational and structural equation modelling (SEM) designs stemming from the studies conducted by Sugianto et al. (2022), Aladini et al. (2024), Balouchi and Samad (2021), Fallah (2014), Joe et al. (2017), Lin (2018), Khajavy et al. (2019), Zhao et al. (2025), Zhou (2025), Borkowska (2022), Nikitina et al. (2022), Mulyono and Saskia (2020), Wahidiyati et al. (2025), and Wu and Lin, (2014). In addition to correlational and SEM design, quasi-experimental intervention studies were also reported by such authors as Ibrahim and Rakhshani (2024), Feng et al. (2023), Hernández-Bravo et al. (2017), Montazeri and Salimi (2019), Bashir et al. (2025), Liu et al. (2023), and Kim (2025). Moreover, other studies, such as those by Eddy-U (2015) and Kruk (2019), utilised qualitative research, and Cao (2022) employed a conceptual review. Detailed characteristics encompassing the research contexts, countries, methodological designs, participant profiles, investigated variables, measurement instruments, extracted effect sizes, and decisions regarding inclusion in the meta-analysis are provided in the supplementary material.

Picture 5. Geographical Representation of the Studies by Country

Moreover, the present study reveals typical pedagogical contexts in which the 29 studies were conducted. The research participants ranged from primary school to higher education (including elderly learners). In this regard, the higher education or university level became the most common context found. In addition, various pedagogical settings were found, ranging from formal classroom settings to (informal) digital or hybrid communication settings (e.g., AI-assisted language assessment, virtual worlds, augmented reality, and social networking platforms). To sum up, the present study reveals that the data derive from different ages, learning modes, instructional designs, and sociocultural contexts. The corpus profile of the present study, therefore, offers a descriptive basis for heuristic contexts within the social psychology of communication, particularly those involving intercultural sensitivity, motivation, and willingness to communicate. The results of the interplay among these three constructs, along with their pedagogical and digital contexts, will be elaborated in the next sub-section.



Picture 6. Numerical Distribution of the Studies by Country

Motivation and Willingness to Communicate

The thematic synthesis demonstrates a positive pattern, indicating that L2 motivation serves as an antecedent of WTC in EFL contexts. In this respect, higher levels of motivation are generally associated with greater WTC, as motivated learners tend to demonstrate stronger engagement, persistence, and willingness to invest effort in communicative activities. However, this relationship was not solely direct; psychological, social, and contextual factors also shaped it. For instance, self-confidence and reduced anxiety appeared to facilitate the translation of motivational dispositions into actual willingness to communicate (Fallah, 2014; Mulyono & Saskia, 2020; C. P. Wu & Lin, 2014). Similarly, supportive classroom environments, teacher support, and positive social climates strengthen learners' motivation and encourage their participation in English communication (Aladini et al., 2024; Borkowska, 2022; Joe et al., 2017; Nikitina et al., 2022). Other factors, which encompass intercultural affective

orientations, language learning beliefs, metacognitive strategies, social-emotional needs, L2 grit, and academic outcomes, further contributed to promoting learners' motivational pathways toward WTC (Khajavy et al., 2019; Lin, 2018; Shahrokhi & Dehaghani, 2025; Wahidiyati et al., 2025; Zhao et al., 2025; Zhou, 2025).

Nevertheless, the relationship between motivation and WTC may vary depending on learners' individual and contextual conditions. In contrast to the dominant pattern, Mohammadian (2013) found a non-significant relationship between motivation and WTC, suggesting that motivation alone may not automatically yield communication readiness when learners encounter barriers such as limited confidence, insufficient interactional opportunities, or contextual constraints. This finding indicates that WTC is a situated construct influenced by the interaction between motivational factors and learners' affective, social, and environmental conditions.

Intercultural Findings

The following section focuses on studies that yield intercultural findings, as well as on motivation or willingness to communicate and other relevant constructs, in EFL contexts. Most of the studies come to a consensus that intercultural sensitivity or other closely related terms, such as intercultural communicative competence or international orientation were found to have an association with or affected by other constructs which encompass motivation (Mirzaei & Forouzandeh, 2013; Sofyan et al., 2025), willingness to communicate (Balouchi & Samad, 2021; Chen et al., 2019), extroversion (Sugianto et al., 2022), intercultural based learning programme (Hernández-Bravo et al., 2017), digital learning augmented reality (Liu et al., 2023), and online social media platform (Kim, 2025). Collectively, these studies outline and demonstrate the interplay among intercultural variables in various forms, including competence, sensitivity, posture, global orientation, and other constructs encompassing motivation, willingness to communicate, and various digitally mediated intercultural contacts.

Pedagogy and Digital Contexts

A group of reviewed articles reveals that the three core constructs were scrutinised across diverse pedagogical strategies and contexts. In this regard, the articles report several pedagogical strategies to promote intercultural sensitivity or competence, motivation, or willingness to communicate. For instance, motivation and willingness to communicate are promoted by using specific activities, including group tasks (Eddy-U, 2015), mind-mapping (Feng et al., 2023), and assessments that incorporate metacognitive corrective feedback (Montazeri & Salimi, 2019) and portfolio assessment (Ibrahim & Rakhshani, 2024). Also, other favourable means for promoting motivation, willingness to communicate or intercultural competence were through using digital settings for the

interactions (Mulyono & Saskia, 2020), CALL-based multimodal pedagogy (Bashir et al., 2025), augmented reality (Liu et al., 2023), AI-assisted language assessment (Aladini et al., 2024), a digital platform such as Second Life virtual world (Kruk, 2019) and SNS (Kim, 2025). Overall, these articles reported that several means and pedagogical contexts significantly affect and foster students' intercultural competence, motivation, or willingness to communicate.

Statistical Findings of Motivation – WTC (Cross Sectional Studies)

The meta-analysis of 13 cross-sectional studies examined the association between motivation and willingness to communicate (WTC). Table 1 below shows the detailed key statistics outputs comprising the lists of the studies, number of sample (N), the strength and direction of the relationship indicated by the correlation coefficient (r), 95% confidence interval (CI) around r which reveals the true population effect might fall, Z value with the p-value to test the null hypothesis (no significant relationship between motivation and WTC), and Weight (W) both for fixed (F) and random (R) models.

Table 1. Statistics Outputs of Motivation—WTC

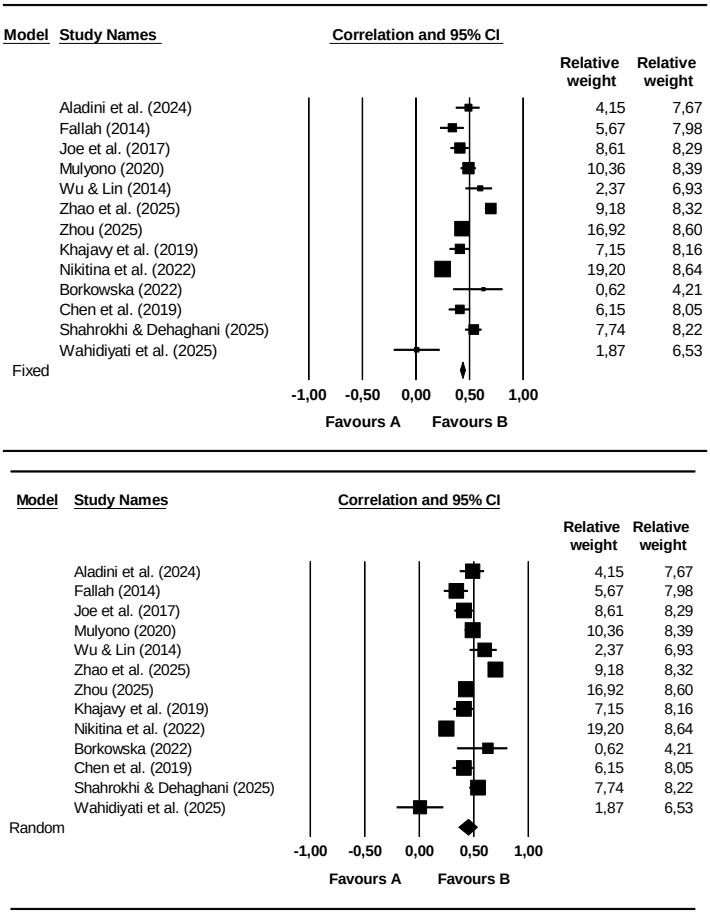
No.	Study Names	N	r	Lower Limit	Upper Limit	Z	P	W (R)	W (F)
1.	Aladini et al. (2024)	185	0.49	0.372	0.592	7.232	<0.001	4.15	7.67
2.	Fallah (2014)	252	0.34	0.226	0.445	5.587	<0.001	5.67	7.98
3.	Joe et al. (2017)	381	0.41	0.323	0.490	8.469	<0.001	8.61	8.29
4.	Mulyono and Saskia (2020)	458	0.49	0.417	0.557	11.435	<0.001	10.36	8.39

 DOI: 10.62079/abjad.v4i2.146

5.	Wu and Lin (2014)	107	0.6	0.463	0.709	7.069	<0.001	2.37	6.93
6.	Zhao et al. (2025)	406	0.699	0.646	0.746	17.372	<0.001	9.18	8.32
7.	Zhou (2025)	746	0.43	0.370	0.487	12.536	<0.001	16.92	8.60
8.	Khajavy et al. (2019)	317	0.41	0.314	0.498	7.719	<0.001	7.15	8.16
9.	Nikitina et al. (2022)	846	0.249	0.185	0.311	7.385	<0.001	19.20	8.64
10.	Borkowska (2022)	30	0.63	0.349	0.807	3.853	<0.001	0.62	4.21
11.	Chen et al. (2019)	273	0.41	0.306	0.504	7.158	<0.001	6.15	8.05
12.	Shahrokhi and Dehaghani (2025)	343	0.54	0.460	0.611	11.140	<0.001	7.74	8.22
13.	Wahidiyati et al. (2025)	85	0.007	-0.206	0.220	0.063	>0.001	1.87	6.53
Fixed			0.436	0.412	0.460	30.965	<0.001		
Random			0.449	0.361	0.530	8.911	<0.001		

Note: N: Number of sample; r: correlation; Z: Z-value; p: p-value; W: Relative Weight; F: Fixed model; R: Random model

Table 1 presents key findings from the 13 studies using a cross-sectional design. In this respect, the range of individual studies' correlations spans from the lowest, with r=0.007, N=85, p=0.949, derived from Wahidiyati et al. (2025), to the highest, with r=0.699, N=406, p=0.000, derived from Zhao et al. (2025). Additionally, all Z-values were positive, and all p-values were 0.000, except for one study. Moreover, regarding the effect model, particularly the fixed one produced a point estimate of r=0.436 (95% CI: 0.412-0.460, Z=30.965, p=0.000), while the random one yielded r=0.449 (95% CI: 0.361-0.530, Z=8.911, p=0.000). Similarly, based on the visual inspection of the Forest Plots in Picture 7, the 13 studies are listed vertically. The graphical representations show that most studies, in both the fixed-effect and random-effects models, were on the right side, indicating positive relationships. Besides, the plots also reveal non-overlapping CIs among the studies, indicating variability. Lastly, there was no study with CIs including 0, except for one study, which indicates that most studies suggest positive associations between motivation and willingness to communicate.



Picture 7. Forest Plots of Motivation—WTC

Table 3. Heterogeneity of Motivation—WTC

Heterogeneity				τ^2			
Q	**df (Q)**	**P**	**I^2**	**τ^2**	**Std Err**	**Variance**	**τ**
140.365	12	0.000	91.451	0.033	0.017	0.000	0.182

Furthermore, Table 3 presents statistical results regarding the heterogeneity of the motivation and WTC datasets. The statistical outputs show high heterogeneity, as indicated by Cochran's Q statistic (12)=140.365, p<0.001, I2=91.5% (rounded). These values reject homogeneity strongly (far below α = 0.05). They also provide a strong justification for the random-effect model as a suitable model. These statistical outputs

show substantial variability across studies (Higgins et al., 2003), which may be due to factors such as cultural contexts, populations, and instruments used.

Additionally, the among-study variance, $\tau 2 = 0.033$, with a standard error of 0.017, indicates modest absolute heterogeneity. The $\tau=0.182$ indicates the standard deviation of true effects on the correlation scale; thus, the true r values vary by approximately 0.182 around the pooled mean (r=0.449). This statistic provides a useful way to predict new study ranges using prediction intervals.

Following the heterogeneity test above, other statistical analyses were employed to justify the results. In this respect, since CMA version 3.3.070 did not provide the feature for estimating prediction interval statistics, the analysis moved to R + metafor software, resulting in Table 4 below.

Table 4. R Output Concerning Interval Prediction Interval of Motivation-WTC.

Statistic	Value
pred	0.4496
ci.lb	0.3557
ci.ub	0.5345
pi.lb	0.0900
pi.ub	0.7055

As shown by Table 4, although the pooled random-effects estimate shows a positive relationship between L2 motivation and WTC (r = .450, 95% CI [.356, .535]), the 95% prediction interval ranged between r = .090 and r = .706, indicating variability in light of the effects across future studies. This would suggest that the magnitude of the relationship may vary considerably across different EFL contexts, e.g., due to participant-related and methodological differences. Therefore, further inspections, including leave-one-out analysis, sensitivity analysis, and subgroup analysis, are also suggested. To begin with, the R + metafor outputs concerning the leave-one-out analysis are illustrated in Table 5 below.

Table 5. Leave-One-Out Analysis of L2 Motivation and WTC

Study omitted	Pooled effect (*r*)	95% CI	I^2 (%)
Aladini et al. (2024)	0.446	[0.343, 0.539]	93.24
Fallah (2014)	0.459	[0.359, 0.548]	92.89
Joe et al. (2017)	0.453	[0.350, 0.545]	92.94
Mulyono and Saskia (2020)	0.446	[0.342, 0.539]	92.80
Wu and Lin (2014)	0.437	[0.338, 0.526]	92.73
Zhao et al. (2025)	0.420	[0.341, 0.494]	87.29
Zhou (2025)	0.452	[0.348, 0.544]	92.38
Khajavy et al. (2019)	0.453	[0.350, 0.545]	93.05

 DOI: 10.62079/abjad.v4i2.146

Nikitina et al. (2022)	0.467	[0.373, 0.551]	90.81
Borkowska (2022)	0.440	[0.343, 0.528]	92.92
Chen et al. (2019)	0.453	[0.351, 0.545]	93.12
Shahrokhi and Dehaghani (2025)	0.441	[0.339, 0.533]	92.76
Wahidiyati et al. (2025)	0.474	[0.395, 0.546]	89.70
Overall model	**0.450**	**[0.356, 0.535]**	**92.34**

Note: r=correlation coefficient, CI=Confidence Interval, I^2=percentage of total variability attributable to between-study heterogeneity

As noted in Table 5 above, the leave-one-out analysis indicates that the pooled correlation remained relatively stable after sequentially removing each study. The recalculated effect sizes range from r=0.420 to r=0.474, compared with the overall pooled effect of r=0.450. Moreover, heterogeneity remains substantial across all iterations (I2=87.29% to 93.24%), showing that no single study is responsible for the observed variability. These statistical findings confirm the robustness of the overall association between L2 motivation and WTC.

In addition to the leave-one-out analysis, a sensitivity analysis comparing fixed-effect and random-effects models was conducted to evaluate the stability of the pooled effect estimate. In this respect, as shown by Table 1, the results indicate that the estimated relationship between L2 motivation and WTC remained highly consistent across analytical models, with an r of 0.436 (95% CI [0.412, 0.460]) for the fixed-effect model and an r of 0.450 (95% CI [0.356, 0.535]) for the random-effects model. Given the minimal difference, the overall conclusion was not substantially affected by the choice of estimation model. Following the sensitivity analysis above, influence diagnostics were conducted using studentised residuals, DFFITS, Cook's distance, and changes in heterogeneity estimates. The results are shown in Table 6 below.

Table 6. Influence Diagnostics for Motivation and WTC

Study	Studentized residual	DFFITS	Cook's distance	Interpretation
Zhao et al. (2025)	2.568	0.811	0.401	Higher influence
Wahidiyati et al. (2025)	-2.546	-0.626	0.297	Higher influence
Nikitina et al. (2022)	-1.269	-0.389	0.141	Moderate influence
Other studies	<	<	<	Limited influence

Note: DFFITS=Difference in Fits; Other studies are provided in Appendix 2

The results of influence diagnostics shown in Table 6 above indicate that two studies demonstrate a stronger influence than the others. In this respect, the strongest positive influence derives from Zhao et al. (2025), whereas the largest negative influence derives from Wahidiyati et al (2025), indicated by their DFFITS values of 0.811 and -0.626, respectively. Following the influence diagnostics above, another inspection concerns subgroup analysis. In this respect, subgroup analysis was conducted based on the geographical context; specifically, macro-cultural regions were employed in the present study. The distribution of subgroups is presented in Table 7 below.

Table 7. Macro Cultural Regions of the Studies Concerning Motivation and WTC

Region	**k**	**Studies**
East Asia	6	Joe et al. (2017), Wu & Lin (2014), Zhao et al. (2025), Zhou (2025), Nikitina et al. (2022), and Chen et al. (2019)
Middle East	4	Aladini et al. (2024), Fallah (2014), Khajavy et al. (2019), and Shahrokhi & Dehaghani (2025)
Southeast Asia	2	Mulyono & Saskia (2020) and Wahidiyati et al. (2025)
Europe	1	Borkowska (2022)

Note: k=number of studies

As shown in Table 7 above, the subgroup sizes are unequal. In this respect, Europe has only one study, Southeast Asia has only two, the Middle East has 4, and East Asia has 6. Therefore, conducting a four-group analysis will have very low statistical power. Accordingly, this would suggest that the subgroup analysis include two broader cultural clusters: East Asia and Other EFL contexts (if the clusters are categorised into Asian and European EFL contexts, the analysis would not be ideal for the European region, as it would report only 1 study). The results are shown in Table 8 below.

Table 8. Subgroup Analysis of Motivation and WTC by Geographical/Cultural Context

Geographical/Cultural Context	**k**	**Pooled Effect (*r*)**
East Asia	6	0.477
Other EFL Contexts	7	0.424
Difference between subgroups		QM(1)=0.315, p=0.575; R^2= 0.00%

Notes: QM constitutes the test of moderators using a mixed effects model in which the reference category is East Asia

 DOI: 10.62079/abjad.v4i2.146

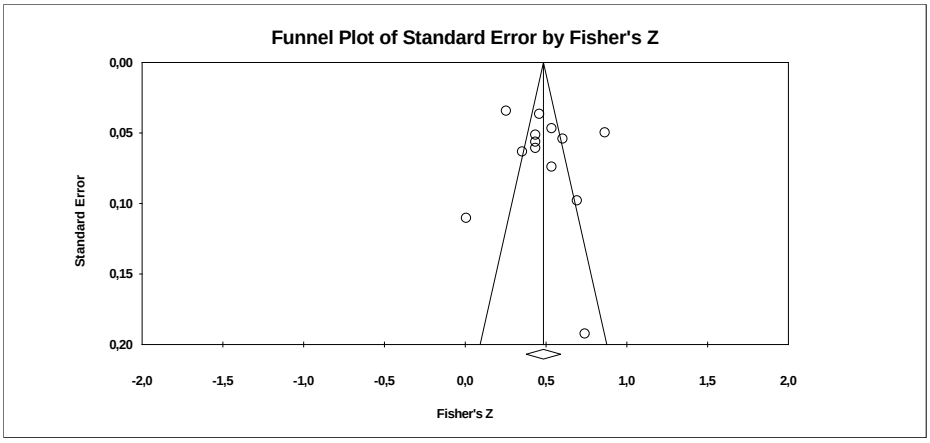
As illustrated in Table 8 above, the subgroup analysis was conducted based on two geographical/cultural contexts, i.e., East Asia and Other EFL Contexts. In this respect, the pooled effect sizes for both contexts are r=0.477 and 0.424, respectively. Albeit the effect size for East Asia is slightly larger than that of Other EFL contexts, the difference between these two groups was not statistically significant, as indicated by QM(1) = 0.315, p = 0.575. Moreover, with an R2 value of 0.00%, the geographical/cultural context explained 0.00% of the observed between-study heterogeneity. The findings suggest that the magnitude of the relationship between motivation and WTC was consistent across different geographical/cultural contexts. Accordingly, the regional classification alone was not sufficient to explain the substantial variability among the reviewed studies. Given the substantial heterogeneity reported in the overall meta-analysis, an additional subgroup analysis was conducted. In this respect, an additional subgroup analysis was conducted by educational level, i.e., higher education/adult learners and secondary education learners. The results are presented in Table 9 below.

Table 9. Subgroup Analysis of Motivation and WTC by Education Level

Educational level	**k**	**Pooled effect (*r*)**
Higher education/adult learners	10	0.482
Secondary education learners	3	0.336
Difference between groups		QM(1)=1.699, *p*=0.192; R^2=1.25%

As reported in Table 9, higher-education/adult learners have a larger pooled effect size than secondary-education learners, with r=0.482 and r=0.336, respectively. However, the difference between these groups was found to be insignificant, as indicated by the moderator test QM(1) = 1.699, p = 0.192. Additionally, the educational levels accounted for 1.25% of the observed heterogeneity, as indicated by R2=1.25%. These statistics indicate that educational level makes a minimal contribution to explaining the variability in effect sizes. Accordingly, although the relationship between motivation and WTC was found to be stronger at the higher education level than at the secondary level, the educational level alone was regarded as insufficient to explain the substantial heterogeneity across the included studies.

Nevertheless, it should be noted that this result may be due to the limited number of studies (k=13), unequal subgroup distribution (10 higher education and 3 secondary studies), and variation within educational contexts (for instance, university majors, proficiency levels, and instructional settings). Following the subgroup analysis, another inspection concerns publication bias. The results are shown in the Funnel Plot presented in Picture 8 and in Table 10 below.



Picture 8. Funnel Plot of Motivation—WTC

The visual inspection in Picture 8 shows a relatively symmetrical inverted-funnel distribution, with the effect sizes appearing to be distributed evenly around the pooled estimate. This visual inspection is consistent with the results from Egger's regression test in Table 10, which indicates an insignificant result (p=0.576). Accordingly, the current meta-analysis did not find clear evidence concerning the small-study effects or funnel plot asymmetry. However, the findings should be interpreted with caution, as substantial between-study heterogeneity may reduce the sensitivity of funnel plot-based methods and Egger's test in detecting potential publication bias.

Table 10. Egger's Regression of Motivation—WTC.

Egger Regression Intercept	**Statistic**
Intercept	1.592
Std error	2.76
95% lower limit (2-tailed)	-4.482
95% upper limit (2-tailed)	7.666
t-value	0.577
df	11
P-value (1-tailed)	0.288
P-value (2-tailed)	0.576

Statistical Findings of ICC/IS – Motivation (Cross Sectional Studies)

Similar to the previous scrutiny of the constructs using a cross-sectional study design, the following section discusses the meta-analysis outputs of intercultural communicative competence/intercultural sensitivity and motivation. It includes an overview of statistical outputs, including effect size estimates, heterogeneity, and publication bias.

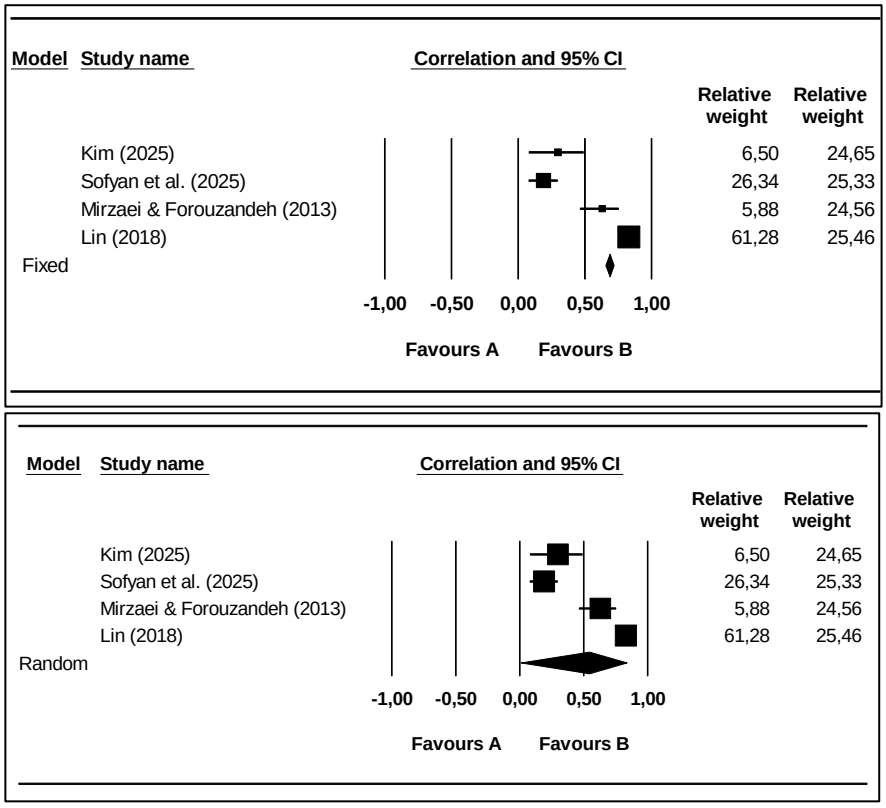
Table 11. Statistics Outputs of IC/IS—Motivation

 DOI: 10.62079/abjad.v4i2.146

No.	Studies	N	R	Lower Limit	Upper Limit	Z	p	W (F)	W (R)
1.	Kim (2025)	77	0.299	0.08	0.49	2,648	0,008	6.5	24.65
2.	Sofyan et al. (2025)	303	0.19	0.079	0.296	3,331	0,001	26.34	25.33
3.	Mirzaei and Forouzandeh (2013)	70	0.63	0.464	0.753	6,069	0.000	5.88	24.56
4.	Lin (2018)	701	0.83	0.805	0.852	31,39	0.000	61.28	25.46
Fixed			0.687	0.655	0.717	28.43	0.000		
Random			0.544	0.009	0.837	1.988	0.047		

N: Number of sample; r: correlation; Z: Z-value; p: p-value; W: Relative Weight; F: Fixed model; R: random model

The general overview of statistical findings on intercultural communicative competence/intercultural sensitivity with respect to motivation is presented through numerical and graphical representations. In this respect, the numerical representation is shown in Table 11. The correlation between the two variables across the 4 cross-sectional studies ranges from 0.19 to 0.83, with sample sizes ranging from 70 to 701 participants. The fixed-effect model generates a general correlation coefficient of r=0.687 (95% CI: 0.655-0.717, Z=28.43, p=0.000), while the random-effect model yields a general correlation coefficient of 0.544 (95% CI: 0.009-0.837, Z=1.988, p=0.047). Additionally, each study has CIs that exclude zero, indicating significance ($p<0.05$). Similarly, the forest plots reveal positive correlations among the four studies, as indicated by the black squares that lie to the right of the central vertical line (r=0.00). Besides, the black diamond, which belongs to the fixed-effect model, positions at r=0.687 (CI:0.655-0.717). The statistics showing the lower and upper limits reflect the homogeneity assumption. This result is different from the random-effect model, which indicates heterogeneity; in this respect, its black diamond is at r=0.544 (CI:0.009-0.837).



Picture 9. Forest Plots of IC/IS—Motivation.

Additionally, due to the limited number of studies included in the synthesis (i.e., only four) and the high heterogeneity, the interpretations of the current findings should be treated with caution. Accordingly, further studies involving more articles are required to provide precise estimates and to scrutinise potential sources of the between-study variability. Additionally, to ensure a suitable effect model, whether fixed or random, a heterogeneity test was conducted. The results are presented in Table 12 below.

Table 12. Heterogeneity of IC/IS-Motivation

Heterogeneity				Tau-squared			
Q	df (Q)	P	I^2	τ^2	Std Err	Variance	T
232,036	3	0.000	98.707	0.367	0.382	0.146	0.606

Based on the statistical outputs in Table 12, a judgment regarding the data's heterogeneity was made. In this respect, the Cochran's statistic Q=232.036 (df=3, p=0.000), I2=98.71% (rounded), and τ2= 0.367 indicate a high heterogeneity; thus, the suitable effect model is the random one. Additionally, the standard deviation of the true effects, $\tau = 0.606$, indicates that the true r values vary by approximately 0.606. These

findings were corroborated by the lowest r value reported by Sofyan et al. (2025), which was r=0.19, significantly different from Lin's (2018) and Mirzaei and Forouzandeh's (2013) studies, which yielded r=0.83 and r=0.63, respectively. Additionally, to assess whether publication bias emerged, a fail-safe N and a trim-and-fill test were utilised. The statistical outputs were shown in Tables 6 and 13.

Table 13. Fail-Safe N statistics of IC/IS-Motivation

Classic Fail-Safe N	**Statistic**
Z-value for observed studies	21.719
P-value for observed studies	0.000
Alpha	0.050
Tails	2
Z for alpha	1.960
Number of observed studies	4
Number of missing studies that would bring p-value to > alpha	488

Orwin's Fail-Safe N	**Statistic**
Correlation in observed studies	0.687
Criterion for a 'trivial' correlation	0.10
Mean correlation in missing studies	0.00
Number of missing studies needed to bring correlation 0,1	30

Based on the Fail-Safe N results shown in Table 13 above, several statistics were computed to assess publication bias. In this case, with N=488 indicating unpublished studies that are required to make a collective p-value above 0.05 (Z=21.719, p=0.000). In a similar vein, through Orwin's Fail-Safe N, the N=30 is required to reduce the combined correlation coefficient r=0.687 under the 'trivial' correlation of 0.10. Such findings indicate a robustness of the pooled positive correlation. Hence, it is very unlikely that publication bias will emerge.

Table 14. *Trim and Fill Test of IC/IS-Motivation*

Duval and Tweedie's Trim and Fill								
		Fixed Effects			**Random Effects**			
	Studies Trimmed	**Point Estimate**	**Lower Limit**	**Upper Limit**	**Point Estimate**	**Lower Limit**	**Upper Limit**	**Q**
Observed Values		0.687	0.655	0.717	0.544	0.009	0.837	232.035
Adjusted Values	2	0.817	0.800	0.833	0.796	0.424	0.424	666.048

 DOI: 10.62079/abjad.v4i2.146

Similar to the findings from Fail-Safe N, the Trim and Fill test, as shown in Table 7, indicates only a limited influence of potentially missing studies, resulting in an adjusted pooled effect size that remained positive. The findings indicate that the overall interplay between IC/IS and motivation was relatively robust after accounting for potential publication bias. Nevertheless, the interpretations of these results should be treated with caution because only four studies were involved; accordingly, with such a small number of studies, the Trim and Fill method has limited ability to detect or adjust for publication bias. Therefore, the adjusted estimates must be deemed exploratory rather than conclusive.

Statistical Findings of ICC/IS (Quasi-Experimental Studies)

The following statistical findings examine how a particular treatment influences intercultural communicative competence (ICC)/intercultural sensitivity (IS). The outputs derive from three studies utilising a quasi-experimental design. The outputs provide general views on the effect size estimates, along with other statistics used to identify heterogeneity and publication bias in the datasets. First, the general statistics for the variable are presented in Table 15 below.

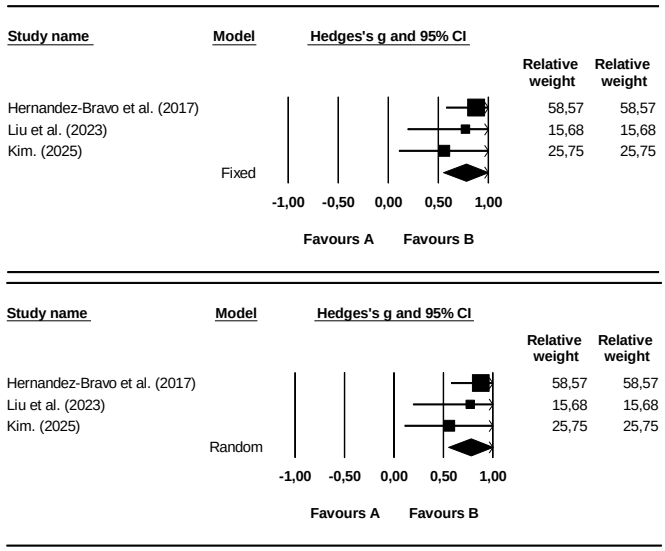
Table 15. Statistical Outputs of ICC/IS

No.	Study Names	N		Hedges's g	Std Error	Lower Limit	Upper Limit	Z	p	W (F)	W (R)
		Exp	Control								
1.	Hernández-Bravo et al. (2017)	94	93	0.878	0.153	0.579	1.177	5.757	0.000	58.57	58.57
2.	Liu et al. (2023)	25	23	0.772	0.295	0.194	1.350	2.617	0.009	15.68	15.68
3.	Kim (2025)	39	38	0.559	0.230	0.108	1.010	2.430	0.015	25.75	25.75
Fixed				0.779	0.117	0.551	1.008	6.675	0.000		
Random				0.779	0.117	0.551	1.008	6.675	0.000		

*)95% Confidence Interval

The statistical outputs in Table 15 above present findings on the effect size for each variable. All three studies report positive Hedges' g statistics, between 0.559 and 0.878. The statistics indicate that, in every case, the experimental group outperformed the control group on ICC/IS outcomes. In a similar vein, Picture 10 reveals that all three studies lie to the right of the central vertical line (g=0.00), indicating that the intervention derived from the experimental groups exceeded that of the control groups with regard to ICC/IS outcomes. Moreover, both the fixed-effect and random effect models yield the

same combined pooled effect sizes represented by Hedges' g=0.779 (95% CI: 0.551-1.008, p=0.000). Similarly, this finding was corroborated by the forest plot in Picture 10: the black diamond lies approximately at 0.779 and to the right of the central vertical line (stretching from 0.551 to 1.008), indicating that the pooled effect was significant. Overall, both statistical and graphical findings suggest that the interventions used have a moderate-to-strong effect on ICC/IS.



Picture 10. Forest Plots of IC/IS

Table 16. Heterogeneity of IC/IS

Heterogeneity				τ^2			
Q	**df (Q)**	**P**	**I^2**	**τ^2**	**Std Error**	**Variance**	**τ**
1.337	2	0.512	0.000	0.000	0.048	0.002	0.000

Moreover, to determine whether the datasets exhibit heterogeneity, a statistical test is used. In this respect, Cochrane's statistic Q=1.337 (df=2, p=0.512, I2=0.000, τ2=0.000) points out that the null of homogeneity is accepted; thus, there was no significant evidence of heterogeneity among studies. All the studies were found to be identical or similar in terms of their true effects. Thus, both fixed-effects and random-effects models will yield nearly the same pooled estimate. Also, the statistics indicate that the collective variation is attributable to true between-study differences, meaning that all variation is due to sampling error only (Borenstein et al., 2021; Higgins et al., 2003). Thus, a fixed-effects model is more suitable than a random-effects model (Borenstein et al., 2021; Higgins et al., 2003, 2019).

Table 17. *Fail-safe N statistics of IC/IS*

Classic Fail-Safe N	**Statistic**
Z-value for observed studies	6.238
P-value for observed studies	0.000
Alpha	0.050
Tails	2
Z for alpha	1.960
Number of observed studies	3
Number of missing studies that would bring p-value to > alpha	28
Orwin's Fail-Safe N	**Statistic**
Hedges's g in observed studies	0.779
Criterion for a 'trivial' hedges's g	0.1
Mean hedges's in missing studies	0.000
Number of missing studies needed to bring hedges's g under 0.1	21

The publication bias of the ICC/IS was tested using Classic Fail-Safe N and Orwin's Fail-Safe N. First, based on Table 17 above, the Classic Fail-Safe N outputs yielded a Z-value of 6.238 (p=0.000). With such statistics, the endeavour to nullify the significant findings of the three observed studies ($p<0.05$) necessarily requires 28 missing studies, which would render the findings insignificant at alpha=0.05, Z=1.960. These statistics of Fail-Safe N surpass the rule of thumb (robustness threshold) of 5k +10=23 (here, k=3 observed studies) (Rosenthal, 1979). Similarly, Orwin's Fail-Safe N shows that there would be 21 missing studies required to change the pooled value of Hedges' g 0.779 to be under 0.100, as shown by a trivial Hedges' g. Such a large N (21 missing studies) indicates that the combined Hedges' g of the observed studies, even with only three studies included, is robust.

Table 18. *Trim and Fill Test of IC/IS*

Duval and Tweedie's Trim and Fill								
		Fixed Effects			**Random Effects**			**Q**
	Studies Trimmed	**Point Estimate**	**Lower Limit**	**Upper Limit**	**Point Estimate**	**Lower Limit**	**Upper Limit**	
Observed Values		0.779	0.551	1.008	0.779	0.551	1.008	1.337
Adjusted Values	0	0.779	0.551	1.008	0.779	0.551	1.008	1.337

Furthermore, another method used to assess publication bias was Trim and Fill. In this respect, Table 18 presents the statistical outputs, which show that no studies were

 DOI: 10.62079/abjad.v4i2.146

trimmed under both fixed-effect and random-effect models, with Q=1.337. This suggests there was no asymmetry in the pooled effect (0.779; 95% CI: 0.551-1.008). These statistics suggest a symmetric funnel plot, which gives strong evidence against publication bias; the positive pooled effect is significant without any corrections.

Statistical Findings of Motivation (Quasi-Experimental Studies)

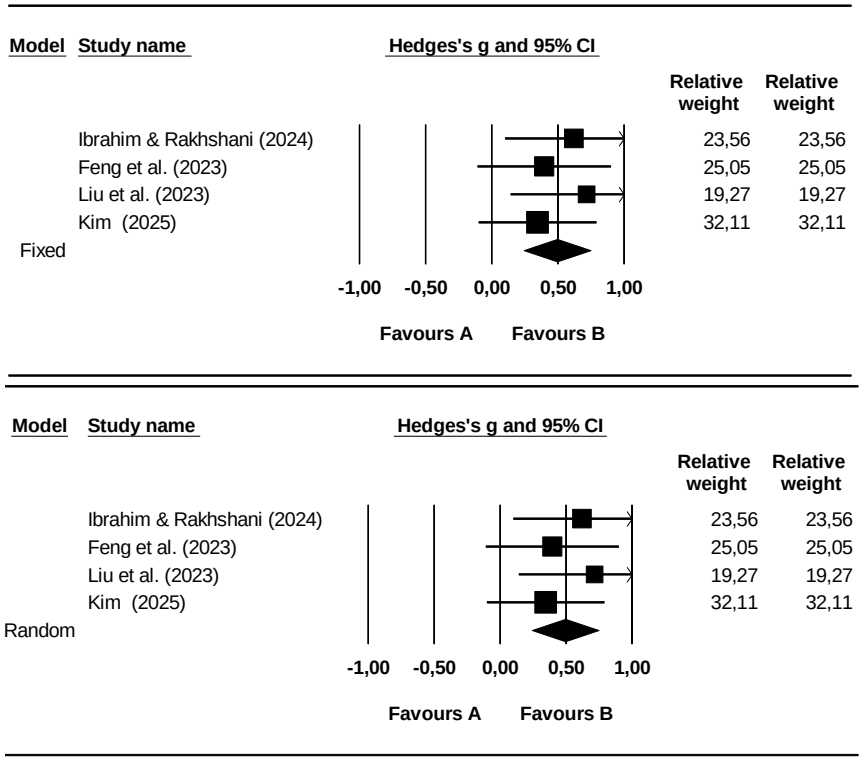
The following statistical findings examine how a particular intervention affects students' motivational outcomes. The outputs derive from four studies which employ a quasi-experimental design. The effect size estimates and the examinations of heterogeneity and publication bias in the datasets were reported. Table 19 below provides a general overview of the four studies.

Table 19. Meta-Analysis Output of Motivation

No.	Study Names	N		Hedges's g	Std Error	Lower Limit	Upper Limit	Z	p	W (F)	W (R)
		Exp	Control								
1.	Ibrahim and Rakhshani (2024)	29	29	0.621	0.265	0.101	1.141	2.339	0.019	23.56	23.56
2.	Feng et al. (2023)	30	30	0.396	0.257	-0.108	0.901	1.539	0.124	25.05	25.05
3.	Liu et al. (2023)	25	23	0.717	0.293	0.142	1.292	2.443	0.015	19.27	19.27
4.	Kim (2025)	39	38	0.346	0.227	-0.100	0.792	1.521	0.128	32.11	32.11
Fixed				0.495	0.129	0.242	0.747	3.840	0.000		
Random				0.495	0.129	0.242	0.747	3.840	0.000		

*)95% Confidence Interval

Table 19 presents various statistics, including sample size, effect size, and other outputs from the datasets of the four studies. In this respect, the effect sizes ranged from 0.346 to 0.717, with standard errors ranging from 0.227 to 0.293. Two studies report significant effects with p-values below 0.05 (Ibrahim & Rakhshani, 2024; Liu et al., 2023), while the other two studies (Feng et al., 2023; Kim, 2025) report insignificant results (p>0.05). The pooled effect size, both in the fixed and random-effect models, was indicated by Hedges' g of 0.495 with a standard error of 0.129 (95% CI: 0.242-0.747). Additionally, the combined Z-value is 3.840 with a p-value of 0.000, indicating significance.



Picture 11. Forest Plots of Motivation

The forest plots in Picture 11 corroborated the statistical outputs shown by Table 12 above. In this regard, both fixed- and random-effects models illustrate the same key visual features. The four studies have black squares, and the combined effect is indicated by the diamond lying to the right of the central vertical line (g=0). These illustrations suggest positive effects of the interventions on motivation. Nevertheless, there are two studies derived from Feng et al. (2023) and Kim (2025) with horizontal lines crossing the central vertical line, indicating that the effects of these two studies were not statistically significant at the 95% confidence level. To further investigate the Heterogeneity of the datasets, Heterogeneity was conducted. The result was shown in Table 20.

Table 20. Heterogeneity Test of Motivation

Heterogeneity				τ^2			
Q	**df (Q)**	**P**	**I^2**	**τ^2**	**Std Err**	**Variance**	**τ**
1.374	3	0.712	0.000	0.000	0.055	0.003	0.000

Table 20 above shows the statistical outputs to examine the heterogeneity of the datasets. In this respect, the heterogeneity test reveals a Cochrane Q statistic of Q=1.374 (df=3, p=0.712, I2=0.000, τ2=0.000), which suggests homogeneity among the included studies. It indicates that the effects of these four studies were similar or identical. To assess

 DOI: 10.62079/abjad.v4i2.146

publication bias, the Fail-Safe N and Trim-and-Fill tests were conducted. The results were shown in Tables 21 and 22 below.

Table 21. Fail-safe N statistics of Motivation

Classic Fail-Safe N	**Statistic**
Z-value for observed studies	3.921
P-value for observed studies	0.00009
Alpha	0.05
Tails	2
Z for alpha	1.960
Number of observed studies	4
Number of missing studies that would bring p-value to > alpha	13
Orwin's Fail-Safe N	**Statistic**
Hedges's g in observed studies	0.495
Criterion for a 'trivial' Hedges's g	0.10
Mean hedges's in missing studies	0.00
Number of missing studies needed to bring Hedges's g under 0.1	16

The statistical outputs shown in Table 21 were obtained to scrutinise the Classic fail-safe N and Orwin's fail-safe N. In this respect, the statistics of Classic Fail-Safe N reports a Z-value of 3.921 ($p=0.00009$). Additionally, to nullify the findings ($p<0.05$), there would need to be 13 missing studies for the p-value to exceed alpha (0.05). Nevertheless, the N value obtained does not exceed the robustness threshold of $5k+10=30$ (here, $k=4$ observed studies), indicating that the findings are likely to be susceptible to publication bias. Notwithstanding, Orwin's fail-safe N reports that substantial missing studies (i.e. 16 studies) with a trivial effect were required to make Hedges' g under 0.1. This statistic indicates that the results have a relatively stable effect and a low likelihood of publication bias (Orwin, 1983). To provide another strong justification, the Trim and Fill test was employed. The results were shown in Table 22 below.

Table 22. Trim and Fill Test of Motivation

Duval and Tweedie's Trim and Fill								
		Fixed Effects			**Random Effects**			**Q Value**
	Studies Trimmed	**Point Estimate**	**Lower Limit**	**Upper Limit**	**Point Estimate**	**Lower Limit**	**Upper Limit**	
Observed Values		0.495	0.242	0.747	0.495	0.242	0.747	1.374
Adjusted Values	1	0.442	0.211	0.673	0.442	0.211	0.673	2.423

 DOI: 10.62079/abjad.v4i2.146

The outputs of the Trim and Fill test, shown in Table 22, found one missing study that might contribute to asymmetry in the data distribution. In this respect, with the adjusted value included, there was a slight decrease from 0.495 to 0.442, along with a narrowing of the 95% confidence interval (from 0.242-0.747 to 0.211-0.673), in both fixed- and random-effects models. However, the effect size remains positive, indicating that the overall findings are still relatively stable.

Statistical Findings of WTC (Quasi-Experimental Studies)

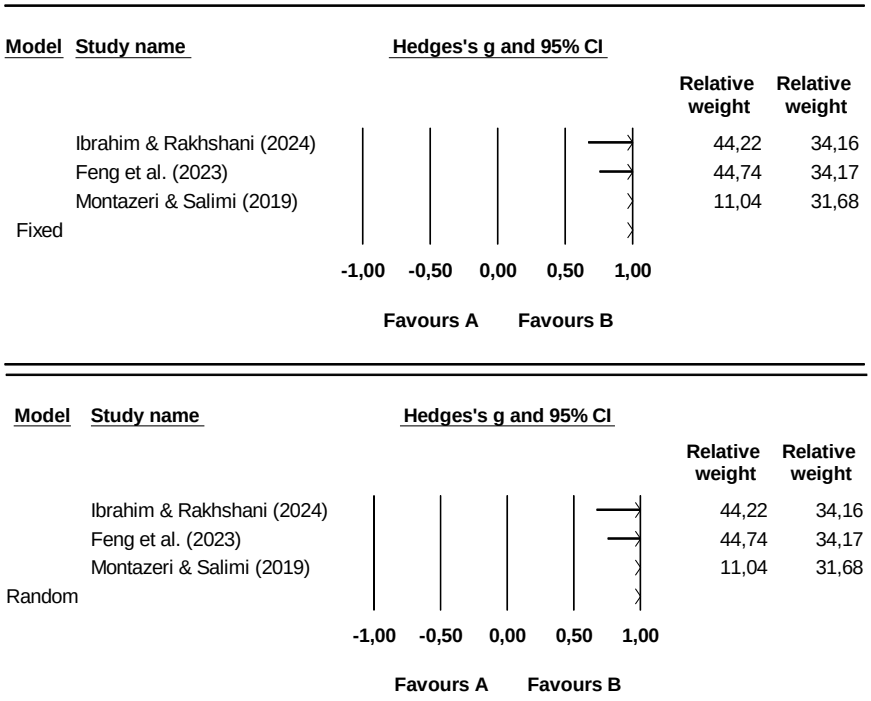
This section presents detailed statistical results regarding Willingness to Communicate. Three studies were included to estimate the pooled effect size, assess heterogeneity, and evaluate publication bias. First, the general overview of the results, presented statistically in Table 23 and corroborated graphically in Picture 12 below.

Table 23. Statistical Output of WTC

No.	Study Names	N		Hedges's g	Std Error	Lower Limit	Upper Limit	Z	p	W (F)	W (R)
		Exp	Control								
1.	Ibrahim and Rakhshani (2024)	29	29	1.225	0.283	0.670	1.779	4.329	0.000	44.22	34.16
2.	Feng et al. (2023)	30	30	1.305	0.281	0.754	1.856	4.639	0.000	44.74	34.17
3.	Montazeri and Salimi (2019)	31	30	5.597	0.566	4.487	6.707	9.884	0.000	11.04	31.68
Fixed				1.743	0.188	1.375	2.112	9.265	0.000		
Random				2.637	1.025	0.628	4.646	2.573	0.010		

*)95% Confidence Interval

 DOI: 10.62079/abjad.v4i2.146



Picture 12. Forest Plots of WTC

Table 23 presents general statistics on the effect sizes from the three studies on Willingness to Communicate. The sample size ranged from 29 to 31 subjects in the experimental groups and from 29 to 30 subjects in the control groups. The individual Hedges' g values for the three studies were 1.225, 1.305, and 5.597, derived from Ibrahim and Rakhshani (2024), Feng et al. (2023), and Montazeri and Salimi (2019), respectively. These statistics indicate varying magnitudes of effects among the studies, with the 95% confidence interval remaining above zero, namely ranging from 0.670 to 4.487, as indicated and corroborated by the horizontal lines in Picture 12. Additionally, the values of the combined fixed-effect and random-effect models are 1.743 and 2.637, respectively. The reported 95% confidence interval statistics for both models did not cross zero, and the Z-values for both models remained statistically significant against the Z critical values ($Z > 1.96$, $p < 0.05$) (Field, 2024). Next, to determine whether the included studies support heterogeneity, a heterogeneity test was conducted. The results obtained were shown in Table 24 below.

Table 24. Heterogeneity Test of WTC

Heterogeneity				τ^2			
Q	**df (Q)**	**P**	**I^2**	**τ^2**	**Std Err**	**Variance**	**τ**
52.099	2	0.000	96.161	2.996	3.460	11.970	1.731

 DOI: 10.62079/abjad.v4i2.146

Table 24 presents the statistical outputs used to test the heterogeneity among the three studies on WTC. In this respect, with Cochrane's Q statistic of 52.099 and p-value of 0.000, the three studies show substantial variability. Similarly, the reported I-squared (96.161%) and tau-squared (2.996) provide stronger evidence of heterogeneity among the three studies. In addition to testing heterogeneity, an examination of the publication bias of the three studies was also conducted. Tables 25 and 26 present statistical outputs from the fail-safe N and the Tri and Fill Tests to assess publication bias.

Table 25. *Fail-safe N statistics of WTC*

Classic Fail-Safe N	**Statistic**
Z-value for observed studies	10.884
P-value for observed studies	0.000
Alpha	0.050
Tails	2
Z for alpha	1.960
Number of observed studies	3
Number of missing studies that would bring p-value to > alpha	90
Orwin's Fail-Safe N	**Statistic**
Hedges's g in observed studies	1.743
Criterion for a 'trivial' Hedges's g	0.10
Mean Hedges's in missing studies	0.000
Number of missing studies needed to bring hedges's g under 0.1	50

Based on Table 25 above, the statistics of classic fail-safe N and Orwin's fail-safe N report some statistics. In this respect, the classic fail-safe N suggests that, with a Z-value of 10.884, 90 missing studies were required to make the p-value equal to alpha (0.05). This statistic exceeds the robustness threshold of 5k+10=35 (here, k=3 observed studies), indicating no potential publication bias. In a similar vein, Orwin's fail-safe N reports 50 missing studies to reduce the overall Hedges' g of 1.743 under the 'trivial' Hedges' g of 0.10. The findings show the robustness of the pooled positive effects on WTC. Thus, it is very unlikely that a publication bias exists. To provide a stronger justification, the Trim-and-Fill test was utilised. The findings were reported in Table 26 below.

Table 26. Trim and Fill Test of WTC

Duval and Tweedie's Trim and Fill								
		Fixed Effects			Random Effects			Q Value
	Studies Trimmed	Point Estimate	Lower Limit	Upper Limit	Point Estimate	Lower Limit	Upper Limit	
Observed Values		1.743	1.375	2.112	2.637	0.628	4.646	52.099
Adjusted Values	0	1.743	1.375	2.112	2.637	0.628	4.646	52.099

Table 26 reports that the statistical outputs from the Trim and Fill analysis are identical for both observed and adjusted values under the fixed- and random-effects models. With a Q value of 52.099, the point estimates and confidence intervals remained stable following the adjustment. The findings indicate that the method did not detect any asymmetry in the effect sizes because no potentially missing studies were found on either side of the funnel plot. To sum up, there was no evidence of publication bias that would influence the pooled effect size estimates.

Discussion

The present study sought to review the 29 included studies to develop a collective understanding of the social psychology of communication, particularly in English-as-a-foreign-language contexts. By incorporating qualitative thematic synthesis and meta-analysis, the following section elaborates and explicates the answers to the three research questions posed. In this respect, the discussion section will be broken down into conceptualising the social psychology of communication in EFL (RQ 1), conceptual issues and research gaps (RQ 2), and the strength of relationships among the constructs (RQ 3). At the end of the section, explanations were provided in light of the present study's implications and limitations.

Conceptualising Social Psychology of Communication in EFL

The findings reveal that previous studies conceptualise the social psychology of communication in EFL as a multidimensional construct in which WTC, L2 motivation, and intercultural sensitivity function as closely related social and psychological constructs, even though the three constructs have frequently been scrutinised separately rather than investigated within a single framework. Of the 29 studies reviewed, WTC was consistently associated with several psychological and contextual factors. For

instance, Aladini et al. (2024), Joe et al. (2017), Zhao et al. (2025), and Zhou (2025) show that factors encompassing learners' satisfaction of psychological needs, teacher support, and positive classroom climate contribute to promoting WTC. These findings are in line with MacIntyre et al.'s (1998) model of WTC, which suggests that communication is influenced not only by immediate situational constructs but also by more enduring motivational and interpersonal constructs that operate across various layers of the communication process. In a similar vein, Fallah (2014), Mulyono and Saskia (2020), and Wu and Lin (2014) demonstrate that learners' self-confidence and anxiety have a significant effect on their communicative behaviour, indicating that affective variables constitute the key antecedents of WTC. In addition, the findings from the reviewed articles were consistent with Gardner's (1985) socio-educational model, which posits that L2 learning motivation is the outcome of learners' attitudes towards the learning environment and the target language community. In this respect, the positive classroom climate, positive teacher-student rapport, and learners' psychological engagement suggest that motivation is developed through interconnections among individual psychological dispositions and social and educational factors, rather than being influenced by a single factor, e.g., learners' personal characteristics. Moreover, because many of these studies scrutinised only one or two of the main constructs (IS, L2 motivation, or WTC), rather than examining the three constructs in an integrated manner, the findings of the present study are necessarily interpreted as indicating a complementary relationship or effects among them. These outcomes underscore the need for further studies to investigate more comprehensive theoretical frameworks that integrate all constructs into a single explanatory model.

Furthermore, motivation, as one of the foci discussed in the current review, was identified as a variable comprising an internal driving force and a socially mediated component. It was, in this respect, shaped by attitudes, beliefs, and contextual affordances (Khajavy et al., 2019; Nikitina et al., 2022; Wahidiyati et al., 2025). Nevertheless, it was found that the association between motivation varied. On the one hand, most of the reviewed studies report a positive relationship; on the other hand, Mohammadian (2013) found no significant correlation between these two variables. The findings reinforce the notion that communicative engagement might not rely solely on motivation but also on emotional and contextual conditions.

Moreover, another psychological social construct to ponder in communication is intercultural sensitivity. Intercultural sensitivity was considered to play a key role, particularly in communication across various pedagogical contexts. The studies conducted by Mirzaei and Forouzandeh (2013), Sofyan et al. (2025), and Chen et al. (2019) report that intercultural sensitivity promotes both motivation and WTC. Additionally, intercultural sensitivity and competence were reported to be effectively

developable through pedagogical and technological approaches, as indicated by intervention-based studies by Hernández-Bravo et al. (2017), Liu et al. (2023), and Kim (2025).

Lastly, the qualitative findings above are also partly supported by statistical findings obtained from meta-analysis. For instance, as reported by the random-effects models, the pooled correlation between motivation and WTC (r=0.449) indicates a moderate, statistically significant relationship. Moreover, a stronger relationship between intercultural sensitivity and motivation (r=0.544) was found. Overall, the findings justify viewing the social psychology of communication in EFL not as a set of independent variables but as a synergistic system.

Conceptual Issues and Research Gaps

Notwithstanding the growing body of literature, the present study reveals some key conceptual and methodological gaps. In this respect, the most prominent gap is that research which simultaneously scrutinises intercultural sensitivity, L2 motivation, and WTC remains limited. Previous studies have largely investigated these three constructs in isolation or in pairs. For instance, some studies mainly conducted investigations on intercultural sensitivity or related intercultural constructs (Hernández-Bravo et al., 2017; Kim, 2025; Liu et al., 2023; Mirzaei & Forouzandeh, 2013; Sugianto et al., 2022) whilst other studies emphasised the interplays between motivation and WTC (e.g., Aladini et al., 2024; Borkowska, 2022; Fallah, 2014; Joe et al., 2017; Khajavy et al., 2019; Mulyono & Saskia, 2020; Nikitina et al., 2022; Shahrokhi & Dehaghani, 2025; Wahidiyati et al., 2025; W.-C. V. Wu et al., 2025; Zhao et al., 2025; Zhou, 2025). Only a few studies, e.g., Balouchi and Samad (2021) and Chen et al. (2019), explicitly investigated associations among more than one of these domains; however, they did not provide a comprehensive theoretical framework that simultaneously encompasses the three constructs. The current findings remain fragmented, which limits understanding of how social and psychological variables jointly shape communication in the EFL context. Such a gap provides a strong rationale for the present review, which constructs a social psychology framework by synthesising evidence on intercultural sensitivity, L2 motivation, and WTC within an integrated conceptual view.

Second, inconsistency in light of the role and interplay of motivation still emerged. There were some studies (e.g., Borkowska, 2022; Wu & Lin, 2014; Zhao et al., 2025) reporting positive, significant, and strong relationships between motivation and WTC, while others (Mohammadian, 2013; Wahidiyati et al., 2025) reported weak or non-significant effects. The findings indicate that, rather than being a universal predictor, motivation is necessarily a mediator or moderator.

Third, the other issues and research gaps to ponder concern the context in which the studies were conducted. The included studies were conducted in Asian contexts (26 of 29). Moreover, higher education levels constitute the most frequent contexts. Besides, digital settings constitute contexts that remain under-explored. Only a few studies have been conducted within the digital settings, such as the use of CALL-based multimodal pedagogy (Bashir et al., 2025), augmented reality (Liu et al., 2023), AI-assisted language assessment (Aladini et al., 2024), a digital platform such as Second Life virtual world (Kruk, 2019) and SNS (Kim, 2025). Therefore, further exploration and investigation of the respective constructs within these contexts remain warranted in future studies.

Lastly, another limitation concerns methodological constraints. In this regard, self-report and cross-sectional designs are the most common designs utilised (see Picture 3, Table 1 and Table 4); however, these designs have limitations for causal inference. Meanwhile, only a few studies employed other designs, such as intervention-based studies, which might indicate the malleability of the constructs (Feng et al., 2023; Hernández-Bravo et al., 2017; Ibrahim & Rakhshani, 2024; Kim, 2025; Liu et al., 2023; Montazeri & Salimi, 2019) and a qualitative design (Eddy-U, 2015), which might yield insights into dynamic processes, albeit limitations emerge in terms of the generalizability. All the aforementioned gaps indicate longitudinal, integrative, and context-sensitive research agendas that might provide comprehensive portraits of the social psychology of communication in EFL and are in demand.

Magnitude of the Relationships among the Constructs

The magnitude of the relationship was assessed using the effect sizes from the meta-analytic findings. The findings provide robust empirical support, alongside the aforementioned qualitative findings, given the associations among intercultural sensitivity, motivation, and WTC. In this regard, based on the cross-sectional studies reviewed, a moderate and significant relationship between motivation and WTC ($r=0.449$) was reported. It shows that motivation is a reliable predictor of communicative readiness, although variability across studies emerged, which might be due to contextual influences. Moreover, a stronger relationship ($r=0.544$) between intercultural sensitivity and motivation was reported. It indicates that intercultural orientation has a key role in driving motivation. In addition to the results of the cross-sectional studies, the findings of intervention-based studies suggest that WTC is a highly malleable variable, as indicated by a large effect size ($g=2.637$), while intercultural sensitivity and motivation were found to have moderate to large effects, indicated by $g=0.779$ and $g=0.495$, respectively. The findings above suggest that communication-related constructs are developable capacities rather than fixed traits, particularly through pedagogical interventions.

Implications and Limitations

Taken together, the findings aforementioned propose some key practical implications. To begin with, since learners' WTC consistently has a positive association with a supportive classroom and reduced language anxiety. Teachers are required to create psychologically safe learning spaces that promote learners' participation and self-confidence and provide meaningful opportunities for interaction (Borkowska, 2022; Fallah, 2014; Joe et al., 2017). Moreover, since intercultural sensitivity consistently has an association with WTC, the present study suggests that curriculum designers are necessarily required to infuse and incorporate intercultural learning experiences, authentic communicative tasks, and collaborative activities into English instruction to promote intercultural awareness and openness (Chen et al., 2019; Sofyan et al., 2025). Additionally, the positive effects of technology-mediated learning environments suggest that digital platforms and technology utilisation in language learning can be integrated to promote learners' motivation, intercultural engagement, and opportunities for communication (Aladini et al., 2024; Bashir et al., 2025; Liu et al., 2023). Lastly, there should be alignment between assessment practices and communicative and intercultural learning objectives, facilitated by formative, portfolio-based assessment and feedback-oriented approaches that encourage learners' reflective development and sustained engagement (Ibrahim & Rakhshani, 2024; Montazeri & Salimi, 2019; Shahrokhi & Dehaghani, 2025).

Notwithstanding the contributions above, it should be noted here that there are some limitations. First, the systematic review was limited to studies indexed in the Scopus database. Although Scopus is recognised as one of the most comprehensive databases for educational studies, limiting the inquiry to a single database may have excluded relevant studies indexed elsewhere, thereby reducing the comprehensiveness of the evidence base. Second, the inclusion of only English-language, open-access publications may introduce language and accessibility bias, potentially underrating high-quality studies in other languages with restricted publication access. Third, although the meta-analysis yields statistically pooled estimates, some analyses showed substantial between-study heterogeneity, indicating considerable variation in observed effect sizes across educational contexts, populations, instructional settings, and methodological designs. Accordingly, the pooled estimates must be interpreted as average effects rather than as universally applicable associations.

Furthermore, some meta-analytic models were based on a relatively small number of primary studies (*k*), which may reduce the precision and stability of the estimated effect sizes and limit opportunities for moderator analyses. Another limitation concerns

the conceptual diversity of the included studies. To facilitate synthesis, some closely related constructs (e.g. intercultural sensitivity, intercultural competence, intercultural communicative competence, and intercultural awareness) were grouped under broader conceptual categories. Although these constructs share substantial conceptual overlap, they are not theoretically identical, and conflating them may obscure important conceptual distinctions. Moreover, although publication bias was scrutinised using statistical analyses, the evaluations generally have limitations in statistical power because the number of included studies was relatively small. Therefore, the absence of statistically significant publication bias must be interpreted with caution rather than taken as definitive evidence that publication bias was not present. Additionally, most of the included studies used cross-sectional designs, which may limit causal inference and introduce methodological bias. In conclusion, the limitations above suggest that future systematic reviews are necessarily required to extend database coverage that includes the studies using multiple languages, scrutinise additional moderators to explain heterogeneity, differentiate related intercultural constructs, and integrate a larger body of primary studies to gain more generalisable meta-analytic findings and promote a more comprehensive framework with respect to social psychology of communication in EFL studies.

Conclusion

To sum up, the present study advances understanding of the social psychology of communication in EFL by synthesising evidence on intercultural sensitivity, L2 motivation, and WTC. Instead of viewing these constructs as a single, integrated model, the studies reviewed demonstrate that the social and psychological constructs are closely related and collectively contribute to learners' communicative readiness, although they have been scrutinised separately. The qualitative synthesis reveals that the interaction of psychological, intercultural, and contextual factors influences communication. At the same time, the meta-analysis provides quantitative evidence of moderate to strong positive associations between L2 motivation and WTC ($r = 0.449$, 95% CI [0.361, 0.530], $p < .001$) and between intercultural sensitivity/intercultural competence and motivation ($r = 0.554$, 95% CI [0.009, 0.837]). Moreover, the experimental studies report positive effects on intercultural sensitivity/intercultural competence ($g = 0.779$, 95% CI [0.551, 1.008]), motivation ($g = 0.495$, 95% CI [0.242, 0.747]), and WTC ($g = 2.637$, 95% CI [0.628, 4.646]), indicating that the constructs are amenable to pedagogical intervention. Notwithstanding, the findings must be interpreted with caution due to the aforementioned limitations, e.g., limited generalizability from the small number of included studies and heterogeneity in some meta-analytic estimates.

The present study also extends socio-psychological views within second language acquisition research. In this respect, the synthesis proposes that intercultural sensitivity

serves as an important antecedent that facilitates learners' motivation and, ultimately, their WTC. In addition, the interplay among these variables is influenced by contextual factors, including classroom climate, teacher support, and technology-mediated learning environments. Moreover, from a pedagogical perspective, the experimental studies demonstrate that technology-enhanced learning environments that incorporate augmented reality, AI-assisted assessment, and digital communication platforms enable learners to enhance their motivation, intercultural engagement, and opportunities for authentic communication. Rather than merely instructional tools, the technologies serve as contextual mediators through which meaningful interaction and support advance communicative competence in EFL contexts.

Finally, future studies are necessarily required to employ longitudinal designs and structural equation modelling to scrutinise the direct, indirect, and reciprocal relationships or effects among the constructs concerned in the present study. Moreover, future studies are also required to include underrepresented geographical regions, diverse populations and educational levels, and interventions involving technology, language proficiency, or achievement, which will yield a more comprehensive understanding of the mechanisms supporting the social psychology of communication in EFL.

Disclosure statement

The authors of the article have no conflict of interest in research or the review.

References

Aladini, A., Mahmud, R., & Ali, A. A. H. (2024). The importance of needs satisfaction, teacher support, and L2 learning experience in Intelligent Computer-Assisted Language Assessment (ICALA): a probe into the state of willingness to communicate as well as academic motivation in EFL settings. *Language Testing in Asia*, *14*(1). https://doi.org/10.1186/s40468-024-00334-9

Balouchi, S., & Samad, A. A. (2021). The effect of perceived competence on second language communication frequency: the mediating roles of motivation, willingness to communicate, and international posture. *Education and Information Technologies*, *26*(5), 5917–5937. https://doi.org/10.1007/s10639-021-10579-z

Bashir, A., Aziz, A., Imran, M., & Almusharraf, N. (2025). Effect of CALL-based multimodal pedagogy on learner motivation and willingness to communicate in English: A study from university students' perspective. *Contemporary Educational Technology*, *17*(2), 1–18. https://doi.org/10.30935/cedtech/15961

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2014). *Comprehensive Meta-Analysis* (3.0). Biostat.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Wiley & Sons, Ltd.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2021). *Introduction to meta-analysis* (2nd ed.). Wiley & Sons, Ltd.

Borkowska, A. (2022). The relationship among age-advanced Learners' in-class willingness to communicate in English, intrinsic motivation, classroom environment and teacher immediacy-a pilot study. *Theory and Practice of Second Language Acquisition*, *8*(2), 1–25. https://doi.org/10.31261/TAPSLA.8989

Cao, G. (2022). Toward the favorable consequences of academic motivation and L2 enjoyment for students' willingness to communicate in the second language (L2WTC). *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.997566

Chabot, C., Stolte, C., & Hanrahan, P. (2026). *Tableu Public*. Salesforce, Inc. https://public.tableau.com/app

Chen, G. M., & Starosta, W. J. (2000). The development and validation of intercultural sensitivity scale. *Human Communication*, *3*(1), 2–14.

Chen, Y. L., Cheng, H. F., Tang, H. W., & Wang, C. (2019). The roles of english varieties and L2 motivation in English learners' willingness to communicate in the internationalization at home (IaH) context. *IRAL - International Review of Applied Linguistics in Language Teaching*, *59*(3), 293–314. https://doi.org/10.1515/iral-2017-0163

Cheung, M. W. L., & Vijayakumar, R. (2016). A guide to conducting a meta-analysis. *Neuropsychology Review*, *26*(2), 121–128. https://doi.org/10.1007/s11065-016-9319-z

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Duval, S., & Tweedie, R. (2000a). A nonparametric "Trim and Fill" method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association*, *95*(449), 89–98. https://doi.org/10.1080/01621459.2000.10473905

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

Eddy-U, M. (2015). Motivation for participation or non-participation in group tasks: A dynamic systems model of task-situated willingness to communicate. *System*, *50*, 43–55. https://doi.org/10.1016/j.system.2015.03.005

Fallah, N. (2014). Willingness to communicate in English, communication self-confidence, motivation, shyness and teacher immediacy among Iranian English-major undergraduates: A structural equation modeling approach. *Learning and Individual Differences*, *30*, 140–147. https://doi.org/10.1016/j.lindif.2013.12.006

Feng, R., Alsager, H. N., Azizi, Z., & Sarabani, L. (2023). Impact of mind-mapping technique on EFL learners' vocabulary recall and retention, learning motivation, and willingness to communicate. *Heliyon*, *9*(6). https://doi.org/10.1016/j.heliyon.2023.e16560

Field, A. (2024). *Discovering statistics using IBM SPSS statistics* (6th ed.). Sage.

Gardner, R. C. (1985). *Social psychology and second language learning: the role of attitudes and motivation*. Edward Arnold.

Hernández-Bravo, J. A., Cardona-Moltó, M. C., & Hernández-Bravo, J. R. (2017). Developing elementary school students' intercultural competence through teacher-

led tutoring action plans on intercultural education. *Intercultural Education*, *28*(1), 20–38. https://doi.org/10.1080/14675986.2017.1288985

Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (2019). *Cochrane handbook for systematic reviews of interventions* (2nd ed.). The Cochrane Collaboration and John Wiley & Sons Ltd.

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ*, *327*, 557–560.

Ibrahim, K. A. A., & Rakhshani, Y. (2024). Portfolio assessment among upper-intermediate learners: effects on learners' willingness to communicate, grit tendencies, and motivation. *Asian Pacific Journal of Second and Foreign Language Education*, *9*(36), 1–23. https://doi.org/10.1186/s40862-024-00257-1

Joanna Briggs Institute. (2026). *Critical appraisal tools*. https://jbi.global/critical-appraisal-tools

Joe, H. K., Hiver, P., & Al-Hoorie, A. H. (2017). Classroom social climate, self-determined motivation, willingness to communicate, and achievement: A study of structural relationships in instructed second language settings. *Learning and Individual Differences*, *53*, 133–144. https://doi.org/10.1016/j.lindif.2016.11.005

Khajavy, G. H., Makiabadi, H., & Navokhi, S. A. (2019). The role of psychological capital in language learners' willingness to communicate, motivation, and achievement. *Eurasian Journal of Applied Linguistics*, *5*(3), 495–513. https://doi.org/10.32601/ejal.651346

Kim, J. (2025). SNS-based intercultural contact in an EMI context: fostering EFL learners' motivation and intercultural competence. *Journal of Asia TEFL*, *22*(3), 557–575. https://doi.org/10.18823/asiatefl.2025.22.3.4.557

Kruk, M. (2019). Dynamicity of perceived willingness to communicate, motivation, boredom and anxiety in Second Life: the case of two advanced learners of English. *Computer Assisted Language Learning*, *35*(1–2), 190–216. https://doi.org/10.1080/09588221.2019.1677722

Lin, Y. T. (2018). Taiwanese EFL Learners' Willingness to Communicate in English in the Classroom: Impacts of Personality, Affect, Motivation, and Communication Confidence. *Asia-Pacific Education Researcher*, *28*(2), 101–113. https://doi.org/10.1007/s40299-018-0417-y

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. SAGE Publications, Inc.

Liu, S., Gao, S., & Ji, X. (2023). Beyond borders: exploring the impact of augmented reality on intercultural competence and L2 learning motivation in EFL learners. *Frontiers in Psychology*, *14*, 1–15. https://doi.org/10.3389/fpsyg.2023.1234905

MacIntyre, P. D., Dornyei, Z., Clément, R., & Noels, K. A. (1998). Conceptualizing willingness to communicate in a L2: a situational model of L2 confidence and affiliation. *Modern Language Journal*, *82*(4), 545–562. https://doi.org/10.1111/j.1540-4781.1998.tb05543.x

Mirzaei, A., & Forouzandeh, F. (2013). Relationship Between Intercultural Communicative Competence and L2-Learning Motivation of Iranian EFL Learners. *Journal of Intercultural Communication Research*, *42*(3), 300–318. https://doi.org/10.1080/17475759.2013.816867

Mohammadian, T. (2013). The effect of shyness on Iranian EFL learners' language learning motivation and willingness to communicate. *Theory and Practice in Language Studies*, *3*(11), 2036–2045. https://doi.org/10.4304/tpls.3.11.2036-2045

Montazeri, M., & Salimi, E. A. (2019). Assessing motivation to speak (MTS) and willingness to communicate through metalinguistic corrective feedback. *Learning and Motivation*, *68*, 101594. https://doi.org/10.1016/j.lmot.2019.101594

Mulyono, H., & Saskia, R. (2020). Dataset on the effects of self-confidence, motivation and anxiety on Indonesian students' willingness to communicate in face-to-face and digital settings. *Data in Brief*, *31*, 1–5. https://doi.org/10.1016/j.dib.2020.105774

Nikitina, L., Lan, G., & Woo, W. S. (2022). L2 motivation and willingness to communicate: a moderated mediation model of psychological shyness. *Linguistics Vanguard*, *8*(1), 225–235. https://doi.org/10.1515/lingvan-2021-0125

Orwin, R. G. (1983). A Fail-Safe N for effect size in meta-analysis. *Journal of Educational Statistics*, *8*(2), 157–159. https://doi.org/10.2307/1164923

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, *372*. https://doi.org/10.1136/bmj.n71

Patole, S. (2021). Systematic reviews, meta-analysis, and evidence-based medicine. In S. Patole (Ed.), *Principles and practice of systematic reviews and meta-analysis* (pp. 1–11). Springer.

Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychological Bulletin*, *86*(3), 638–641. https://doi.org/10.1037/0033-2909.86.3.638

Shahrokhi, M., & Dehaghani, F. J. (2025). Mediating role of students' self-assessment in the relationships among EFL students' L2 grit, motivation, and willingness to communicate. *BMC Psychology*, *13*(1052). https://doi.org/10.1186/s40359-025-03414-4

Sharma, P., & Scholler, J. (2026). *Napkin*. Napkin. https://www.napkin.ai/

Sofyan, S., Habibi, A., Attar, R. W., Hendra, R., Alhazmi, A. H., Guillén-Gámez, F. D., & Sembiring, D. A. E. P. (2025). The impact of global-intercultural competence and social capital on students' motivation to participate in exchange programs: A PLS-SEM approach. *Power and Education*, *0*(0), 1–19. https://doi.org/10.1177/17577438251331917

Sugianto, A., Kurniawan, E., & Sukyadi, D. (2022). Zooming in on the Indonesian EFL primary school students' intercultural sensitivity and their extroversion. *Studies in English Language and Education*, *9*(1), 258–277. https://doi.org/10.24815/siele.v9i1.21478

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor. *Journal of Statistical Software*, *36*(3), 1–48.

Wahidiyati, I., Sabiq, A. H., Handayani, M. U., & Sartika, E. (2025). Speak up! How attitudes, motivation, and beliefs shape EFL learners' willingness to communicate. *Issues in Language Studies*, *14*(1), 37–54. https://doi.org/10.33736/ils.9268.2025

Wahyudi, L. (2018). *Watase Uake*. Watase Uake Indonesia. https://www.watase.web.id/home/index.php

Wu, C. P., & Lin, H. J. (2014). Anxiety about speaking a foreign language as a mediator of the relation between motivation and willingness to communicate. *Perceptual and Motor Skills: Learning and Memory, 119*(3), 785–798. https://doi.org/10.2466/22.PMS.119c32z7

Wu, W.-C. V., Lin, R. P., & Darmawansah, D. (2025). Effects of augmented reality-supported experiential learning on low-achieving EFL students in rural elementary school. *Innovation in Language Learning and Teaching*. https://doi.org/10.1080/17501229.2025.2510522

Zhao, M., Li, T., Yu, Y., Sin, K. F., & Zheng, E. (2025). Metacognitive strategies and willingness to communicate in L2 learning: a mediation model of self-efficacy and intrinsic motivation. *SAGE Open*, *15*(2), 1–17. https://doi.org/10.1177/21582440251340830

Zhou, S. (2025). Empowering academic communication in EMI higher education: Understanding the role of social-emotional need satisfaction in motivation and willingness to communicate. *Journal of English for Academic Purposes*, *78*, 1–12. https://doi.org/10.1016/j.jeap.2025.101582